A Theory of Ellipsis

MARJORIE J. McSHANE

OXFORD UNIVERSITY PRESS

A Theory of Ellipsis

This page intentionally left blank

A Theory of Ellipsis

MARJORIE J. McSHANE

OXFORD
UNIVERSITY PRESS

2005

OXFORD
UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further
Oxford University's objective of excellence
in research, scholarship, and education.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press

Library of Congress Cataloging-in-Publication Data
McShane, Marjorie Joan, 1967–
A theory of ellipsis / Marjorie J. McShane.
p. cm.
Includes bibliographical references and index.
ISBN-13 978-0-19-517692-6
ISBN 0-19-517692-8
1. Grammar, Comparative and general—Ellipsis.
I. Title.
P291.3.M38 2005
415—dc22 2004043497

2 4 6 8 9 7 5 3 1

Printed in the United States of America
on acid-free paper

For Ruth McShane

This page intentionally left blank

PREFACE

This book proceeds from the assumption that progress in describing and processing ellipsis in any of the world's languages is best achieved using an extensible parameters-and-values approach. Such a parameter-oriented description should incorporate factors from many branches of linguistics and, when fully developed, should be explicit enough to permit a computer to interpret and generate elliptical utterances. Thus, in contrast to available descriptions of ellipsis, the question is not only what *can* be elided in a language but also in what circumstances a given category *would* be elided, including the factors that might render ellipsis mandatory or unacceptable and the effect of eliding one category on the ellipsis potential of others. Since seeing the big picture can facilitate work on any corner of it, this framework encompasses unexpressed elements at the syntactic, semantic, and morphological levels and bridges work on ellipsis with the larger study of reference.

In principle, all descriptive material on any language, presented within any branch of linguistics, could contribute to the development of this theory. But for this book exhaustive depth and breadth of coverage was not the main goal—though the coverage of phenomena in the chapters that follow is quite representative. Instead, the book concentrates on:

- developing a methodology for describing ellipsis,
- supporting that methodology and the description it licenses in terms of potential applications,
- creating an inventory of relevant phenomena (which, for this work, derives primarily from a study of English, Russian, and Polish but incorporates select phenomena from other languages),

- formulating sample algorithms and heuristics for the analysis and synthesis of elliptical expressions, and
- presenting samples of the type of close description that would result from the current and future research efforts.

The practical justification for grounding the approach in parameterization is that it will facilitate the application of descriptions, heuristics, algorithms, and even specific language-processing programs from one language to others—an important benefit for practical natural language processing applications.

If one had to name the field to which this work belongs, it might best be called descriptive computational linguistics. The fine grain size of description answers the rigorous demands involved in developing computer programs to manipulate natural language, and the sample algorithms show how one might convert descriptions into processing routines. However, the issues involved in turning such algorithms into programs, as well as the literature devoted to extant ellipsis-resolution programs, lie outside the scope of this book. Here the focus is on the *linguistics* of computational linguistics—the exhaustive description of how elliptical phenomena in language work and the framing of such descriptions in practically useful, rule-oriented terms. When using this book, natural language processing system implementors can select specific phenomena and affecting factors to include in any given implementation. The material presented in this book can be viewed as an important component of the descriptive foundation for a program of work on developing a comprehensive computational approach to the treatment of ellipsis and, more broadly, reference. The descriptions in the book will, however, not only benefit computational linguists but also be useful for theoretical linguists studying ellipsis, reference, and cross-modular language phenomena, field linguists pursuing descriptions of the grammar and the lexicon of a language, and anyone involved in learning or teaching a foreign language.

Like all work, this book reflects the input, academic and other, of people too numerous to list, all of whom I thank sincerely albeit implicitly. Here I would like to mention those people whose contributions directly affected the content of this book and the development of the thinking that led to it. Special thanks to Leonard Babby, Sophia Lubensky, and Charles Townsend, for their guidance and support in my days of Slavic linguistics; to Mary Besemeres and Ron Zacharski for the same during the writing of this manuscript; to the many speakers of Russian and Polish—particularly Sophia Lubensky and Katarzyna Hagemajer—who generously shared their time and intuitions; and to Sergei Nirenburg, for turning my thinking to systems, applications, and theories of another kind.

CONTENTS

ABBREVIATIONS

∅	Empty (elided) category; says nothing about the nonelliptical variant
( )	The category may optionally be elided.
no special marking	The category must be overt.
*	The sentence is ungrammatical.
?	The sentence is of questionable grammaticality.
⌋	The sentence is stylistically infelicitous but not ungrammatical.
⌋ ⌋	The sentence is stylistically highly marked but not ungrammatical.
NOM	Nominative case
ACC	Accusative case
GEN	Genitive case
DAT	Dative case
INSTR	Instrumental case
PREP	Prepositional case
NP	Noun phrase
VP	Verb phrase
PP	Prepositional phrase
$PP_{NOM, GEN, etc.}$	PP whose complement NP has NOM, GEN, and so on, case marking
SG	Singular
PL	Plural
COORD	Coordinating (for conjunctions)
CONTR	Contrastive (for conjunctions)

CONDIT	Conditional particle
NEG	Negative particle
IMPER	Imperative
IMPERS	Impersonal
DO	Direct Object
R	Russian
P	Polish
Cz	Czech

NOTES ABOUT EXAMPLES

Most of the examples used in this book are from real texts, meaning that they tend not to be as conveniently short as examples invented for illustrative purposes. Under these circumstances, full glossing of each element in non-English examples could quickly obfuscate the point, as could lining up each word with its transliteration and translation. Therefore, transliterations and word-for-word translations are typed as running text and only select elements are supplied with their immediately relevant grammatical information. For examples that are really long or that contain idioms too complex to be rendered word for word, the word-for-word translation is limited to the crucial bit, with ellipses surrounded by braces { . . . } to show that material has been left out. In most cases, Russian examples are presented in both Cyrillic and transliteration; however, when Russian examples are being directly compared with Polish and Czech equivalents, the Cyrillic is not provided.

Sources of examples are referenced variously: for literary works, by the author's last name (plus a distinguishing number if more than one of his or her works is cited) and the page number; for linguistic sources, by the author's last name, the year of the work, and the page number, if applicable (formatting has been changed for consistency in some cases); for print media, by the name of the newspaper or journal and the year; for examples from the Tübingen on-line Russian corpora (found at http://www.sfb441.uni-tuebingen.de/b1/en/korpora.html), by the author's first initial and last name; for examples invented by informants, no source is cited.

Some literary works are translations from other languages into Russian or Polish; all examples drawn from such works were judged by native speakers to represent standard usage. The translations are treated as original works, meaning that my

glosses of such examples are based on the translation, not on the original work, even if it was in English. Plays and more colloquial prose provide more examples of ellipsis, giving them priority as sources to be mined for examples.

The abbreviations used most commonly throughout the book are provided in the abbreviations list. Note that slash notation is used to show different judgments for elliptical and nonelliptical variants only in select instances. In general, the text that discusses each example provides information about variants not explicitly covered by the presentation of the example. Although the combination of an asterisk and parentheses (*) has been used by some to indicate that a category cannot be elided, I find this notation opaque and chose not to use it, instead leaving a category without any marking at all when it must be overt. However, in those cases in which examples from other sources use this notation, it is retained.

A Theory of Ellipsis

This page intentionally left blank

1

Getting Started

Syntactic ellipsis is the nonexpression of a word or phrase that is, nevertheless, expected to occupy a place in the syntactic structure of a sentence. For example, in *Mary **got** an A on the math test and Louise ∅ a B*, the verb 'got' in the second conjunct is elided.[1] Although for many linguists syntactic ellipsis has come to represent the default interpretation of the term "ellipsis," ellipsis is actually a much broader phenomenon whose many aspects vie for immediate attention, particularly in the realm of natural language processing (NLP). For example, semantic ellipsis—the nonexpression of elements that, while crucial for a full semantic interpretation, are not signaled by a syntactic gap—occurs in *I forgot my keys* and *He is reading Tolstoy*, since the meanings are actually "I forgot to take/bring my keys" and "He is reading a book written by Tolstoy." The fact that a component of semantics is, in fact, missing in sentences like these can be detected, among other ways, by cross-linguistic comparison: in Chinese, for example, one cannot use the elliptical *read an author*.

Ellipsis is a universal property of natural language, but its scope and means of realization differ substantially from language to language. Considering the ubiquity of this phenomenon, it may seem rather surprising that ellipsis studies are relatively undeveloped or, at least, lack breadth and depth of coverage. This state of affairs can be explained by the dizzying complexity of the phenomenon, which is not readily recognizable until one attempts to solve the ellipsis problem. Working on ellipsis—which requires reference to syntax, lexical semantics, discourse, prosody, semantics, and stylistics—is a prime example of doing linguistics across language modules.

The interaction between factors from different modules, while long of interest in linguistics, has not been a strong suit for linguistic theories, which—with the

exception of Optimality theory, recently expanded from phonology into morphology and syntax—tend to wander little outside their respective domains (phonology, syntax, semantics, lexicology). The raw material of theoretical linguistics, language description, also lags behind on cross-modular issues, with the lion's share of time and energy devoted to the traditional domains.

This is not to say that ellipsis has not been productively studied; it most definitely has. For example, syntacticians have delineated licensing and recoverability conditions for instances of ellipsis subject to a syntax-oriented treatment, discourse-theoreticians have incorporated ellipsis into models of theme-rheme structure and shared spaces of current concern between interlocutors, and descriptive and typological linguists—in addition to describing ellipsis potential in the world's languages—have used ellipsis as one of the diagnostics for the presence or absence of so-called discourse orientation in language.

All current treatments have one thing in common: they address those subtypes or aspects of ellipsis that are most salient for a given established framework, be it theoretical or descriptive. I have a different interest that correspondingly requires a new framework: painting the whole canvas of ellipsis at a single go, considering all the factors from all the realms that determine when a speaker of some language can, should, or even virtually must elide categories. Moreover, although research in the trenches must be oriented toward some finite number of languages (here those languages are Russian, Polish, and English, with other languages treated in connection with specific phenomena), the results should be generalizable, with parametric variation across languages. With these goals in mind, in this work I develop an extensible theory of ellipsis that:

1. takes a bold step toward the crucial but, as yet, not commonly sought goal of formally addressing ellipsis as a *cross-disciplinary* linguistic phenomenon;
2. presents all elliptical phenomena within a *single descriptive framework*, rather than limiting the scope of inquiry to those phenomena or even aspects of phenomena that neatly fall under the purview of some theory or modular approach;
3. incorporates data and phenomena from *many languages* and exploits literature from many domains, with the goal of ultimately creating a comprehensive inventory of cross-linguistic elliptical patterns;
4. develops a *parameter-and-value*-oriented description methodology that can be utilized to account for ellipsis potential in any natural language;
5. incorporates *syntactic*, *semantic*, and, as yet to a lesser degree, *morphological ellipsis* and relates all of these to the larger task of reference resolution;
6. is *independent of established schools* and thus, presumably, able to feed into many applications over a longer period of time than theoretically restricted approaches;
7. limits the complexity of the *metalanguage* of description yet provides means of extending this metalanguage dynamically in order to handle unforeseen phenomena;

8. *embraces useful partial solutions* if complete ones are currently out of reach, thus rejecting the all-or-nothing constraints imposed by many frameworks;
9. maximally *formalizes* the descriptions while writing in plain English, avoiding obfuscating jargon;
10. orients the description toward the goal of *NLP*, since doing so both imposes rigor on the description and in no way excludes its application to other realms; and
11. presents *sample processing algorithms* that show how the descriptive generalizations can be turned into useful rules.

1. Theory across language modules

In common usage, the word 'theory' tends to carry connotations of abstraction or hypothesis—and, indeed, in the realm of hard science this is the case. But for linguistic inquiry, the first definition of 'theory' in *Collins English Dictionary* seems a better fit: "a system of rules, procedures, and assumptions used to produce a result."[2] For example, although certain schools of theoretical syntax derive from the hypothesis that there exists an independent, innate language mechanism in the human brain devoted exclusively to syntax, most of the work done in such schools revolves around accounting for language phenomena using rules supported by notational devices. There is a similar lack of hypothetical or abstract quality to, say, lexicology, which is the theory of dictionary making that provides principles and guidelines for the more practical task of lexicography. In short, many existing theories of linguistics represent a means toward a formal description. With theory conceived of this way, it is hardly radical to propose a new theory to answer a new set of research goals.

However, not every formal description represents a theory, a statement that begs the question, what *is* a theory, after all? Fortunately, I am saved the labor of independently answering that question by adopting the definition of theory from Nirenburg and Raskin's "Prolegomena to the Philosophy of Linguistics" (2004, chapter 2). Such adoption, of course, can reflect neither the depth nor the well-argued rationale of the original work; however, even in bare form the tenets convey enough self-evident logic to act as a scaffolding for this theory of ellipsis.[3]

According to Nirenburg and Raskin, a theory can be defined as a combination of its function and its components. Its function must be to promote (1) selection of the best description methodology and (2) subsequent evaluation of the quality of the description, leading to iterative improvements in both methodology and description. The components of the theory—purview, premises, body, and justification—as well as methodologies to support theory building are discussed below as applied to this nascent theory of ellipsis.

1.1. The purview of the theory

The purview of this theory is all instances of syntactic, semantic, and morphological ellipsis in natural language that can be accounted for by rules of syntax, lexical and

compositional semantics, pragmatics, and combinations thereof.[4] By "rules" I mean statements subject to formalization in a way suited for machine processing in the foreseeable (non–science-fiction) future. This excludes, for example, unexpressed material in highly colloquial speech, in which grunts and pointing can count for a speech act or in which background information known only to the interlocutors themselves renders unexpressed material recoverable. Not every phenomenon to be discussed here *fully* falls within this purview: for example, an elliptical pattern that has clear-cut syntactic and lexico-semantic components might have an intangible pragmatic aspect that provides the only available explanation for differing grammaticality judgments in minimal pairs. However, since one goal of the theory is to describe factors from all realms that bear on ellipsis and clearly delineate between those that can currently be formalized and those that lie beyond reach such hybrid phenomena are quite in place.

Obviously, this purview is too vast to be fully treated in one book, especially when the goal of the theory is, ultimately, to ensure coverage of all elliptical phenomena in all natural languages. So, here I lay down the foundations for the theory (its conceptual structure and raison d'être), treat select phenomena in detail, then show how the same approaches can be applied to other relevant phenomena.

Needless to say, there is no single correct way to divide up and treat elliptical phenomena and, as such, there is no self-evident logic in the organization of its analysis. As a means of orientation into my approach, I summarize here the top-level classification of elliptical phenomena and the representative sample of subtypes to be discussed in this book (chapters and subsections are noted in square brackets as chapter.subsection). Although all the material in the book is required to fully develop and motivate this classification, this snapshot should act as an initial lay of the land, helping the details of the exposition to fall into place.

Syntactic ellipsis with coreference (*i.e., with an accessible syntactic antecedent*)

- Accusative object ellipsis with an Accusative object antecedent [3, 12]
- Accusative object ellipsis with a Nominative antecedent [4]
- Accusative object ellipsis with an oblique antecedent [5]
- Oblique object ellipsis with any syntactic antecedent [6]
- Head noun ellipsis [8]
- Gapping [9.1]
- Stripping [9.2]
- Sluicing [9.3]
- Verb Phrase Ellipsis [9.4]
- Multilicensor Verbal Ellipsis [10]
- The ellipsis of conjunctions and relative pronouns [11.1]
- The ellipsis of prepositions [11.2]
- The ellipsis of conditional particles [11.3, 12]
- The ellipsis of reciprocal and reflexive particles [11.4, 12]

Subject ellipsis [13.1.1]

Object ellipsis with an extralinguistic antecedent [13.1.2]

Nonfinite clauses [13.1.3]

Syntactic ellipsis without coreference (*i.e., with no syntactically accessible antecedent*)

The ellipsis of objects due to clause modality [7]

The ellipsis of objects with a generalized-human referent [7]

The ellipsis of objects in a series of actions [7]

Multilicensor Verbal Ellipsis [10]

The ellipsis of conjunctions [11.1]

The ellipsis of relative pronouns [11.1]

The ellipsis of subjects with a generalized-human referent [13.2]

Semantic ellipsis (*ellipsis of meaningful elements, but with no syntactic gap*)

Unexpressed agents in passives [13.3.1]

Agentive impersonals [13.3.2]

Unexpressed experiencers and possessors [13.3.3]

Unexpressed arguments in derived nominals [13.3.4]

Unexpressed morphemes

Haplology [13.4]

Morpheme Ellipsis [13.4]

Morpheme loss during incorporation [13.4]

Language strategies

Dialogue strategies [13.5.1]

Sentence fragments [13.5.2]

Nominal sentences [13.5.3]

Unagentive impersonals [13.5.4]

One aspect of this inventory that might strike native speakers of English as unexpected is the many references to object ellipsis. Since standard English does not employ object ellipsis, the largely Anglo-oriented literature likewise makes little reference to it. However, I will show that object ellipsis, like many of the other phenomena listed here but not employed in English, offers particularly rich insights into the difficult problems of ellipsis as well as material to stimulate the search for practical and integrated approaches to their resolution.

1.2. The premises of the theory

Premises determine which questions a theory should ask and what would count as a satisfactory answer. For example, given the existence of syntactic structures, we can look for gaps in them; given the existence of different types of phrases, we can specify them as elided or overt sentence constituents; given the notion that elided entities must refer to something in the linguistic or real-world context, we can seek the rules that determine how to recover that referent. All of these givens, drawn from mainstream linguistics, not only help to define what questions to ask and how to answer them but also provide a vocabulary for talking about language and a body of previous work to use as a jumping-off point. A sampling of questions to ask about ellipsis, the types of answers to seek, and some of the premises that underlie both follow (L stands for any language).

Question 1: Which elements can be left out in grammatical sentences of L?
Answer 1: A list of the syntactic categories in L that can be elided.
Premises: Language has syntax that determines grammatical versus ungrammatical sentences; ellipsis occurs in language; grammatical categories (N, NP, etc.) exist; and so on.

Question 2: What licenses the ellipsis of category *x* in the given configuration?
Answer 2: Some syntactic configuration, lexical item, pragmatic factor, and so on.
Premises: Syntactic, lexical, pragmatic, and other factors, in isolation or combination, license ellipsis in all languages.

Question 3: How is the referent for the elided category recovered (understood)?
Answer 3: The referent is syntactically accessible, understood pragmatically, or understood through world knowledge.
Premises: Elided categories require referents, which have various grammatical and real-world sources.

Question 4: In a syntactic configuration that in principle permits the ellipsis of category *x*, can anything block or make mandatory the ellipsis?
Answer 4: Factors and their interactions.
Premises: Many aspects of language affect ellipsis; ellipsis-supporting and ellipsis-impeding factors are often in competition.

Question 5: When given different explanations of an elliptical structure, which is best?
Answer 5: For application A, . . . ; for application B, . . .
Premises: Applications have available to them different sources of information, different expressive means for rule writing, and so on.

Although this theory supports descriptions that incorporate morphological, syntactic, lexico-semantic and pragmatic factors, any given application is expected to make only partial use of the full description.

1.3. The body of the theory and the description it supports

The body of a theory defines the format of the description that the theory licenses. For this theory of ellipsis, the body comprises: (1) an inventory of the parameters

that define elliptical configurations and their value sets, (2) sets of value combinations (configurations) that actually occur in language, and (3) the factors that affect ellipsis potential in each configuration. Tables 1.1 and 1.2 show a sampling of parameters and their value sets. Possible configurations and factors that affect them are best illustrated through examples, like the Russian ones that follow:

(1) a. Он придерживается **своих прежних взглядов** и менять **их** не собирается.
On priderživaetsja **svoix prežnix vzgljadov** i menjat' **ix** ne sobiraetsja.
he adheres-to **self's former views**$_{\text{GEN}}$ and change **them**$_{\text{ACC}}$ NEG intends
'He adheres to his former views and has no intention of changing them.'

b. Он придерживается **их** и менять (**их**) не собирается.
On priderživaetsja **ix** i menjat' (**ix**) ne sobiraetsja.
he adheres-to **them**$_{\text{GEN}}$ and change (**them**)$_{\text{ACC}}$ NEG intends
'He adheres to them and has no intention of changing them.'

(2) Она налила себе **молока** и выпила (**его**)**.**
Ona nalila sebe **moloka** i vypila (**ego**).
she poured self **milk**$_{\text{GEN}}$ and drank (**it**)$_{\text{ACC}}$
'She poured herself some milk and drank it.'

(3) a. Он отнесся к **этому делу** легкомысленно и отложил **его**.
On otnessja k **ètomu delu** legkomyslenno i otložil **ego**.
he related to **this matter**$_{\text{DAT}}$ frivolously and put-off **it**$_{\text{ACC}}$
'He didn't take this matter seriously and put it off.'

b. Он отнесся к **этому делу** легкомысленно и отложил (**его**) на завтра.
On otnessja k **ètomu delu** legkomyslenno i otložil (**ego**) na zavtra.
he related to **this matter**$_{\text{DAT}}$ frivolously and put-off (**it**)$_{\text{ACC}}$ until tomorrow
'He didn't take this matter seriously and put it off until tomorrow.'

TABLE 1.1 Parameters and values relevant for ellipsis on the whole

Parameters	*Values*
Elided category	Head verb
	Verb phrase
	Head noun
	Noun phrase functioning as a subject
	Noun phrase functioning as an object
	Conditional particle
	Etc.
Referent type	Syntactically accessible antecedent
	Syntactically accessible postcedent
	Referent understood from the discourse context
	Referent understood from real-world knowledge
	Generalized referent

TABLE 1.2 Parameters and values relevant for nominal ellipsis with a syntactically accessible antecedent

Parameters	*Values*
Case marking of the elided category	ACC, GEN, DAT, etc.
Case marking of the antecedent	ACC, GEN, DAT, etc.
Nature of the antecedent	Referential expression
	Pronoun
	Elided
Syntactic role of the antecedent	NP complement of a verb
	NP complement of a preposition
	Etc.
Semantic correspondence between the antecedent and the elided category	Identity
	Whole-part
	Generic-specific

The type and grain size of analysis supported by this theory is exemplified by the following considerations that apply to this set of examples. All of the preceding sentences contain VP-coordinate structures in which the antecedent occurs in the first conjunct and the potentially elided category in the second. In all instances, the antecedent is an object with non-Accusative case marking and the coreferring category is a direct object with Accusative case marking (which is the default for Russian direct objects). The contrast between (1a) and (1b) shows that, whereas a lexically case-marked Genitive *referential-expression* (*R-expression*) antecedent blocks ellipsis, a *pronominal* antecedent with the same properties permits it. The contrast between (1a) and (2) shows that whereas a *lexically* case-marked Genitive R-expression antecedent blocks ellipsis, a *configurationally* case-marked Genitive (i.e., Genitive of negation) R-expression permits it. The contrast between (3a) and (3b) shows that the addition of an adverbial modifier to the sentence can permit ellipsis in a configuration in which ellipsis would otherwise be blocked. Factors like these must be teased apart using minimal pairs of examples and similar methodologies if one is to gain a fundamental understanding of the workings of ellipsis in any language.

Taken to its natural conclusion, the body of this theory will be a template for the description of ellipsis in any language, in the same way as field linguistic surveys (e.g., Comrie and Smith 1977, Longacre 1964) are for language in general. That is, there will be an inventory of elliptical configurations and the factors that might affect them. In addition, as in field linguistics, unexpected data must always be expected and, when encountered, should be reduced to generalized principles and incorporated back into the theory.

This book, as it stands, can serve as motivated guidelines for describing, then processing elliptical phenomena in any language. Ultimately, the material could also be turned into a structured questionnaire, whose lines of questioning might look as follows (assume the person or interactive system asking the questions already has some information about L, like that it has case marking and that objects can be elided in some configurations; also assume that examples from L would be provided or concocted during the interactive knowledge elicitation):

Can a direct object with ACC case marking and a definite referent ever be elided? (Yes)
Can it be elided in VP-coordinate structures with an overt coordinating conjunction? (Yes)
. . . with a contrastive conjunction? (Yes)
. . . with no conjunction? (No)
Can the referent follow the elided category as a postcedent? (No)
Can the referent be case-marked NOM? (No)
. . . ACC? (Yes)
. . . GEN? (No)

The results of work on a new language could lead to several eventualities: (1) the suggested inventory of parameters, values, and their combinations covers all the known elliptical phenomena in that language; (2) the parameter inventory is sufficient, but additional values need to be supplied for certain parameters; (3) the parameter inventory needs to be supplemented; (4) the inventory of parameters and values is adequate, but some important combinations are not represented by examples; and so on. The first eventuality effectively justifies the theory. The others point out areas for improvement.

The *description*, which is arrived at by applying the body and methodology to a language, is the rules of ellipsis for that language, including all relevant factors from all branches of linguistics and their interactions.

1.4. The justification for the theory

The justification for a theory is that it is correct and sufficient—something that cannot be proven but only disproven by counterevidence, and usually only over time. The description and analysis in the chapters to follow will show that the theory covers much of what arguably needs to be covered to describe, then practically use or implement computer programs to process ellipsis in a subset of languages. Delayed justification will come when the theory is used by descriptive, theoretical, and computational linguists in the future.

1.5. The methodology employed in developing the theory

The methodology used during development of this theory includes:

1. culling relevant generalizations and analysis from the linguistic literature;
2. collecting elliptical examples from print sources, and from native speakers when minimal pairs are needed;
3. analyzing those examples and hypothesizing what factors might affect the employment of ellipsis, using linguistic parameters and their values as an orientation;
4. constructing, based on those hypotheses, minimal pairs or other means of teasing apart the influence of individual factors;
5. presenting such minimal pairs to native speakers for evaluation;

6. constructing generalizations and rules based on potentially competing factors; and
7. to the extent possible, explaining the rules and interactions, thus incorporating findings into the broader work of linguistic inquiry.

A crucial methodological contribution of the proposed research is that it will bring into play as much evidence concerning ellipsis phenomena as possible, irrespective of the source (e.g., language or world knowledge) or specific linguistic theory, approach, or level (syntax, semantics, discourse, etc.). This "integrational" research methodology promises a road that leads toward convergence of the earlier disjoint findings, thus revealing the full, multidimensional picture.

The importance of "real" (not invented) examples in this work cannot be overstated. The examples used as evidence for generalizations represent only a small fraction of the data collected and analyzed, most of which were extracted from print sources, not taped speech, since the orientation is on language devoid of performance lapses. When one is creating a full description of ellipsis for some language, however, many examples of each phenomenon should be recorded. There are three reasons for this. First, it is possible that the factors that initially appear to account for ellipsis in certain configurations actually do not or do so but only in part. Revising analysis can be done only on the basis of concrete evidence. Second, ellipsis-resolution programs must be tested, and the same examples that support a description of ellipsis can contribute to the test corpus. (By the way, examples of many types of ellipsis would be extremely difficult to find automatically, using typical corpus search techniques, since the categories of interest are not overt in the text and finding them would require a high-quality ellipsis detection program.) Third, since the analysis of ellipsis can be affected by the lexical items in question and their lexical specification, it is not superfluous to record and later test multiple examples that may, on the surface, look quite similar but, in fact, could harbor lexically oriented complexities.

2. Why develop a new theory of ellipsis?

Treatments of ellipsis already exist in reference grammars, modules of syntactic and pragmatic theories, and programs for the computational processing of language. This work has produced more than domain-specific results; it has generated a descriptive knowledge base about elliptical phenomena. However, this knowledge base must be pieced together from widely disparate sources whose purviews, formalisms, and ultimate goals have little in common. In addition, an incomplete picture of ellipsis is common, with descriptions tending to focus on what can, in principle, be elided but saying much less about practical rules of usage.

It is essential to break through this ceiling of results because understanding and generating elliptical constructions is critically important for descriptive and theoretical linguists, foreign language learners, and computer programs that aspire to "understand" and produce texts. For example, they all must know that *I could go out with him, but I don't want to* ∅ is acceptable, while **I wanted to go out with someone but didn't know who to* ∅ is not.[5] While many aspects of ellipsis in *English* have been

well described, the current state of the art in the study of ellipsis does not provide any comprehensive support for this activity or present a broad-coverage descriptive framework for treating ellipsis in more highly elliptical languages. In the grammars and textbooks of virtually all languages, descriptions of ellipsis are simplified and incomplete (if they exist at all), not foreseeing the pitfalls of real-world decisions about when and when not to elide.

It may be the perceived complexity of ellipsis that, at least in part, discourages deep and broad-reaching accounts. But if descriptive linguists had access to a guide to ellipsis that pointed out what to look for, suggested a methodology for collecting and analyzing examples, provided a framework for organizing phenomena, and advised how to work with informants, the coverage of ellipsis in grammars and textbooks might soon become much more adequate and concomitant with work in other spheres.

Ellipsis is not a peripheral language phenomenon that one can do without—at least not if one wants to analyze, speak, or computer-generate real language. Imagine how heavy language would be if we produced sentences like (4a) rather than (4b):

(4) a. I bought two chic designer dresses made of Chinese silk, my mother bought three chic designer dresses made of Chinese silk, and my grandmother bought four chic designer dresses made of Chinese silk.
b. I bought two **chic designer dresses made of Chinese silk**, my mother bought three Ø, and my grandmother bought four Ø.

Moreover, ellipsis is not an entity unto itself; it is part of a larger branch of linguistics whose import cannot be questioned: reference.[6]

Resolving reference in linguistics refers to establishing a link between an overt or elided text element and some entity in the real or conceptual world. References to instances of real-world objects and events can be made in language using a variety of expressive means, such as:

- direct reference by name: *Last week Secretary of State Colin Powell went to Belgium*;
- pronominalization and other deictic phenomena, as in *The goal of his visit there was to take part in a meeting of the foreign ministers of the NATO countries*;
- indefinite and definite descriptions of various kinds, as in *This was the Secretary's third overseas trip in less than two months*;
- ellipsis, as in *Defense Secretary Rumsfeld planned to travel* [*to Belgium*] *with him but had to cancel* [*his trip*] *at the last moment*; and
- nonliteral language (that is, metaphors, metonymies and other tropes), as in *The White House* [*metonymy*] *hopes that the visit will tip the balance* [*metaphor*] *of European opinion in favor of sending NATO forces to Iraq*.

In order to create a full semantic representation of a sentence, as is necessary for knowledge-rich NLP, it is not enough for text elements to be interlinked; all referring expressions must be linked to an "anchor" in the real or conceptual world—or so we

believe in the Ontological Semantic research group (see, e.g., Nirenburg and Raskin 2004). That is, it is not sufficient for *his* to be coreferenced with *Secretary of State Colin Powell* in the text, but both of these must be linked to the anchor for *Colin Powell* in the world model—which includes many types of other information about him as well. The difference in processing the various types of referring expressions, then, reduces to carrying out different procedures to establish this text-element to world-model linking.

Most ellipsis studies endeavor to go partway, delineating rules for establishing *co*reference with an antecedent (or, occasionally, a postcedent) but not a reference link to any concrete object or event in the real world. For example, in the sentence *Peter* ***finished*** *at five and Paul* ***∅****, at six*, coreference between the elided verb in the second conjunct and overt *finished* in the first conjunct can be established based on syntactic rules, but that does not tell us anything about what it is that Peter and Paul finished doing. Thus, although *co*reference has been established, reference has not.

Obviously, one method of resolving reference for elided categories is by first establishing a coreference relation and using it to trace back to a real-world referent. However, not all elided elements are in a coreference relation with another text element; there are two other possibilities. First, they can have an extralinguistic antecedent, as in (5), where the auxiliary *begin* licenses the ellipsis of a verb that means "writing, working, taking the test," and so on:

(5) [Having just passed out test papers] Begin!

Second, they can have a generalized referent understood on the basis of real-world knowledge. For example, the semantically obligatory but grammatically optional agent-adjunct in passives can refer to nonspecific humans (6) or to some contextually implied subset of people or animals (7):

(6) Promises are commonly made, then broken.

(7) During the dry season in Africa, watering holes are desperately sought out and fought for.

These latter two examples represent semantic ellipsis, which this theory—unlike most approaches to ellipsis—covers.

Figure 1.1 shows one way to view ellipsis in the larger picture of reference. Domain A contains "surfacy" phenomena and is the purview of traditional syntactic treatments, which have provided much of the formal work on ellipsis, especially as pertaining to English. Domain B, like Domain A, uses syntactic gaps as triggers for reference resolution but appeals to any branch of linguistics for factors that affect ellipsis. In Domain C, the triggers for reference resolution include unfilled slots in a meaning representation, constructions known to suppress some types of information, and so on (see section 4). And in Domain D, lexical categories of certain types (pronouns, some definite descriptions, etc.) act as triggers for reference resolution.

Within any of these realms, some phenomena lend themselves more easily to formal description than others. For example, straightforward syntactic rules deter-

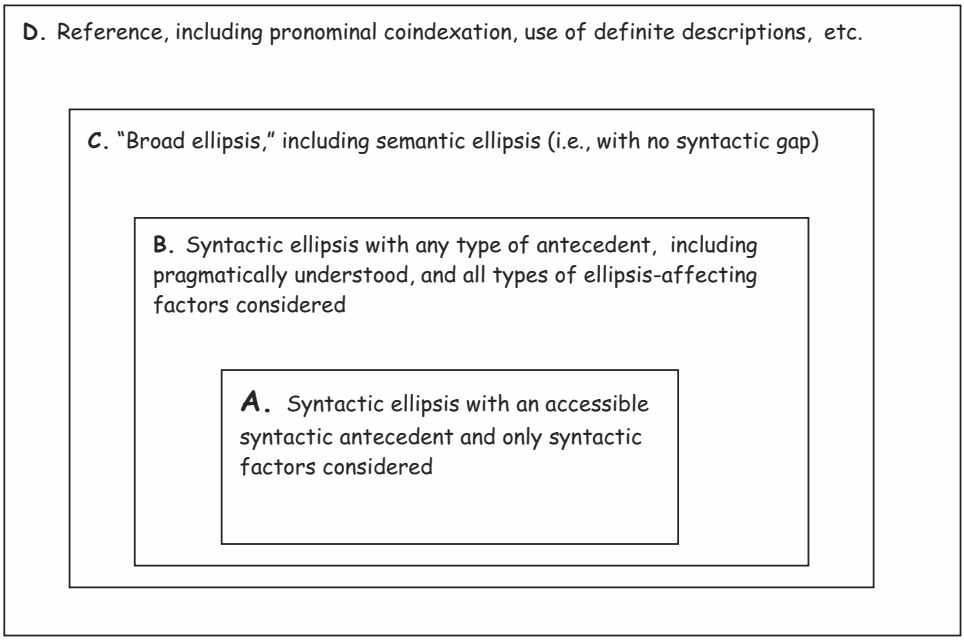


FIGURE 1.1 Ellipsis in the larger picture of reference issues

mine the contextual distribution of English *he*, *him*, and *himself*, but it is much more difficult to generalize about when the use of metaphor or metonymy is appropriate and what the referents for such entities are. As for elliptical phenomena, the more completely they are determined by syntax, the easier it is to formulate rules to account for them, whereas the more they rely on factors from mixed domains or on intangible aspects of human knowledge, the harder they are to capture by rules.

3. Syntactic ellipsis: Background

Syntactic ellipsis is the nonexpression of a syntactically obligatory category whose referent can be recovered by syntactic rules or discourse cues. Syntactically obligatory sentence components include the verb (or other predicate word), its arguments, and certain elements that belong to minor parts of speech, like conjunctions. For example, *The boy is running* is a complete sentence, since the verb *run* requires only one argument, the agent. One can add any number of adjuncts that provide extra information, as in *The boy is running like the wind along the beach in his bare feet*, but since they are not grammatically necessary they are not subject to syntactic ellipsis.[7] Although syntactic ellipsis is a universal property of language, the specific types permitted and the restrictions on them differ from language to language. The following subsections present some of the foundational ideas and terminology that will serve as the syntactic starting point for this theory of ellipsis.

3.1. Licensing and recoverability

In order for syntactic ellipsis to be possible, two conditions must be met: the language must license (permit) ellipsis in the given configuration, and the content of the elided category must be recoverable (understandable). The most common licensing strategies are licensing by a particular type of lexical category and licensing by syntactic parallelism. For example, in English, verbal ellipsis can be licensed by an auxiliary verb, underlined in (8) and (9), or by the type of parallelism shown in (10) (cf. chapter 9):

(8) "If you're going to **procrastinate**, I will ∅, too."

(9) [The speaker, eyeing two slabs of chocolate cake] Shall we ∅?

(10) By midnight Joan **had finished** her term paper and Jason ∅ his math homework.

A licensing strategy not discussed in the mainstream literature but prevalent, for example, in Russian, is licensing by a combination of lexical categories, as in (11) (cf. chapter 10):

(11) Я ∅ не об этом.
Ja ∅ ne ob ètom.
I_{NOM} ∅ NEG about that
'That's not what I'm talking about.'

Referents for syntactically elided categories can be recovered from the linguistic context, as in (8) and (10), the extralinguistic context, as in (9), or one's world knowledge in conjunction with the semantics of the overt categories, as in (11).

3.2. Sentence grammar versus discourse grammar

Licensing and recoverability conditions can be viewed in the framework of sentence grammar or discourse (also called functional) grammar.[8] Sentence grammar—which is the realm of syntactic theories—focuses on the formal, purely syntactic properties of individual sentences. Discourse grammar is a more comprehensive approach to language, studying discourses larger than the minimal sentence and addressing issues related to the speech context, the speaker's and interlocutor's presuppositions and belief systems, grammar, semantics, intonation, and virtually all other factors that affect people's use of language. Some linguists working within the discourse framework believe that discourse grammar is the only valid approach to language, considering the syntactic approach unjustifiably narrow in scope and subsumable under discourse grammar.[9] However, as Morgan (1982) explains, there is no *decision* to be made regarding whether or not sentence grammar exists: it either does exist as an independent component of the human brain (the language faculty) or does not. If it does, it must have special properties that other cognitive systems lack and its workings must be independent and formally definable.

The most compelling evidence in favor of pursuing sentence grammar independently of discourse grammar, for whatever scope of phenomena it can cover, is that rules of discourse are difficult to formalize and implement. This is natural because it is notoriously difficult to quantify notions like "the speaker's presuppositions." As Morgan (1982: 202) puts it, we are still far from achieving the goal of creating a discourse grammar that is as explicit as the sentence grammars currently being developed because of the "almost miraculous complexity of the mental systems that underlie our ability to produce and understand discourse." Therefore, the most productive approach to language study appears to be to account for the maximum number of phenomena within the sentence grammar, then employ discourse grammar to account for the rest.

A number of linguists have recognized the necessity to merge, rather than choose between, these approaches to language study. For example, Lambrecht (1994: 11) says: "The issue which ultimately divides the 'formal' and the 'functional' approach is not so much disagreement about facts but the question of what constitutes explanation in linguistics."[10] Kuno (1987) finds no theoretical conflict between them, despite obvious practical conflicts, and concisely captures the state of affairs:

> In theory there is no conflict between functional syntax and, say, the government and binding theory of generative grammar. Given a linguistic process that is governed purely by syntactic factors, this process will be described in the syntactic component of grammar both by pure syntacticians and by functional syntacticians. On the other hand, given a linguistic process that is governed by both syntactic and, say, discourse factors, the syntactic aspect will be formulated in the syntactic component, while discourse factors that interact with this syntactic characterization will be described in the discourse component of grammar. Pure syntacticians would concentrate on the former characterization and functional syntacticians on the latter. There need not be any disagreement between the two. In practice, however, there are numerous conflicts between the outlooks of pure and functional syntacticians with respect to how to analyze a given linguistic phenomenon. Pure syntacticians tend to give characterizations to linguistic phenomena which are in fact controlled by non-syntactic factors. Or they label them as non-syntactic phenomena and brush them aside, without attempting to find out what kind of nonsyntactic factors are in control. (Kuno 1987: 1–2)

Here I integrate the syntactic and discourse/functional approaches to language, attempting to cover the greatest number of phenomena with the most specific, potentially machine-tractable rules possible, while not excluding from the application-independent description factors that are less easily formalized.

3.3. Discourse-oriented languages

One way of typologically dividing natural languages is into discourse-oriented and non-discourse-oriented groups. In discourse-oriented languages, the grammatical structure of any given sentence is significantly influenced by the surrounding discourse. Such languages tend to have at least some of the following properties: theme-rheme or topic-comment structure, free word order, morphological case marking, and expanded use of ellipsis.[11] English is not a discourse-oriented

language and, as expected, does not employ nearly as many forms of ellipsis as discourse-oriented languages do. Therefore, it would be impossible to create a cross-linguistically useful theory of ellipsis using English as the sole source of phenomena. For this reason, material from discourse-oriented languages is widely used in this book, with Russian and Polish figuring most prominently. Ultimately, phenomena from many more languages will need to be incorporated if true cross-linguistic coverage is to be achieved. However, the proposed sampling is sufficient to develop the core of the theory, which can—indeed, must—then be extended as evidence suggests.

Russian and Polish are both Slavic languages, from the East and West Slavic groups, respectively. They have a highly developed system of morphological case marking (six cases for Russian, seven for Polish, and two numbers each), free word order, and theme-rheme structure as the realization of their discourse orientation.[12]

Theme and rheme deal with the logical flow of information in a discourse. The theme can be described in various ways: as the starting point for the message, what the message is about, what the speaker's attention is focused on, and so on. The rheme says something about the theme and contains the most new information. The theme tends to precede the rheme, since one first establishes or confirms the topic of conversation and then asserts something about it, as shown in (12):

(12) «Ну что, будем готовить ужин?» — «Нет, ужин приготовит Маша».
«Nu čto, budem gotovit' užin?» — «Net, užin prigotovit Maša».
well what will$_{1PL}$ prepare$_{INFIN}$ dinner$_{ACC}$ no dinner$_{ACC}$ prepare$_{3SG.FUT}$ Masha$_{NOM}$
'"So, are we going to make dinner?" "No, *Maša's* making dinner."'

The fact that the object *dinner* precedes the subject *Maša* in Russian shows that it is the theme.

Theme-rheme discourse structure relates to ellipsis in two ways: generally, only thematic (already understood) elements can be elided; and thematic elements are more likely to be elided when some other element in the sentence is strongly emphasized. The latter point is illustrated in (13), where the direct object practically must be elided in the last three utterances because of the emphasis on the verb itself:

(13) «Дальше . . . Что там у тебя дальше!» — «**Ригодон**» . . . — «Выучил ∅?» . . . — «Я ∅ учил». — «Выучил ∅ или нет?» (Токарева: 55).
«Dal'še . . . Čto tam u tebja dal'še!» — «**Rigodon**» . . . — «Vyučil ∅?» . . . — «Ja ∅ učil». — «Vyučil ∅ ili net?» (Tokareva: 55).
further what there at you$_{GEN}$ further **Rigodon** memorized$_{2SG.MASC}$ ∅$_{ACC}$ I$_{NOM}$ ∅$_{ACC}$ studied$_{2SG.MASC}$ memorized$_{2SG.MASC}$ ∅$_{ACC}$ or not
'"Go on . . . What's next?!" "Rigodon." . . . "Did you memorize it?" . . . "I studied it." "But did you *memorize* it?"'

To summarize, discourse orientation is important in this work for the following reasons:

1. Since discourse-oriented languages typologically permit more ellipsis than non-discourse-oriented languages, a non-discourse-oriented

language like English does not provide many of the types of data required for a cross-linguistic theory of ellipsis.
2. Discourse orientation provides grammatical reflexes of the theme/topic (i.e., information activated in the discourse) and the rheme/comment (i.e., new or most relevant information in the current utterance). One such reflex is the widespread ellipsis of thematic elements.
3. In many instances, ellipsis is not merely an optional grammatical process but an important means of supporting the discourse structure.

3.4. Direct and reverse valency

Direct valency describes a configuration in which a head implies its elided complement(s), whereas reverse valency describes a configuration in which one or more complements imply their elided head.[13] Of the types of ellipsis widely discussed in the English literature, direct valency obtains, for example, in VP Ellipsis (where an auxiliary signals its elided verbal complement) and subject ellipsis (where the finite verb form signals the absence of its external argument). Reverse valency obtains, for example, in head noun ellipsis (where a modifier signals the absence of the selecting noun: *If you're making **coffee** anyway, I'll have some* ***Ø***).

The terms *direct valency* and *reverse valency*, while applicable to phenomena found in English, neither come from nor are used in the English literature. Rather, they come from the Russian literature, where they have particular import. For example, if a host holds out a piece of cake to a guest and asks:

(14) ∅ Хочешь ∅?
∅ Хочеš' ∅?
$\emptyset_{\textbf{NOM}}$ $\text{want}_{2.\text{SG}}$ $\emptyset_{\textbf{ACC}}$
'Would you like some?'

the verb implies the presence of both of its complements, which are instances of direct valency. However, if the host is pouring tea and asks his guest:

(15) Тебе крепкий ∅?
Tebe krepkij ∅?
you_{DAT} $\text{strong}_{\text{NOM.SG.MASC}}$ ∅
'Would you like yours strong?'

the adjectival modifier implies the head noun *tea*, an instance of reverse valency. In colloquial Russian, a single short sentence like (16) can contain elided categories licensed in both ways:

(16) [The interlocutor is serving cake]
Можно ∅ ∅ с розочкой? (Zemskaja, Kitajgorodskaja, and Širjaev 1981: 206).
Možno ∅ ∅ s rozočkoj? (Zemskaja, Kitajgorodskaja, and Širjaev 1981: 206).
$\text{possible}_{\text{IMPERS}}$ ∅ ∅ with $\text{rose}_{\text{INSTR}}$
'May I have one <a piece> with a rose?'

In this example, *possible*$_{IMPERS}$ implies a verb like *get, have* and the first-person experiencer by direct valency, whereas *with a rose* implies *piece of cake* by reverse valency.

Reverse valency is especially important for the analysis of what I call Multilicensor Verbal Ellipsis, used in sentences like (17):

(17) Я ∅ в кино.
Ja ∅ v kino.
I_{NOM} ∅ to movies
'I'm going to [or: I'm off to, I'm heading to] the movies.'

Here the combination of a Nominative-case noun phrase that means "I" and the directional prepositional phrase *to the movies* both licenses the ellipsis and ensures recoverability of the verbal meaning. This type of ellipsis, which to my knowledge is discussed only in Slavic sources, is the subject of chapter 10.

3.5. Parallelism

The grammatical effects of parallelism have long been recognized and are addressed, for example, in the literature on Gapping and coordinate structures. But the most interesting aspect of parallelism with respect to ellipsis, which has not yet attracted attention, is that it functions on many levels simultaneously. The syntactic and semantic parallelism constraints on Gapping are well known. For example, whereas *John rides his bike in the park and Mary ∅ on the boardwalk* is fine, **John rides his bike in the park and Mary ∅ with pleasure* is not, because the semantic classes of the prepositional phrases do not correspond.[14] However, the ellipsis effects of parallelism go much deeper. Apart from syntactic and lexico-semantic parallelism, morpho-syntactic and phonetic parallelism also affect ellipsis judgments. Morpho-syntactic parallelism obtains when antecedents and their coreferential categories have the same features (e.g., case, number). Phonetic parallelism obtains when coreferential categories have the same phonetic form (which does not imply that they must have the same features, since case syncretism is common in case languages). Evidence from Russian (especially in chapter 3) and Polish (chapter 11) will show that if many types of parallelism co-occur in a given structure, ellipsis may shift from being optional to being virtually obligatory.

3.6. Coordination

Coordinate structures figure centrally in the analysis of many elliptical patterns, so their properties deserve mention.[15] For ellipsis to occur in coordinate structures, certain types of syntactic and semantic parallelism must obtain. However, coordinate structures in principle do not require such parallelism. As Sag, Gazdar, Wasow, and Weisler (1985: 140) note: "The conventional wisdom on this topic has it that conjuncts must all be of the same category, say α, and that the mother of these conjuncts will also be of category α. But the conventional wisdom is wrong, or at best, seriously incomplete." Amid supporting evidence for this are grammatical

sentences like (18a) and (18b) (Sag et al. 1985), in which the conjuncts are of different syntactic categories:

(18) a. Pat is either stupid or a liar.
b. Pat is a Republican and proud of it.

In addition, one can say things like *She teased me **and** you know what else? She pulled my hair,* in which the coordinating conjunction is used to join quite incomparable entities.

Conjuncts in a coordinate structure can be joined by an overt conjunction (syndetic coordination) or not (asyndetic coordination).[16] Although analyses of coordination in English tend to focus on coordination with an overt conjunction, many languages permit coordinate structures without an overt conjunction. As Sag, Gazdar, Wasow, and Weisler (1985: 133, footnote 8) note, the presence or absence of a coordinating morpheme is "a highly parochial matter."[17]

In her cross-linguistic survey of coordination, Mithun (1988) notes that some languages have obligatory coordinating conjunctions, some mark coordination morphologically, some have no syntactic markers of coordination at all, and some use a combination of syntactically overt and syntactically null means of indicating coordination. She explains that some languages that initially had no overt conjunctions (and juxtaposed clauses entirely through intonation) have diachronically adopted conjunctions from other languages. This has occurred under the relatively recent influence of the written language on the spoken language, since written language cannot adequately reproduce the pauses, intonation, rhythm, and pitch of spoken language. Mithun's work includes two points that are particularly relevant to this study. First, coordination can be signaled by intonation alone, a strategy used far more widely in Russian than in English.[18] Second, coordination without an overt conjunction relies heavily on intonational cues and is employed principally in spoken language.

Determining whether or not juxtaposed categories represent asyndetic coordination can be tricky. Quirk, Greenbaum, Leech, and Svartrik (1972) suggest that an *and*-insertion test is sufficient to determine whether juxtaposed elements are asyndetically coordinated.[19] According to this test, (19b) would contain asyndetically coordinated adverbs.

(19) a. Slowly and stealthily, he crept towards his victim (Quirk et al. 1972: 550).
b. Slowly, stealthily, he crept towards his victim (Quirk et al. 1972: 550).

However, the *and*-insertion test, while a valuable heuristic, is not a litmus test for asyndetic coordination, as shown by (20): the (a) variant is a coordinate structure that blocks ellipsis, while the (b) variant, which lacks a conjunction, permits ellipsis:

(20) a. Он больше не катается на **велосипеде** и, наверно, скоро **его** продаст.
On bol'še ne kataetsja na **velosipede** i, naverno, skoro **ego** prodast.
he$_{NOM}$ anymore NEG rides on **bicycle$_{PREP}$** and probably soon **it$_{ACC}$** will-sell
'He doesn't ride his bike anymore and probably will sell it soon.'

b. Он больше не катается на **велосипеде**. Наверно, скоро (**его**) продаст.
On bol'še ne kataetsja na **velosipede**. Naverno skoro (**ego**) prodast.
he_{NOM} anymore NEG rides on $\textbf{bicycle}_{\textbf{PREP}}$ Probably soon $(\textbf{it})_{\textbf{ACC}}$ will-sell
'He doesn't ride his bike anymore. He'll probably sell it soon.'

The natural explanation for the difference in ellipsis potential is that the second sentence is not a coordinate structure with a hidden conjunction but, rather, a different discourse structure: an assertion followed by its elaboration (see section 3.7).

For this theory of ellipsis, three aspects of coordination are key. First, although most of the literature on ellipsis in coordinate structures assumes that conjuncts are linked by an overt conjunction, this need not be the case, and the effects of conjunction use on ellipsis require investigation (see especially chapter 3, section 1, and chapter 9, section 1). Second, it is often difficult to determine which structures represent asyndetic (conjunctionless) coordination and which represent other types of clause juxtaposition. In many cases, elliptical results suggest one or the other analysis, but from the point of view of analysis, this is rather like putting the cart before the horse. The problem—which I will not attempt to fundamenally solve, although I will suggest some heuristics—is how to automatically determine the correspondence between juxtaposed clauses and, based on that, predict ellipsis potential. Third, it is not trivial to automatically establish semantic comparability between elements of text, which is required for the analysis and generation of elliptical structures whose missing elements are licensed by parallelism. Ontological Semantics (briefly sketched in section 4.3) offers significant promise in this area, since the concepts instantiated by each relevant text element can be compared with respect to their properties as well as their placement in the tangled tree of concepts that constitutes the ontology.

3.7. In opposition to coordination: Assertion and elaboration

Whereas coordinate configurations generally contain two functionally equivalent conjuncts joined by an overt conjunction, what I call Assertion and Elaboration (A and E) configurations contain two clauses that are not functionally equivalent and are generally not joined by any lexical items—though some form of punctuation in the written language is expected.[20] In A and E configurations, the first clause states something and the second explains or otherwise comments upon it:

(21) Малый не стал возиться с **сумкой**: схватил (**ее**) и дал дёру.
Malyj ne stal vozit'sja s **sumkoj**: sxvatil (**ee**) i dal dëru.
guy_{NOM} NEG began fool with $\textbf{bag}_{INSTR}$ grabbed $(\textbf{it})_{\textbf{ACC}}$ and took-to-his-heels
'The guy didn't fool with the bag: he grabbed it and took to his heels.'

A and E configurations are best considered clause complexes rather than sentences because this terminology allows us to circumvent the thorny issue of what, precisely, constitutes a sentence, punctuation, of course, playing no role in this determination.[21]

Especially important for our purposes is that A and E configurations permit object ellipsis in many instances where similar coordinate structures do not. Their ellipsis-promoting power is particularly clear in instances when the antecedent does not match the elided category in case marking (see especially chapter 4, section 1, and chapter 5, section 2), suggesting two things: first, that a mismatch in case marking is tolerated only when the antecedent structure and elliptical structure are, themselves, not parallel; and second, that the semantic and pragmatic aspects of following up an assertion by its elaboration support ellipsis in ways that coordinate structures do not.

3.8. Does obligatory syntactic ellipsis exist?

Traditional definitions of syntactic ellipsis require that it be possible to overtly express the elided category. So, (22a) has syntactic ellipsis because it means the same as (22b) and (22b) is grammatical:

(22) a. Mary **went** to the beach and John **Ø** to the mountains.
b. Mary **went** to the beach and John **went** to the mountains.

If the category cannot be overtly expressed, there is no syntactic ellipsis. For example, the obligatory nonexpression of the arguments of nonmatrix verbs is not syntactic ellipsis; it is semantic ellipsis. So, although a full semantic representation requires the establishment of *Mary* as the agent of *to go* in (23) and *returning* in (24), grammatical rules block the possibility of referring to her overtly as associated with those nonmatrix verbs:

(23) Mary likes to go to the beach.

(24) Returning home from the beach, Mary was in a good mood.

The possibility or impossibility of an overt-category variant is one heuristic for syntactic versus semantic ellipsis. What this diagnostic masks, however, is that although an overt-category variant may be grammatically possible, it is not necessarily usable in the language that people actually speak and write. If, for example, a person was announcing the seating arrangements for a wedding, it is hardly likely that he or she—unless a pedantic bore—would pronounce (25a) in full. Some elliptical variant like (25b) would be expected.

(25) a. The McCoys will be seated at table 1, the Burkes will be seated at table 2, the Clanceys will be seated at table 3, the Conroys will be seated at table 4, the Deans and Bakers will be seated at table 5 . . .
b. The McCoys will be seated at table 2, the Burkes ∅ at table 2, the Clanceys ∅ at ∅ 3, the Conroys ∅ at ∅ 4, the Deans and Bakers ∅ at ∅ 5 . . .

So, although a valid diagnostic for syntactic ellipsis is *Can the category be inserted?* the question still remains *Would the category be inserted?* This is a question not generally

taken up by syntacticians because the grammaticality of a process is of more relevance for their purposes than rules of usage. However, speakers do have intuitions about what constitutes too much repetition and what sounds overly elliptical, just as they have intuitions about grammaticality as such. Although these intuitions can be affected by the speech situation (e.g., more repetition can be used for emphasis) and although these intuitions are not as consistent as many of those used for grammaticality judgments, they are quite strong for each individual speaker and are relatively stable from speaker to speaker. Thus, although *syntactically* obligatory syntactic ellipsis cannot, by definition, exist, there is evidence for *pragmatically obligatory syntactic ellipsis*.

The situation, however, is even more complex than this. Apart from pragmatically *obligatory* syntactic ellipsis, there is also highly preferred, preferred, completely optional, less desirable, and so on, syntactic ellipsis. So, ellipsis judgments are not a yes-no option; they represent a scale of acceptability.

The examples of ellipsis one finds in texts tend to be solidly in the "can/must elide" camp, with the occasional "should not but may" that is used for stylistic effects or to represent a high degree of colloquialism. Delineating the perimeter of normal ellipsis usage, however, requires asking of native speakers that they make judgments in the gray area around this perimeter. The challenge set to them is exemplified by the following set of English examples:

(26) a. Republicans **are** conservative and Democrats ∅ liberal.
b. John **hates** liberalism and Jane ∅ surfing.
c. John **hates** liberalism and Jane ∅ to go surfing.
d. John **hates** liberalism and dogs ∅ to go surfing.

All speakers of English must agree that (26a) is grammatical: it juxtaposes two comparable traits possessed by two comparable groups. Example (26b) is slightly deviant, because *liberalism* and *surfing* are not semantically comparable. Example (26c) is more deviant, because *liberalism* and *to go surfing* are not only semantically incompatible but categorially different, too. And (26d) is quite bad, perhaps even uninterpretable if not primed by the preceding examples. The complication arises in labeling (b)–(d): where does "strange but grammatical" end and "ungrammatical" begin? In this work, I will attempt to push the envelope in answering this question, bringing to center stage real usage of ellipsis, rather than limiting investigation to generalized patterns.

3.9. What has understandability got to do with it?

The fact that speakers may be able to recover a category if it is elided does not mean that ellipsis of that category is grammatical. If, for example, a non-native speaker of English were to utter (27a), it would be understood, even though (27b) is the grammatical way of saying it:

(27) a. *Mom accidentally let out the bird, but Dima caught.
b. Mom accidentally let out the bird, but Dima caught it.

Ellipsis is blocked in (27a) because English does not permit object ellipsis (outside of telegraphic speech, recipe contexts, etc.). In languages that permit object ellipsis, however, understandability is still not sufficient to make ellipsis grammatical, as is shown by the contrast between (28a) and (28b), which derives from the different function and case marking of the antecedent in each instance:

(28) a. Мама случайно выпустила **птицу**, но Дима (**ее**) поймал.
Mama slučajno vypustila **pticu**, no Dima (**ee**) pojmal.
Mom accidentally let-out **bird**$_{\text{ACC}}$ but Dima (**it**)$_{\text{ACC}}$ caught
'Mom accidentally let out the bird, but Dima caught it.'

b. В комнату влетела **птица** и Дима поймал **ее/*∅**.
V komnatu vletela **ptica**, i Dima pojmal **ee/*∅**.
into room flew **bird**$_{\text{NOM}}$ and Dima caught **it/*∅**$_{\text{ACC}}$
'A bird flew into the window and Dima caught it.'

As we see, although recoverability of the referent is a required condition for the use of ellipsis, it is not a sufficient condition. The second required condition is that the given type of ellipsis be licensed *in the given configuration*. As the contrast between (28a) and (28b) shows, ellipsis licensing does not reduce to permitting categories of type *x* (e.g., direct objects) to be elided wherever they can be understood; it also constrains the contexts that are determined by the language system to ensure content recoverability. For this reason, for each type of category that can be elided, every conceivable type of antecedent must be tested for its fitness to support ellipsis. This generalization is fleshed out and supported here by the three chapters devoted to the ellipsis of Accusative objects, for which Accusative (chapter 3), Nominative (chapter 4), and oblique (chapter 5) case-marked categories can serve as antecedents in certain configurations.

4. Semantic ellipsis, and reference as distinct from coreference

I have just reviewed some of the foundational aspects of syntactic ellipsis and their import for this work. Now I turn to semantic ellipsis, which I define as the nonexpression of information that, although syntactically not required, is necessary for a full semantic representation of the sentence. Semantic ellipsis is as important for the machine processing of natural language as syntactic ellipsis is and thus occupies an equally central role in this theory.

Semantic ellipsis differs from syntactic ellipsis in two ways: (1) Whereas the immediate goal in resolving syntactic ellipsis is to establish a *co*reference relation with some contextual element (which does not guarantee, however, that reference has been resolved, since that contextual element could, itself, require reference resolution), the immediate goal in resolving most instances of semantic ellipsis is to directly establish a link to a real-world referent. (2) With syntactic ellipsis, there is a gap in the syntax to trigger the search for reference resolution, whereas with semantic

ellipsis there is not. Instead, the triggers for semantic ellipsis are (1) words and collocations that are lexically flagged as requiring reference resolution, (2) syntactic structures that are known to imply unexpressed elements, and (3) in the case of Ontological Semantic text processing, unfilled slots in instances of ontological concepts. Delineating these and other formal triggers for ellipsis constitutes a separate and important first step toward reconstructing the unexpressed meaning.

4.1. Words and collocations that require reference resolution

Certain closed-class lexical items predictably require reference resolution and thus represent semantic ellipsis, like the following in English:

- predicative possessors: *Yuk! This isn't my glass. That one's* ***mine***.
- negators: *How come they offered brandy to you but* ***not*** *to me*?
- referential "do": *You packed in more than I thought you'd* ***do.***
- referential "so": *He thinks it's time to leave, but I don't think* ***so.***

These are not instances of syntactic ellipsis, because reference-establishing lexical items cannot be inserted and still yield a grammatical structure (* . . . *that one's mine glass*). Other closed-class lexical items can also trigger reference resolution when used without a nominal complement, as in: *If it's a choice between cookies and ice cream, I'll take* ***both****! That liqueur isn't my favorite; I'd prefer* ***this***.

A challenge in processing such text elements, however, is that many of them are ambiguous, requiring syntactic analysis to determine whether reference resolution is required. Instances where reference resolution is not required include:

- possessors as predicates in generalizing statements: *What's* ***mine*** *is* ***yours***.
- negators in their nonreferential usage: *She's silly,* ***not*** *stupid.*
- "do" in its nonauxiliary function: *Do you have to* ***do*** everything *well?!*
- "so" as an adverb of degree: *I like him* ***so*** *much!*

Thus, the inventory of closed-class items is one source of potential triggers for semantic ellipsis. Apart from these, there exist lexically licensed "shorthands" that can be more difficult to recognize since they do not constitute any obvious class. In fact, they might only be recognized when the given structures are subject to cross-linguistic comparison or the necessity of creating a full semantic representation. For example, missing semantic elements must be recovered in instances of metonymy (*I like reading* [*books written by*] *Tolstoy*), main verbs implied by the combination of an auxiliary and object (*I forgot* [*to take*] *my wallet, She started* [*singing*] *the song fortissimo*), and so on.[22] In some instances, one could write lexico-syntactic rules to cover such sentences: for example, if the object of *read* is a proper name *or* a personal pronoun with a person as its referent *or* a common noun that refers to a writer, playwright, poet, and so on, then the structure will be restored to *read* [*a reading material written by*] *that person*—with the nature of the reading material delimited, if possible, on the basis of knowledge about the writer in the world model (e.g., James

Michener wrote mostly historical novels). Lists of generalizable constructions of this type can be complied over time when developing and refining NLP systems. Since they are quite lexically idiosyncratic, I will not at this time seek cross-linguistically generalizable parameters and values for their description.

4.2. Syntactic structures known to imply unexpressed information

Different syntactic structures have different functions. Although *Marcy kissed Linus* and *Linus was kissed by Marcy* have the same basic meaning, the first emphasizes Marcy's agency while the latter emphasizes Linus's patiency. Thus, the contrast between active and passive voice can be a contrast in focus or point of view. However, the passive voice also permits the agent to go unspecified, in which case it can have various types of referents:

- a definite, contextually recoverable referent: *When Marcy and Linus were walking in the park, Charlie saw Linus get kissed* (by Marcy);
- a nonspecific referent of a certain class, delimited by world knowledge: *By the age of five Linus had already been kissed* (by some girl);
- a generalized-human referent: *Poor Linus has never been truly loved!* (by any human being or, by extension, any domestic pet); and
- any agent or action that can have agentlike effects: *Her grandparents were killed before she was old enough to know them* (by a robber, a train wreck, a storm, disease . . .).

So, one use of the passive voice is to permit a speaker not to state the agent or cause (for nonagentive entities, like a storm or disease), leaving open the possibility for generalized interpretations in some contexts.

Just as agents in passives can go unexpressed, so can experiencers in certain predictable syntactic contexts. Evaluative statements like *It's cold* and *This is boring* imply that, for the speaker and others in his or her real or imagined position, the temperature is uncomfortably low and the activity is not engaging. However, these evaluations are not statements of some higher truth: a Hawaiian might experience coldness at 50 degrees, whereas a resident of northern Siberia would consider that temperature balmy. In English, these hidden-experiencer constructions are often signaled by *It is* or *This is*, whereas in many languages impersonal verb forms have this function. Diagnosis of such structures, however, is often not possible using syntax alone, which leads to the necessity of semantic analysis. Chapter 13 explores a number of syntactic structures known to imply unexpressed information, including unexpressed agents in passives, agentive impersonals, unexpressed experiencers and possessors, and unexpressed arguments in nominalizations and nonmatrix verbs.

4.3. Unfilled slots in instances of ontological concepts

An NLP-oriented ontology is a model of the world used to support the kinds of reasoning necessary at every step in language processing. Some common language tasks that require world knowledge are:

- disambiguation, as is necessary for distinguishing between the senses of *bank* in *I opened up a checking account at the bank* versus *He slid down the bank into the river*;
- relating commonly connected objects and events, as is necessary for understanding that the flowers are affected by the falling of the vase in *Her boyfriend gave her a dozen roses for her birthday, but the cat promptly knocked over the vase*;
- understanding subclass/superclass relations, as is necessary to understand why the Doberman would be fearful in *The dogcatcher was roaming the neighborhood and the Doberman instinctively felt fear*; and
- placing events in scripts, as is necessary to understand why the first instance of *waiter* in a text can have a definite article in English in a context like *I wandered into the empty restaurant and the waiter looked up disinterestedly*.[23]

The formalism and expressive power of every existing formal ontology is different, as is the correlation between it and other knowledge resources (e.g., lexicons and repositories of real-world facts), and the types of programs written to employ it. In a work of this type it would be preferable to consistently speak of ontological text processing in general terms. However, in practice that is very difficult, because some grounding details are necessary. I will use as a sample ontological model the Ontological Semantic text-processing environment implemented in NLP systems like Mikrokosmos and its successor, OntoSem.[24] The description here is, by necessity, very brief but should suffice for the current purposes.

The Ontological Semantics ontology currently consists of a tangled tree of about fifty-five hundred concepts, divided at the highest level into OBJECTS, EVENTS, and PROPERTYS. Concepts are denoted using English words and phrases, written in small caps to differentiate them from English words (i.e., the language-independent concept DOG is not equivalent to the English word *dog*). The tree structure is determined by IS-A links. Properties are inherited unless locally blocked or specified as having different values from their parents, grandparents, and so on. So, one branch of the tree contains the IS-A hierarchy:

DOG
IS-A CANINE
IS-A MAMMAL
IS-A VERTEBRATE
IS-A ANIMAL
IS-A ANIMATE
IS-A PHYSICAL-OBJECT
IS-A OBJECT
IS-A ALL (ALL is the base node of the tree)

Concepts are defined by their inventory of property-facet-value triplets. Properties represent things like size, weight, case roles, and component parts. Facets define what, exactly, the value represents. For example, one can make a distinction between the default value and all "regular" values of a property by using the facet DEFAULT

for the former and SEM for the latter. Similarly, one can block certain fillers by using the NOT facet (e.g., FLY-ANIMATE has AGENT: SEM: BIRD, AGENT: NOT: PENGUIN). Values can be represented by literals or other ontological concepts. Some examples of property definition and inheritance follow. Each concept is actually defined by an average of 16 properties, both locally specified and inherited from ancestors, so these examples are just snapshots for illustration:

DOG is locally specified as

AGENT-OF	SEM	BARK
WEIGHT	DEFAULT	(<> 4, 40) ; a range with the default measure being kilograms
	SEM	(<> 1, 70)

DOG inherits from its parent, CANINE

AGENT-OF	SEM	HOWL
HAS-AS-PART	SEM	TAIL, CLAW, TOOTH . . .

DOG inherits from its grandparent, MAMMAL

HAS-AS-PART	SEM	HEAD, EYE, MOUTH . . .

DOG inherits from its great-grandparent VERTEBRATE,

HAS-AS-PART	SEM	SPINE

DOG inherits from its great-great-grandparent ANIMAL

AGENT-OF	SEM	BREATHE, THINK, and so forth

. . . and so on.

Multiple inheritance is possible, so STORM is a child of both NATURAL-HAZARD and WEATHER, inheriting relevant properties and reflecting the necessary local constraints.

Lexical items in the Ontological Semantic lexicons for specific languages are mapped to the relevant concepts. Since concepts are generally less specific than any single lexical item, a many-to-one mapping is typical. For example, since the Ontological Semantic ontology has not yet been used in applications devoted to zoology, that branch of the ontology is not highly developed and all lexical entities that represent types of lizards are mapped to the concept LIZARD. Should further specification of the properties of different types of lizards become important, they can be expressed either by developing this branch of the ontology (creating differentiated children for the parent LIZARD) or by constraining properties of given lexical items in the relevant lexicon entries for a given language. This latter method is useful, for example, to describe words like German *Schimmel* 'white horse', which maps perfectly well to the concept HORSE but needs its color to be lexically constrained to white (there is no reason to have a concept WHITE-HORSE in the language-independent ontology).

During text processing, each lexical item instantiates an *instance* of the concept to which it is linked. So, in the sentence *The wall was painted*, a numbered instance of the event PAINT and a numbered instance of the object WALL are created (e.g., PAINT-23 and WALL-355, if this is the twenty-third time the concept PAINT is instantiated, and the three hundred and fifty-fifth time the concept WALL is instantiated, in the

world model used when processing this text). The ontological frame for PAINT has, among others, slots for:

AGENT	SEM	HUMAN
THEME	DEFAULT	ARTIFACT
	SEM	OBJECT
INSTRUMENT	DEFAULT	PAINT-BRUSH
	SEM	IMPLEMENT

The input sentence supplies a valid filler for the THEME, since ARTIFACT is a great-grandparent of WALL. It does not, however, supply any filler for AGENT or INSTRUMENT. However, the fact that the concept frame includes reference to an AGENT and an INSTRUMENT is a sort of trace that permits reasoning based on ontological, rather than textual, information—in this case, there had to be a human involved in the event and some implement had to be used to carry out the process.

It is namely this type of ontological information that can be used to detect and often resolve semantic ellipsis. Take, for example, *Salmonella jeopardized the farm*. A semantic analysis would require making inferences using the following types of interrelated, ontologically encoded information:

SALMONELLA:	EXPERIENCER	SEM	HUMAN
	IS-A	SEM	ANIMAL-DISEASE, NATURAL-HAZARD
ANIMAL-DISEASE:	THEME-OF	SEM	DIAGNOSE, HEAL
	HAS-EVENT-AS-PART	SEM	ANIMAL-SYMPTOM, TREAT-ILLNESS
NATURAL-HAZARD:	IS-A	SEM	DISASTER-EVENT
DISASTER-EVENT:	DEATH-TOLL	SEM	(>= 0)
	PROPERTY-DAMAGE	SEM	(>= 0)
FARM:	LOCATION-OF	SEM	PRODUCE
PRODUCE:	IS-A	SEM	WORK-ACTIVITY, PHYSICAL-EVENT
	AGENT	SEM	HUMAN
JEOPARDIZE:	SOURCE	SEM	EVENT
	THEME	SEM	OBJECT

So, even though the original sentence says nothing about humans, the knowledge resources provide the information that HUMANS are the only possible experiencers of SALMONELLA; that SALMONELLA is an ANIMAL-DISEASE that involves ANIMAL-SYMPTOMS; that a PRODUCE event goes on at farms that requires HUMAN agents; that SALMONELLA is a NATURAL-HAZARD, which in turn is a DISASTER-EVENT that can cause a DEATH-TOLL and PROPERTY-DAMAGE, and so on. In effect, the ontology tells us general facts about events that are part of the world knowledge a human speaker or hearer would bring to a speech event. If we can ultimately incorporate enough rules of reasoning to leverage such ontological information—something we are working on in

earnest at present—such reasoning should add significantly to the quality of NLP systems.

Apart from "simple" ontological frames, the ontology also contains scripts, which are complex events that represent typical world events, like getting ready for work (showering, eating breakfast, getting dressed), going to a concert (getting dressed up, sitting quietly in rows of seats, applauding, throwing flowers), and preparing dinner (washing and chopping vegetables, cooking food, setting the table). Scripts, like simple ontological concepts, can help to detect and resolve semantic ellipsis, not to mention syntactic ellipsis.

5. Ontological semantic text processing and practicality

Ontological Semantic processing for text analysis relies on the results of a battery of presemantic text-processing modules: the tokenizer, the morphological analyzer, the lexical look-up module, and the syntactic analyzer. The output of these modules provides input to and background knowledge for semantic analysis. The tokenizer module deals with any mark-up symbols in the input text, finds boundaries of sentences and words, and detects and recognizes dates, numbers, named entities, and acronyms. Morphological analysis uses the results of tokenization.

A morphological analyzer accepts a string of word forms as input and for each word form outputs a record that contains its citation form in the lexicon and a set of morphological features and their values that correspond to the word form from the text. Once the morphological analyzer has generated the citation forms for word forms in a text, the system can look them up in its lexicons, including the onomasticon (a lexicon of names), and thus activate the relevant lexical entries. The task of syntactic analysis in ontological semantics is, essentially, to determine clause-level dependency structures for an input text and assign syntactic valency values to clause constituents (that is, establish subjects, direct objects, obliques, and adjuncts).

Semantic analysis proper takes as input results from the earlier stages of processing and produces a *text meaning representation* (TMR). The central task for semantic analysis is to construct unambiguous propositional meaning by processing selectional restrictions, listed in the ontology and the semantic zones of lexicon entries. Other issues include treating such phenomena as aspect and modality, nonliteral language (which, incidentally, is important for the treatment of reference as well), and building a discourse structure associated with the basic propositional structure of the text. For details of this process see Nirenburg and Raskin 2004, chapter 8.

Ontological Semantic text processing has been implemented in systems that work on unrestricted text, meaning that simple and difficult aspects of text processing have equal priority. However, since the difficult problems—like ellipsis, metaphor, and ambiguity—cannot be solved quickly or completely, partial solutions must be not only accepted but also embraced. In practical terms, this means supplementing principle-based analysis with heuristics that are shown to improve results in system trials. Although heuristics-supported approaches may be less theoretically compelling than some approaches discussed in the literature and although they may not cover

rare and borderline cases, they have the benefit of being realistically implementable *now*. The force of the moment cannot be overstated based on the feverish demand for fast improvements in the performance of things like search engines, machine translation systems, and tools to support intelligence analysts.

In response to this need, the Ontological Semantics environment is continuously under development. Participation in this work encouraged me both to extend my earlier theoretical and comparative work (McShane 1998, 1999a,b, 2000, 2002a,b) to the computational realm, and to orient this research effort toward real-time, practical ends. However, this orientation lies in a sparsely populated no-man's-land between statistical approaches, where even 70% accuracy on certain tasks is considered quite acceptable, and theoretical approaches, for which full coverage is the goal but for which nontrivial presuppositions call into question whether implementation will ever be possible. Naturally, research efforts must not be reined in such that implementations can keep pace. However, there is much room (and a pronounced practical need) for high-quality and broad-coverage applications of text processing that can be used right now.

2

Object Ellipsis: Preliminaries

Cross-linguistically, objects can be categorized on the basis of parameters like their function, the category that selects them, their syntactic status, their case marking, the means of realizing their case marking, and the definiteness of their referent. Values for these parameters include those listed in table 2.1.[1]

Since objects, even within a given language, can be defined by many combinations of parameter values, one cannot treat object ellipsis as a single phenomenon. Especially in languages with a highly developed system of morphological case marking, the elidability of an object depends not only on the nature of the object itself but also on the nature of its antecedent as well as many other factors. When all combinations of parameter values for the potentially elided object are combined with all combinations of parameter values for its antecedent, the matter of determining ellipsis potential becomes quite complex—enough to provide material for chapters 3 through 7 in this book.

None of the available literature on object ellipsis attempts this depth of analysis, one reason being that most of the languages best represented in the literature either lack object ellipsis altogether, use it sparingly, or use it but lack a sophisticated system of case marking (see section 4 for some highlights from the literature). English, for its part, permits object ellipsis only in stylistically marked genres, like recipes (*Crack **eggs**; beat ∅ well; pour ∅ into pan; stir ∅ vigorously over high heat*),[2] stage directions (*Holmes walks up to **door**. Opens ∅ slowly. Leaps back in horror*), and other types of telegraphic writing (*Went to store. **Dinner** in fridge. Heat ∅ and eat ∅*). A close study of the factors that determine object ellipsis potential requires data from languages that both use object ellipsis widely and have a developed system of case marking.

TABLE 2.1 Parameters and values for categorizing objects

Parameters	*Values*
Function	Direct object
	Indirect object
	Object of a preposition or postposition
	Lexically case-marked object of a verb
Syntactic status	Full NP
	Pronoun
	Clitic
	Unrealized in surface text
Type of case marking	ACC, DAT, INSTR, etc.
Source of case marking	Configurational
	Lexical (quirky)
Realization of case marking	Paradigmatic morphology
	Agglutinating affixes
	Isolating words
	None
Referent for object	Definite
	Indefinite
	Etc.

1. Categorizing objects by function

In broad terms, direct objects represent the person or thing directly acted upon by the verb (*He devoured **the sandwich***) and indirect objects express recipients, experiencers, or beneficiaries (*The postman delivered the package **to my neighbor***). Apart from these functions, objects can be complements of prepositions or postpositions or complements of verbs that have a meaning or case marking that lies outside of the traditional delineation of direct/indirect object.

Based on the kinds of objects a verb selects, it can be classified as intransitive, transitive, optionally transitive, or ditransitive.[3]

- Intransitive verbs, like *sleep*, do not take a direct object: *The cat is sleeping*.
- Transitive verbs, like *respect*, require a direct object: *I respect your feelings.*
- Optionally transitive verbs, like *eat*, can but need not take a direct object: *He is eating macaroni. Don't bother the dog while she's eating*. Optionally transitive verbs either express processes that one can be involved in over some period of time (*eat, read, write, sing*) or are used in the expression of a generalized characteristic, in which case a modifier is required (*She sings nicely*). When an optionally transitive verb does not select an object, this is not ellipsis; it is the exploitation of a lexically available option.
- Ditransitive verbs, like *give*, take both a direct object and an indirect object: *The cleaning lady gave me the key.*[4]

The types of objects a verb takes, their mandatory versus optional status, and any other unpredictable grammatical facts about them (e.g., quirky case marking) are reflected in the verb's subcategorization frame. Subcategorization frames in conjunction with selectional restrictions should contain all the grammatical and some of the lexical information about a verb that any native speaker unconsciously knows and any NLP system needs to access.

2. Case marking: Types, sources, and realizations

Languages with morphological case marking generally have default case markings for direct and indirect objects. In some languages, there is a different default for direct objects in positive and negated clauses (e.g., direct objects in Polish are ACC in positive clauses but GEN in negated ones). These defaults are assigned configurationally, that is, based on the object's syntactic position. However, many languages also employ lexical (quirky) case marking, which is semantically vacuous, oblique case marking assigned by the particular preposition/posposition or verb that selects the object. All available knowledge about case marking in a given language should serve as the starting point for research into the interplay between case marking and object ellipsis. A brief overview of the current state of knowledge about case marking in Russian follows, which will supply the test set of data for the object-ellipsis portion of the theory (for a similar treatment of Polish, see McShane 1998). This is leading toward questions like *Are all objects subject to the same rules of ellipsis*? and *Does the case marking of the antecedent need to match the case marking of the elided category?*

Case marking in Russian can be assigned configurationally, lexically, or semantically. *Configurational case* is assigned based on a noun phrase's syntactic position: NOM for subjects, ACC for garden-variety direct objects, GEN for adnominal noun phrases and direct object–like noun phrases that fall under the scope of negation or quantification, DAT for recipients and experiencers, and INSTR for certain types of predicate nominals.[5] Here only configurational ACC direct objects will be called direct objects, since direct object–like entities with other case markings have different ellipsis-related properties and therefore must be treated separately. *Lexical* (*quirky*) *case* is imposed by lexical-case-assigning verbs and all prepositions/postpositions. It is unpredictable and free of semantic implications. *Semantic case* derives from the semantic function of a noun phrase in context.[6] For example, the noun phrase *woods* in the Instrumental case can mean "through the woods" in Russian, as in (1):

(1) Он шёл лесом.
On šël lesom.
he_{NOM} walked $\text{woods}_{\text{INSTR}}$
'He was walking through the woods.'

NP objects of verbs in Russian can be case-marked ACC, GEN, DAT, or INSTR, and NP objects of prepositions can be case-marked ACC, GEN, DAT,

INSTR, or PREP. Issues described in the literature that are salient for ellipsis include the following.

2.1. Accusative

ACC case can be assigned configurationally or lexically. It is assigned configurationally to Russian direct objects that neither are lexically case-marked nor fall under the scope of negation or the scope of a quantifier. It is assigned lexically to the objects of certain prepositions. Franks (1995: 53) refers to ACC as "the least marked case" in Russian—the case that is assigned by a verb or a preposition in the absence of feature specifications. Evidence for this includes the following: in most declensional patterns, the ACC is like the NOM or the GEN in form; quantifiers often appear in fixed forms that look like the ACC case; when a preposition assigns two different cases with related meanings, one of those cases is always the ACC (e.g., *в город*$_{ACC}$/*v gorod*$_{ACC}$ 'to the city' ~ *в городе*$_{PREP}$/*v gorode*$_{PREP}$ 'in the city') (Franks 1995: 53). Configurational and lexical ACC are almost the same in morphological and phonetic form—pronominal objects of prepositions take an epenthetic *n*—but they occupy different structural positions, the lexical ones being embedded in a prepositional phrase.

2.2. Genitive

GEN case can be assigned configurationally or lexically. It is assigned configurationally to objects that would have received ACC case had they not fallen under the scope of negation or the scope of a quantifier. It is assigned lexically by certain verbs and prepositions. GEN is arguably the most ACC-like of all the oblique cases. Similarities between GEN and ACC have been suggested in various analyses, some of which make no distinction between configurational and lexical GEN and others of which consider only one or the other. Such studies include the following:

- Jakobson (1984), focusing on semantics, includes GEN among the central cases (NOM, ACC, GEN); these stand in opposition to the peripheral cases (INSTR, PREP, DAT).
- Chvany (1996b) further supports the central nature of the GEN: it shares forms with the animate ACC; it is highly differentiated morphologically (unlike the other oblique cases); it is far more frequent than PREP, DAT, or INSTR; it alternates with ACC in the negative.
- Bailyn (1995) suggests that noun phrases with configurational GEN case marking (GEN of negation and partitive GEN) occupy the same structural position as configurational ACC noun phrases but receive GEN case marking by falling under the scope of negation or the scope of a quantifier.[7]
- Fowler (1996) points out a functional similarity between configurational ACC, lexical GEN, and lexical INSTR, that noun phrases with all of these types of case marking can undergo passivization: "Genitive and Instrumental complements of verbs which assign oblique lexical case are

eligible for passivization because in syntactic terms, they are actually direct objects, with a superficial veneer of oblique morphological case" (538).

2.3. Instrumental

INSTR case can be assigned configurationally or lexically. It is assigned configurationally to noun phrases with attributive function and lexically by certain verbs and prepositions. Since attributive noun phrases are not referential, they will not be considered here. The only noteworthy point about lexical Instrumentals is that they, like lexical Genitives, can become subjects under passivization—a property that unites them with configurational ACC direct objects.

2.4. Dative

The DAT case presents challenges for analysis because the line between lexical and configurational DAT is anything but clear. Lexical DAT, if it exists, should be semantically vacuous, since the very definition of a lexical case precludes semantic nuances. However, Fowler (1996: 537) shows that all Datives belong to certain characteristic semantic groups: benefactives (*помогать/pomogat'* 'help', *льстить/l'stit'* 'flatter', *служить/služit'* 'serve'), malefactives (*мешать/mešat'* 'bother', *досаждать/dosaždat'* 'annoy', *угрожать/ugrožat'* 'threaten'), recipients (*кивать/kivat'* 'nod', *maxat'/maxat'* 'wave'), and DAT recipients/benefactives in three-place predicates (*давать/davat'* 'give'). Moreover, apart from these semantic features, Russian Datives have an idiosyncratic syntactic property that distinguishes them from all other direct object–like entities: they are the only oblique objects not located within a prepositional phrases that systematically cannot become subjects under passivization (Fowler 1996). Thus, it is unclear whether any Datives can be considered purely lexical, in the sense of unpredictable and quirky.

2.5. Prepositional

The PREP case in Russian must be assigned by a preposition. The primary function of the PREP case is to indicate location (*в саду/v sadu* 'in the garden'), but certain verbs have quirky selection of PP complements as well (*жениться на / ženit'sja na* + NP_{PREP} 'marry').

While noun phrases with all of these case markings will be studied with respect to their own elidability and their ability to antecede the ellipsis of other objects, not all possible functions of each case will be considered. Restrictions include the following: All noun phrases under consideration must be referential. This excludes as antecedents attributive INSTR noun phrases (2),[8] semantically case-marked INSTR noun phrase adverbials (3), and noun phrases within a prepositional phrase that indicates measure (4).[9]

(2) Он работает врачом.
On rabotaet vračom.

he$_{NOM}$ works doctor$_{INSTR}$
'He works as a doctor.'

(3) Мы пошли всей группой в музей.
My pošli vsej gruppoj v muzej.
we$_{NOM}$ went whole$_{INSTR}$ group$_{INSTR}$ to museum
'We went to the museum as a group.'

(4) сын ростом с отца
syn rostom s otca
son $_{NOM}$ height$_{INSTR}$ of father$_{GEN}$
'a son about the same height as his father'

In addition, the noun phrase must represent the verb's maximal degree of transitivity. This excludes as antecedents noun phrases that show what Chvany (1996a: 163) calls "reduced transitivity"—that is, noun phrases that have INSTR case marking instead of their canonical ACC case marking:

(5) a. швырять камни
švyrjat' kamni
throw$_{INFIN}$ rocks$_{ACC}$
'throw rocks'

b. швырять камнями
švyrjat' kamnjami
throw$_{INFIN}$ rocks$_{INSTR}$
'throw rocks'

When the object in such pairs is ACC, it is presented as a distinct participant in the action, whereas when it is INSTR, it is presented "either as part of the agent or of the agent's action" (Chvany 1996a: 163). The requirement of maximal transitivity also excludes as antecedents DAT complements of the preposition *po* in constructions like:

(6) стрелять по демонстрантам
streljat' po demonstrantam
fire$_{INFIN}$ at demonstrators$_{DAT}$
'fire at demonstrators'

Firing at demonstrators does not imply affecting all or any of them. If one wanted to present them as affected, they would be ACC-case direct objects.

Finally, none of the noun phrases to be considered is a complement of other nominals—like 'grandma' in (7)—since these can neither be elided nor serve as ellipsis antecedents:

(7) дом бабушки
dom babuški

house$_{NOM}$ Grandma$_{GEN}$
'Grandma's house'

For purposes of building a theory of ellipsis, the point is this. For languages that have a highly developed system of case marking, there is much to be known about how case marking works. The more one understands beforehand or learns along the way, the easier it will to be to draw confident conclusions about the role of case marking in ellipsis.

3. Elided objects and their antecedents

For determining the ellipsis potential of an object, both the nature of the object and the nature of its antecedent are important. In principle, one could analyze every combination of case marking (function, case, source, meaning) for the potentially elided category with every combination for the antecedent. However, in practical terms, some principled shortcuts not only can but also probably should be taken, based on known facts about case marking in a language. For example, since quirky case marking is, by definition, semantically vacuous, it is possible that all instances of it—regardless of what case is ultimately assigned—behave the same with respect to their role in elliptical structures. A hypothesis like this, which speeds work at the outset, can be amended later if available data suggests the need for further refinement.

For Russian, a reasonable first-cut delineation of antecedents and potentially elided categories is as follows.

Antecedents

Direct object
Quirky oblique object
Indirect object
Object of a preposition
Nominative subject
Nominative nonsubject
Extralinguistic

Potentially Elided Objects

Direct object
Quirky case-marked oblique object
Indirect object

There are more types of antecedents than types of potentially elided objects because antecedents for object ellipsis can but need not themselves be objects.

Upcoming chapters will consider contexts that contain the most common combinations of antecedents and potentially elided objects: Chapter 3 considers DO

ellipsis with a DO antecedent; chapter 4, DO ellipsis with a NOM antecedent; chapter 5, DO ellipsis with an oblique antecedent; chapter 6, oblique ellipsis with any type of syntactically accessible antecedent; chapter 7, unexpressed objects that do not or may not represent ellipsis. Ellipsis with a pragmatically understood antecedent is considered briefly in chapter 13. Of course, it is not only case marking as such that is in question but also the syntactic structures and semantic roles that are reflected by the case marking and their combined effects on ellipsis potential.

4. Elided objects cross-linguistically: A snapshot of the literature

Object ellipsis has not attracted the same attention in the literature as the ellipsis of subjects or verbs. One exception is a flurry of interest in the realm of generative syntax in the mid-eighties to early nineties, starting with an article by Huang (1984) that treated the ellipsis of subjects and objects in Chinese. Huang proposed that all of the world's languages belong to one of two groups: those whose null arguments are licensed by the verb's agreement morphology (e.g., Italian, which permits subject ellipsis) and those that, lacking agreement, rely on null or overt discourse topics to license argument ellipsis (e.g., Chinese).[10] The notion of null discourse topic captures the intuition that one can elide only those arguments that reflect the discourse topic—informally, what the discourse is about.[11]

Subsequent studies of object ellipsis adopted or modified Huang's account when applied to other languages. For example, Raposo (1986: 380) follows Huang in analyzing null objects in European Portuguese as variables bound by a null operator; however, since the distribution of null objects in European Portugese shows constraints similar to those of *wh*-traces, he suggests that in this language null-object variables are created by movement of the empty category to the complementizer position of the root clause as opposed to being base generated along with a base-generated null-discourse topic. Suñer and Yépez (1988) adopt Huang's proposal unaltered, analyzing null direct objects in Quiteño Spanish as base-generated variables bound by base-generated topics. Bahan, Kegl, Lee, MacLaughlin, and Neidle (2000) argue that pro-drop in American sign language is licensed by agreement, whose sources include hand movements, head tilt, and eye gaze (this analysis differs from the one presented in Lillo-Martin 1986, 1991—which had posited different licensing strategies for "agreeing" and "non-agreeing" verbs—by incorporating non-manual sources of agreement).

All of these works provide at least a brief description of the syntax of object ellipsis in the respective languages, making them informative for building a theory of ellipsis. However, the study that arguably takes the most significant step in the direction that I am pursing is Rögnvaldsson's (1990) treatment of DO ellipsis in Modern Icelandic. In order for DO ellipsis to be possible in Modern Icelandic, the direct object must be located in the second conjunct of a coordinate structure, the subject of the second clause must be elided in addition to the object, and the antecedent must be the direct object of the preceding clause. Rögnvaldsson even opens up, but does not pursue, the case-marking issue, writing at the end of his paper (after

having talked about "objects" throughout and showing examples only of direct objects): "It appears to be much easier to drop accusative than dative or genitive objects" (Rögnvaldsson 1990: 377). For my purposes, there are four important aspects of DO ellipsis in Modern Icelandic:

1. The elided direct object must be in a coordinate structure that contains a syntactically overt object antecedent. The fact that Modern Icelandic permits DO ellipsis only within coordinate structures suggests that ellipsis within coordinate structures might be fundamentally different from ellipsis within other sentential configurations.
2. In order for DO ellipsis to occur, the ellipsis clause generally must have an elided subject as well. On the basis of such examples, one might suspect that DO ellipsis in Modern Icelandic is possible only in VP-coordinate structures, not in clausal coordinate structures. However, Rögnvaldsson explains that this would be an incorrect assumption, because there are instances in which Modern Icelandic same-subject clauses must be analyzed as clausal coordination. Thus, Modern Icelandic provides evidence for a same-subject versus different-subject parameter that affects DO ellipsis in coordinate structures.
3. DO ellipsis in Modern Icelandic requires that the antecedent be an object, which suggests that parallelism constraints between elided categories and their antecedents might play a role in ellipsis cross-linguistically.
4. Lexico-semantic factors affect DO ellipsis in Modern Icelandic; for example, there exist sentences in which a pronominal antecedent can support DO ellipsis, but an R-expression antecedent cannot.

The grain size of analysis carried out in the following chapters, when ultimately applied to many languages, will provide the raw data upon which to test the adequacy of the developing theory of ellipsis.

3

Direct Object Ellipsis with a Like Antecedent

The process of determining how different factors affect ellipsis involves holding certain factors stable while changing others, and seeking differences in ellipsis potential in minimal pairs. In this chapter, the control factors for studying object ellipsis in Russian are as follows:

- the potentially elided categories are direct objects with basic configurational (not quirky, lexical) case marking, which is ACC for Russian;
- all antecedents are, likewise, direct objects with configurational ACC case marking; and
- all antecedents are located in the clause that directly precedes the ellipsis clause, making them unquestionably syntactically accessible.

The parameters and their value sets relevant for this class of ellipsis are listed in table 3.1. They cover syntactic, lexico-semantic, and pragmatic factors that can affect ellipsis potential.

The status of direct objects ranges from "can't be elided" to "virtually must be elided," with many gradations between. For different applications, different approaches to object status might be appropriate. For example, an NLP system used to generate text for information analysts might be designed to produce overt objects in all instances, not seeking the stylistic refinement of variations on this theme. However, a system built to analyze text must be prepared for all instances of object ellipsis.

This chapter first closely analyzes Russian data, with the description organized at the top level according to syntactic structures. For each syntactic structure, the ellipsis effects of nonsyntactic factors are explored, like the nature of the antecedent

TABLE 3.1 Parameters and values for DO ellipsis with a DO antecedent

Parameters	*Values*
Syntactic structure	VP coordination Clausal coordination (the verbs have different subjects) A clause and its elaboration Main clause and gerund phrase Main clause and subordinate clause
Number of sequential clauses with coreferential objects	2, 3, 4, etc.
Nature of the antecedent	Referential expression Pronoun Unexpressed in surface text
Type of connector between clauses or conjuncts	Basic coordinating conjunction Contrastive coordinating conjunction Subordinating conjunction (each with different properties) Punctuation mark (each with different properties) None
Semantic correlation between referent of antecedent and referent of potentially elided category	Identical referent Generic versus specific Whole versus part 'Something' versus 'anything'
Nature of the ellipsis clause verb's selectional restrictions	Typical Narrow
Correspondence between selecting verb in antecedent and ellipsis clauses	Same verb Same root Of same semantic class Members of same ontological script None
Idiomaticity of verb + object collocation in potentially elided category	None Idiom or fixed
Stylistic force	Neutral Emphatic

(R-expression or pronoun) and the choice of lexical items. After the major syntactic structures are covered, some additional lexico-semantic and discourse factors that apply to all of them are discussed. Next, cross-linguistic comparisons are drawn that provide initial validation for the inventory of parameters and values. Finally, sample processing algorithms are presented, showing how the description deriving from this theory might be applied in practice.

1. Coordinate structures

Coordinate structures consist of two or more conjuncts that occur in series and are or at least can be joined by a conjunction like *and*, *or*, or *but*. In the following English sentences, the conjuncts are bracketed for clarity:

(1) a. He [grabbed her pigtail] and [pulled it].
b. He either [bought the pumpkin], [borrowed it], or [stole it].
c. I [made a nice dinner] but [didn't eat it].

In Russian coordinate structures that contain coreferential direct objects, the latter direct object(s) often can be and sometimes virtually must be elided. For example, whereas ellipsis is optional in (2), it is strongly preferred in (3). (The symbol ⫫ indicates that the variant is stylistically highly marked albeit not ungrammatical.)

(2) Он поднял **мяч** и бросил (**его**)**.**
On podnjal **mjač** i brosil (**ego**).
he$_{NOM}$ picked-up **ball**$_{ACC}$ and threw (**it**)$_{ACC}$
'He picked up the ball and threw it.'

(3) Он поднял **его** и бросил **∅/⫫его**.
On podnjal **ego** i brosil **∅/⫫ego**.
he$_{NOM}$ picked-up **it**$_{ACC}$ and threw **∅/⫫it**$_{ACC}$
'He picked it up and threw it.'

A number of factors contribute to determining DO ellipsis potential in coordinate structures. These factors can be thought of as parameters that have more ellipsis-promoting [+ ellipsis] and less ellipsis-promoting [(–) ellipsis] values, as shown in table 3.2. (I use [(–) ellipsis] rather than [– ellipsis] for the less ellipsis-promoting values in order not to imply that these values necessarily work against ellipsis; they simply do not promote it like the values marked [+ ellipsis]:

The information in this table can be recast as a series of logical generalizations. They are numbered according to the keys in the table.

1. The more coreferential direct objects there are in a sentence, the more desirable it is to elide one or more of them. For example, in English one would say *I saw **a firefly**, chased **it** around the yard, caught **it** in my hands, studied **it** carefully, then let **it** go.* Four overt *it*s in a row is inevitable in English and thus sounds fine. However, in Russian a corresponding degree of pronominal repetition would be considered excessive. Therefore, in sentences composed of three or more con-

TABLE 3.2 Parameters and values involved in determining DO ellipsis potential in coordinate structures

Key	*Parameter*	[+ *Ellipsis*]	[(–) *Ellipsis*]
1.	Number of conjuncts that contain the same DO	Three or more	Two
2.	Status of the conjunction	Overt	Absent
3.	Status of the antecedent	Pronoun	Referential expression
4.	Type of conjunct	VP	Clause (different-subject conjuncts)

juncts with coreferential direct objects, one or more of the objects should generally be elided. *Which* one(s) should be elided requires further analysis.

2. Overt conjunctions promote ellipsis by emphasizing the parallel nature of the conjuncts, since it is generally parallel elements that are joined by a coordinating conjunction (although see exceptions noted in chapter 1, section 3.6). All kinds of parallelism promote ellipsis in Russian.
3. There are at least two reasons that ellipsis is more strongly preferred when the antecedent is a pronoun than when it is an R-expression. The first reason is phonetic. Pronouncing two phonetically identical direct objects in close succession (*it . . . it*) sounds more repetitive than pronouncing two direct objects with different phonetic forms (*ball . . . it*). The second reason derives from properties of discourse: in order for you to refer to someone or something using a pronoun, that entity must either already be established in the language context or be visible/audible in the real-world context. Thus, a pronoun is merely a reminder of an already relevant and discourse-activated person or thing. When two pronominal reminders occur in close succession, the second tends to be elided in order to avoid undue repetition of the obvious.
4. When the subjects of the conjuncts are different, the second subject shifts the topic of discourse and decreases the expectation that the direct object will remain the same. Decreased expectation significantly impedes object-ellipsis potential in Russian.

1.1. Coordinated verb phrases

This section considers eight types of sentences that contain coordinated verb phrases (predicates). Each type is listed here showing the most common elliptical pattern for Russian. Summaries of the most common two-conjunct and three-conjunct patterns are found in tables 3.3 and 3.4. Further discussion and alternative ellipsis patterns are presented in the corresponding subsections to follow.

I. Он поднял **мяч** и бросил (**его**)**.**
On podnjal **mjač** i brosil (**ego**).
he$_{NOM}$ picked-up **ball**$_{ACC}$ and threw (**it**)$_{ACC}$
'He picked up the ball and threw it.'

II. Он поднял **его** и бросил ∅.
On podnjal **ego** i brosil ∅.
he$_{NOM}$ picked-up **it**$_{ACC}$ and threw ∅$_{ACC}$
'He picked it up and threw it.'

III. Он поднял **мяч**, бросил **его.**
On podnjal **mjač**, brosil **ego**.

he$_{NOM}$ picked-up **ball**$_{ACC}$ threw **it**$_{ACC}$
'He picked up the ball and threw it.'

IV. Он поднял **его**, бросил ∅.
On podnjal **ego**, brosil ∅.
he$_{NOM}$ picked-up **it**$_{ACC}$ threw ∅$_{ACC}$
'He picked it up and threw it.'

V. Он нашёл **мяч**, поднял (**его**) и бросил ∅.
On našël **mjač**, podnjal (**ego**) i brosil ∅.
he found **ball**$_{ACC}$ picked-up (**it**)$_{ACC}$ and threw ∅$_{ACC}$
'He found the ball, picked it up, and threw it.'

VI. Он нашёл **его**, поднял ∅ и бросил ∅.
On našël **ego**, podnjal ∅ i brosil ∅.
he$_{NOM}$ found **it**$_{ACC}$ picked-up ∅$_{ACC}$ and threw ∅$_{ACC}$
'He found it, picked it up, and threw it.'

VII. Он нашёл **мяч**, поднял **его**, бросил ∅.
On našël **mjač**, podnjal **ego**, brosil ∅.
he$_{NOM}$ found **ball**$_{ACC}$ picked-up **it**$_{ACC}$ threw ∅$_{ACC}$
'He found the ball, picked it up, and threw it.'

VIII. Он нашёл **его**, поднял ∅, бросил ∅.
On našël **ego**, podnjal ∅, brosil ∅.
he$_{NOM}$ found **it**$_{ACC}$ picked-up ∅$_{ACC}$ threw ∅$_{ACC}$
'He found it, picked it up, and threw it.'

Type I 2 conjuncts ◆ R-expr. ant. ◆ Overt conj. ◆ Optional DO

Он поднял **мяч** и бросил (**его**).
On podnjal **mjač** i brosil (**ego**).
he$_{NOM}$ picked-up **ball**$_{ACC}$ and threw (**it**)$_{ACC}$

Sentences of Type I are extremely common in Russian, and ellipsis of the second direct object is always optional:

TABLE 3.3 Two-conjunct patterns for DO ellipsis with a DO antecedent

No.	*Antecedent*	*Conjunction?*	*Coreferential NP*
I	Referential expression	Conjunction	(Pronoun)
II	Pronoun	Conjunction	∅
III	Referential expression		Pronoun
IV	Pronoun		∅

TABLE 3.4 Three-conjunct patterns for DO ellipsis with a DO antecedent

No.	*Antecedent*	*Coref. NP-1*	*Conjunction?*	*Coref. NP-2*
V	Referential expression	(Pronoun)	Conjunction	∅
VI	Pronoun	∅	Conjunction	∅
VII	Referential expression	Pronoun		∅
VIII	Pronoun	∅		∅

(4) С большим трудом преступник одолел следователя, вырвал у него **приказ на собственный арест** и спрятал ∅ к себе в карман (Брагинский и Рязанов: 155).
S bol'šim trudom prestupnik odolel sledovatelja, vyrval u nego **prikaz na sobstvennyj arest** i sprjatal ∅ k sebe v karman (Braginskij i Rjazanov: 155).
with great effort criminal$_{NOM}$ overpowered inspector$_{ACC}$ tore from him **order for self's arrest$_{ACC}$** and hid ∅$_{ACC}$ to self in pocket
'With great effort the criminal overpowered the inspector, grabbed the order for his own arrest and hid it in his pocket.'

Ellipsis is possible in sentences of Type I even if a parenthetical phrase separates the antecedent conjunct and the elliptical conjunct, as shown in (5):

(5) Она подняла **ладонь**, словно чтобы защитить волосы от ветра, но сразу же опустила ∅ — её стрижка лишала это движение всякого смысла (Пелевин: 105).
Ona podnjala **ladon'**, slovno čtoby zaščitit' volosy ot vetra, no srazu že opustila ∅ — eë strižka lišala èto dviženie vsjakogo smysla (Pelevin: 105).
she raised **palm$_{ACC}$** as-if in-order-to shield hair from wind but immediately PARTICLE lowered ∅$_{ACC}$ her haircut$_{NOM}$ deprived that movement$_{ACC}$ all$_{GEN}$ sense$_{GEN}$
'She held up her hand, as if to shield her hair from the wind, but immediately lowered it—her short haircut made that motion entirely unnecessary.'

Type II 2 conjuncts ◆ Pron. ant. ◆ Overt conj. ◆ Ellipsis preferred

Он поднял **его** и бросил ∅.
On podnjal **ego** i brosil ∅.
he$_{NOM}$ picked-up **it$_{ACC}$** and threw ∅$_{ACC}$

Sentences of Type II are also very common in Russian. Since *pronoun . . . pronoun* tends to sound overly repetitive, ellipsis of the second direct object is consistently preferred. In fact, all of the collected examples of this type show ellipsis.

(6) «Запомни: только об одном мы и думаем, только одна у нас цель и есть — освободить **тебя** и увезти ∅ с собой» (Шварц 1: 81).
«Zapomni: tol'ko ob odnom my i dumaem, tol'ko odna u nas cel' i est' — osvobodit' **tebja** i uvezti ∅ s soboj» (Švarc 1: 81).

remember$_{IMPER}$ only about one-thing we PARTICLE are-thinking only one at us goal PARTICLE is free$_{INFIN}$ **you$_{ACC}$** and take-away$_{INFIN}$ **∅$_{ACC}$** with self
‘‘“Remember, we’re only thinking about one thing, we have only one goal: to free you and take you away with us.”’

Type III 2 conjuncts ◆ R-expr. ant. ◆ No conj. ◆ Overt DO

Он поднял **мяч**, бросил **его.**
On podnjal **mjač**, brosil **ego**.
he$_{NOM}$ picked-up **ball$_{ACC}$** threw **it$_{ACC}$**

Sentences of Type III, which lack a coordinating conjunction, are somewhat less common in Russian than sentences of Types I or II. However, they are entirely possible and are less stylistically marked than the corresponding conjunctionless English sentences would be.[1] Conjunctionless coordination represents the same semantic relationship between conjuncts as regular coordination; there is simply no overt marker of this relationship.[2] In some conjunctionless structures, like (7), eliding the direct object sounds perfectly natural:

(7) «На!» — Мишка снял **пиджак**. Не глядя, протянул ∅ Таньке (Токарева 1994: 489).
«Na!» — Miška snjal **pidžak**. Ne gljadja, protjanul ∅ Tan’ke (Tokareva: 489).
here Mishka$_{NOM}$ took-off **jacket$_{ACC}$** NEG looking held-out **∅$_{ACC}$** Tan’ka$_{DAT}$
‘‘“Here!” Mishka took off his jacket. Not looking, he held it out to Tan’ka.’

However, in other conjunctionless structures ellipsis makes the utterance sound telegraphic or overly elliptical. Whether or not ellipsis will sound natural in a conjunctionless structure depends upon word selection, the overall context, and any number of other factors. For example, (8a), like our primary example for this pattern, sounds excessively elliptical if the direct object is elided; this is indicated by ⌋, which shows stylistic markedness. The sentence can be made stylistically neutral in one of three ways: by adding an overt conjunction and keeping the ellipsis (8b), by leaving out the conjunction but making the direct object overt (8c), or by adding both a conjunction and the direct object (8d).

(8) a. ⌋Рыцарь поднимает **меч**, протягивает ∅ королю.
b. Рыцарь поднимает **меч** и протягивает ∅ королю.
c. Рыцарь поднимает **меч**, протягивает **его** королю.
d. Рыцарь поднимает **меч** и протягивает **его** королю.
a. ⌋Rycar’ podnimaet **meč**, protjagivaet ∅ korolju.
b. Rycar’ podnimaet **meč** i protjagivaet ∅ korolju.
c. Rycar’ podnimaet **meč,** protjagivaet **ego** korolju.
d. Rycar’ podnimaet **meč** i protjagivaet **ego** korolju.
knight$_{NOM}$ picks-up **sword$_{ACC}$** and holds-out **it/∅$_{ACC}$** king$_{DAT}$
‘The knight picks up the sword and holds it out to the king.’

A comparison of stylistically neutral and stylistically marked sentences of Type III reveals the following tendency: presenting many actions in series increases the

likelihood that ellipsis will be stylistically acceptable, even if only two of the conjuncts in question contain coreferential direct objects. This is not unexpected, since ellipsis is generally promoted when conveying a series of actions (cf. chapter 7, section 9). For example, in (9)–(10) only two of the conjuncts in each sentence contain coreferential direct objects, but since these conjuncts occur within a longer series of actions ellipsis is acceptable:

(9) Казя снова на секунду умолкла, закрыла **глаза**, открыла Ø, вздохнула глубоко (Хмелевская 2: 263).
Kazja snova na sekundu umolkla, zakryla **glaza**, otkryla Ø, vzdoxnula gluboko (Chmielewska 2: 263).
Kazya$_{NOM}$ again for second fell-silent closed **eyes$_{ACC}$** opened Ø$_{ACC}$ sighed deeply
'Once again, Kazya fell momentarily silent, closed her eyes, opened them, and sighed deeply.'

(10) Взяла **ружьё** (отец был колхозным лесником), зарядила Ø двумя патронами, два положила в карман халата (Московские новости 1999).
Vzjala **ruž'ë** (otec byl kolxoznym lesnikom), zarjadila Ø dvumja patronami, dva položila v karman xalata (Moskovskie novosti 1999).
took$_{3SG.FEM}$ gun (father was kolkhoz forest-ranger) loaded Ø$_{ACC}$ two cartridges$_{INSTR}$ two$_{ACC}$ put in pocket robe$_{GEN}$
'She took the gun (her father was a forest ranger on the kolkhoz), loaded it with two cartridges, and put two more in the pocket of her robe.'

Type IV 2 conjuncts ◆ Pronominal ant. ◆ No conj. ◆ Ellipsis preferred

Он поднял **его**, бросил Ø.
On podnjal **ego**, brosil Ø.
he$_{NOM}$ picked-up **it$_{ACC}$** threw Ø$_{ACC}$

Type IV represents the same syntactic structure as Type III, but the antecedent is a pronoun. Because of the strong ellipsis-promoting power of pronominal antecedents, examples of this type tend to favor DO ellipsis, with the ellipsis creating no stylistic markedness:

(11) [Танька] подскочила к проигрывателю. Поставила **его** на подоконник. Включила Ø на полную мощность (Токарева: 482).
[Tan'ka] podskočila k proigryvatelju. Postavila **ego** na podokonnik. Vključila Ø na polnuju moščnost' (Tokareva: 482).
[Tan'ka]$_{NOM}$ raced to record player put$_{3.SG.FEM}$ **it**$_{ACC}$ on windowsill turned$_{3SG.FEM}$ Ø$_{ACC}$ to full power
'Tan'ka raced to the record player. She put it on the windowsill and turned it up full blast.'

Type V 3 conjuncts ◆ R-expr. ant. ◆ Overt conj. ◆ 2nd DO optional, 3rd elided

Он нашёл **мяч**, поднял (**его**) и бросил Ø.
On našël **mjač**, podnjal (**ego**) i brosil Ø.
he$_{NOM}$ found **ball** picked-up (**it**)$_{ACC}$ and threw Ø$_{ACC}$

In Russian structures of Type V, which contain three coordinated predicates with coreferential direct objects, at least one direct object is generally elided. Table 3.5 lists, in descending order of frequency, the possible combinations of overt and elided objects in three-clause structures with an R-expression antecedent. The last column indicates how widely applicable the given ellipsis pattern is.

By far the most common ellipsis pattern is (a), which is acceptable for all examples of this type. Pattern (b) is also generally possible, though not as common, and (c) may or may not be acceptable depending on the context. Pattern (d), however, tends to be undesirable, as it has three overt direct objects in a row. The following examples illustrate these patterns of ellipsis:

(12) . . . Он медленно снял **изорванный нанковый зипун**,
a. тщательно сложил **его** и повесил ∅ на спинку стула (Толстой 2: 38).
b. тщательно сложил ∅ и повесил ∅ на спинку стула.
c. тщательно сложил ∅ и повесил **его** на спинку стула.
d. ⌋ тщательно сложил **его** и повесил **его** на спинку стула.
. . . On medlenno snjal **izorvannyj nankovyj zipun**,
a. tščatel'no složil **ego** i povesil ∅ na spinku stula (Tolstoj 2: 38).
b. tščatel'no složil ∅ i povesil ∅ na spinku stula.
c. tščatel'no složil ∅ i povesil **ego** na spinku stula.
d. ⌋ tščatel'no složil **ego** i povesil **ego** na spinku stula.
he$_{NOM}$ slowly took **torn nankeen peasant's-coat$_{ACC}$** carefully folded **it/∅$_{ACC}$** and hung **it/∅$_{ACC}$** on back$_{ACC}$ chair$_{GEN}$
'He slowly took off the torn nankeen peasant's coat, carefully folded it and hung it on the back of the chair.'

Although pattern (a) is by far the most common, it is not always the first choice of speakers and authors, as is shown by (13) and (14), which employ patterns (b) and (c), respectively:

(13) . . . [Она] обхватила **брата** левою рукой за шею, быстро притянула ∅ к себе и крепко поцеловала ∅ (Толстой 1: 81).
. . . [Ona] obxvatila **brata** levoju rukoj za šeju, bystro pritjanula ∅ k sebe i krepko pocelovala ∅ (Tolstoj 1: 81).
[she]$_{NOM}$ grabbed **brother$_{ACC}$** left$_{INSTR}$ hand$_{INSTR}$ by neck quickly drew **∅$_{ACC}$** to self and hard kissed **∅$_{ACC}$**
'. . . She grabbed her brother by the neck with her left hand, quickly drew him toward her and kissed him hard.'

TABLE 3.5 Patterns of DO ellipsis in three-conjunct coordinate structures

Key	*Clause 1*	*Clause 2*	*Clause 3*	*Judgment*
a.	Referential expression	Pronoun	∅	Always possible
b.	Referential expression	∅	∅	Generally possible
a.	Referential expression	∅	Pronoun	Often possible
a.	Referential expression	Pronoun	Pronoun	Usually undesirable

(14) Привратники поднимают с травы **огромый блустящий ключ**, вкладывают ∅ в замочную скважину и поворачивают **его** в замке (Шварц 3: 536).
Privratniki podnimajut s travy **ogromyj blestjaščij ključ**, vkladyvajut ∅ v zamočnuju skvažinu i povoračivajut **ego** v zamke (Švarc 3: 536).
gatekeepers$_{NOM}$ pick-up from grass **huge shiny key$_{ACC}$** put ∅$_{ACC}$ in keyhold and turn ∅$_{ACC}$ in lock
'The gatekeepers take a huge shiny key from the grass, put it into the keyhole, and turn it in the lock.'

Finally, in some contexts one ellipsis pattern is strongly preferred over the others, usually based on rhythmic or semantic considerations. For example, in (15) the object must be overt in the second conjunct because if it were not, the conjunct would sound unfinished, leaving the listener waiting for its overt specification—a judgment that is practically impossible to capture by rules:

(15) Наконец Зигфрид нашёл **нужную ему книгу**, взял **её** с полки и принялся неторопливо ∅ перелистывать (Херриот: 33).
Nakonec Zigfrid našël **nužnuju emu knigu**, vjal **eë** s polki i prinjalsja netoroplivo ∅ perelistyvat' (Herriot: 33).
finally Siegfried$_{NOM}$ found **necessary$_{ACC}$ him$_{DAT}$ book$_{ACC}$** took **it$_{ACC}$** from shelf and started leisurely ∅$_{ACC}$ page-through$_{INFIN}$
'Finally Siegfried found the book he needed, took it from the shelf and started to leisurely page through it.'

Type VI 3 conjuncts ◆ Pronominal ant. ◆ Overt conj. ◆ Elide 2nd and 3rd DOs

Он нашёл **его**, поднял ∅ и бросил ∅.
On našel **ego**, podnjal ∅ i brosil ∅.
he$_{NOM}$ found **it$_{ACC}$** picked-up ∅$_{ACC}$ and threw ∅$_{ACC}$

When the antecedent in three-clause examples is a pronoun, ellipsis of the second and third direct objects is preferred, as would be expected considering the strong ellipsis-promoting power of pronominal antecedents, especially in multiconjunct structures:

(16) [About a registration document]
Лейтенант взял **её**, раскрыл ∅ и тщательно изучил ∅, посвечивая себе карманным фонариком (Азимов: 132).
Lejtenant vзjal **eë**, raskryl ∅ i tščatel'no izučil ∅, posvečivaja sebe karmannym fonarikom (Asimov: 132).
lieutenant$_{NOM}$ took **it$_{ACC}$** opened ∅$_{ACC}$ and carefully studied ∅$_{ACC}$ flashing self$_{DAT}$ flashlight$_{INSTR}$
'The lieutenant took it, opened it and carefully studied it, switching on his flashlight now and again.'

Type VII 3 conjuncts ◆ R-expr. ant. ◆ No conj. ◆ 2nd DO overt, 3rd elided

Он нашёл **мяч**, поднял **его**, бросил ∅.

On našël **mjač**, podnjal **ego**, brosil ∅
he$_{NOM}$ found **ball$_{ACC}$** picked-up **it$_{ACC}$** threw ∅$_{ACC}$

Sentences of Type VII present actions in series with no coordinating conjunction. The lack of a conjunction adds a narrative flavor and a heightened sense that the actions follow in succession:

(17) Тамара снимает с головы **платок**, встряхивает **его**, снова надевает ∅ (Московские новости, 1999).
Tamara snimaet s golovy **platok**, vstrjaxivaet **ego**, snova nadevaet ∅ (Moskovskie novosti, 1999).
Tamara$_{NOM}$ takes-off from head **kerchief$_{ACC}$** shakes **it$_{ACC}$** again puts-on ∅$_{ACC}$
'Tamara takes her kerchief off her head, shakes it out, and puts it on again.'

(18) «Во дворце сегодня праздник. Вы понимаете, какое великое дело — праздник! Порадовать **людей**, повеселить ∅, приятно удивить ∅ — что может быть величественнее?» (Шварц 3: 535).
«Vo dvorce segodnja prazdnik. Vy ponimaete, kakoe velikoe delo — prazdnik! Poradovat' **ljudej**, poveselit' ∅, prijatno udivit' ∅ — čto možet byt' veličestvennee?» (Švarc 3: 535).
at palace today holiday you understand what big deal holiday make-happy$_{INFIN}$ **people$_{ACC}$** amuse$_{INFIN}$ ∅$_{ACC}$ pleasantly surprise$_{INFIN}$ ∅$_{ACC}$ what can be more-stately
'"There's a holiday at the palace today. Do you understand what a big deal a holiday is? Making people happy, amusing them, giving them nice surprises—what can be more stately?"'

Since stylistic effects are always present in sentences of Type VII, there are no formal usage diagnostics—except, perhaps, the directive not to risk computer-generating this structure, although it must be prepared for in an analysis system.

Type VIII 3 conjuncts ◆ Pronominal ant. ◆ No conj. ◆ Elide the 2nd and 3rd DOs

Он нашёл **его**, поднял ∅, бросил ∅.
On našël **ego**, podnjal ∅, brosil ∅.
he$_{NOM}$ found **it$_{ACC}$** picked-up ∅$_{ACC}$ threw ∅$_{ACC}$

Type VIII is structurally just like Type VII, but the antecedent is a pronoun instead of an R-expression. As expected, ellipsis of all direct objects that follow the pronominal antecedent is preferred:

(19) Увидев Юрия Ивановича, он заулыбался, обнял **его**, повёл ∅ к кожаному креслу, усадил ∅ (Брагинский и Рязанов: 98–99).
Uvidev Jurija Ivanoviča, on zaulybalsja, obnjal **ego**, povël ∅ k kožanomu kreslu, usadil ∅ (Braginskij i Rjazanov: 98–99).
having-caught-sight-of Yurii Ivanovich he$_{NOM}$ smiled embraced **him$_{ACC}$** led ∅$_{ACC}$ to leather armchair seated ∅$_{ACC}$
'Catching sight of Yurii Ivanovich, he smiled, embraced him, led him over to a leather armchair, and sat him down.'

The most important insights that fall out of this analysis are how important it is to elide in certain configurations and how formally predictable many of those configurations are. As suggested at the beginning of this chapter, although one cannot say that syntactic ellipsis is ever strictly necessary *grammatically*, one *can* say that it is *pragmatically* necessary in certain configurations.

1.2. Coordinated clauses

Coordinated clauses, as defined here, must contain different subjects; if they had the same subject, they would be analyzed as coordinated verb phrases.[3] The new subject in the second clause shifts the topic of the discourse, thus decreasing the expectation that the direct object will remain the same. This, in turn, significantly impedes object-ellipsis potential. Ellipsis is permitted in coordinated clauses only if all of the following properties obtain:

- the conjunction is contrastive: *a/a* (mild contrast) or *но/no* (stronger contrast);
- the sentence is uttered with contrastive intonation; and
- there is a typical semantic relationship between clauses, determined in knowledge-rich text processing by ontological scripts.

Because of these restrictions, DO ellipsis in coordinated clauses is not particularly common. Therefore, rather than splitting the structures into the same subtypes as earlier, the discussion will be organized around broader issues and general tendencies.

The third prerequisite—that there be a typical semantic relationship between clauses—is actually an ellipsis-promoting factor language-wide, since the more predictable the context, the more readily recoverable any elided elements. The ellipsis-promoting force of the first two prerequisites derives from a different source: both a contrastive conjunction and contrastive intonation help to orient the listener with respect to the discourse structure, which in turn promotes ellipsis. For example, (20a) contains the coordinating conjunction *u/i* and blocks ellipsis, whereas (20b) contains the contrastive conjunction *a/a* (with assumed contrastive intonation) and permits ellipsis:

(20) a. Муж купил **картину**, и дочь повесила **её** на стену.
Muž kupil **kartinu**, i doč' povesila **eë** na stenu.
husband$_{NOM}$ bought **painting$_{ACC}$** and$_{COORD}$ daughter hung **it$_{ACC}$** on wall
'My husband bought a painting and my daughter hung it on the wall.'

b. Муж купил **картину**, а дочь повесила (**её**) на стену.
Muž kupil **kartinu**, a doč' povesila (**eë**) na stenu.
husband$_{NOM}$ bought **painting$_{ACC}$** and$_{CONTR}$ daughter hung (**it**)$_{ACC}$ on wall
'My husband bought a painting and my daughter hung it on the wall.'

The following examples, like (20b), permit DO ellipsis because of the interaction of a contrastive conjunction, contrastive intonation, and predictable semantics. Ellipsis is further promoted in (21) by the fact that the verbs in both clauses have the same root (*гад-/gad-*), which makes it clear that they are semantically related.

(21) Он мне **загадку** загадал, а я (**её**) разгадала.
On mne **zagadku** zagadal, a ja (**eë**) razgadala.
he_{NOM} me_{DAT} $\textbf{riddle}_{\textbf{ACC}}$ set and I $(\textbf{it})_{\textbf{ACC}}$ figured-out
'He set me a riddle and I figured it out.'

(22) «В любом случае завтра, нет, уже сегодня, сменю замок. Здесь этим занимается сторщж, куплю **что-нибудь сильно замысловатое**, а он ∅ поставит» (Хмелевская 2: 156).
«V ljubom slučae zavtra, net, uže segodnja, smenju zamok. Zdes' ètim zanimaetsja storož, kuplju **čto-nibud' sil'no zamyslovatoe**, a on ∅ postavit» (Chmielewska 2: 156).
{ . . . } $will\text{-}buy_{1SG}$ $\textbf{something fancy}_{\textbf{ACC}}$ and he $\emptyset_{\textbf{ACC}}$ will-install
'"In any case, tomorrow—no, today—I'll change the lock. Around here the guard takes care of that sort of thing. I'll buy something fancy and he'll install it."'

As always, a pronominal antecedent further promotes ellipsis potential:

(23) Вдруг Инна полюбит другого и уйдёт? Тогда тесть вышвырнет **его** с дачи, а неверная жена выкинет ∅ на ходу из машины (Брагинский и Рязанов: 73).
Vdrug Inna poljubit drugogo i ujdët? Togda test' vyšvyrnet **ego** s dači, a nevernaja žena vykinet ∅ na xodu iz mašiny (Braginskij i Rjazanov: 73).
suddenly $Inna_{NOM}$ will-fall-in-love-with another and will-leave then $father\text{-}in\text{-}law_{NOM}$ $kick\text{-}out_{FUT}$ $\textbf{him}_{\textbf{ACC}}$ from summer-house and unfaithful $wife_{NOM}$ $chuck_{FUT}$ $\emptyset_{\textbf{ACC}}$ on road from car
'What if Inna fell in love with someone else and left him? Then his father-in-law would kick him out of the summer house, and his unfaithful wife would chuck him out of a moving car.'

In all of the previous examples, the conjunction is contrastive, as it must be. But what if there is no overt conjunction? In this case, the necessary contrastive nature of the clauses can be harder to convey and interpret, which works against ellipsis. The following minimal pairs show the general tendency for DO ellipsis to be rejected if clauses are coordinated without an overt conjunction. The (a) variants are conjunctionless structures that impede DO ellipsis; the (b) variants contain an overt contrastive conjunction and permit ellipsis. (All of these sentences would be grammatical and stylistically neutral if the direct object were overt)

(24) a. (?) Я сняла с него **плащ**; Мария повесила ∅ на вешалку.
(?) Ja snjala s nego **plašč**; Marija povesila ∅ na vešalku.
I_{NOM} took from him_{GEN} $\textbf{raincoat}_{\textbf{ACC}}$ $Mary_{NOM}$ hung $\emptyset_{\textbf{ACC}}$ on hanger
'I took his raincoat. Mary hung it on a hanger.'[4]

b. Я сняла с него **плащ**, а Мария повесила ∅ на вешалку.
Ja snjala s nego **plašč**, a Marija povesila ∅ na vešalku.
I_{NOM} took from him_{GEN} $\textbf{raincoat}_{\textbf{ACC}}$ and_{CONTR} $Mary_{NOM}$ hung $\emptyset_{\textbf{ACC}}$ on hanger
'I took his raincoat and Mary hung it on a hanger.'

(25) a. * Я случайно выпустила **птицу**, брат ∅ поймал.
* Ja slučajno vypustila **pticu**, brat ∅ pojmal.

I_{NOM} accidentally let-out **bird$_{ACC}$** brother$_{NOM}$ **∅$_{ACC}$** caught
'I accidentally let out the bird, but my brother caught it.'

b. Я случайно выпустила **птицу**, но брат **∅** поймал.
Ja slučajno vypustila **pticu**, no brat **∅** pojmal.
I_{NOM} accidentally let-out **bird$_{ACC}$** but brother$_{NOM}$ **∅$_{ACC}$** caught
'I accidentally let out the bird, but my brother caught it.'

The reason that (25a) is ungrammatical, whereas (24a) is merely infelicitous, must derive from a combination of pragmatics and semantics, since without the conjunction, the clauses in (25) seem oddly juxtaposed.

The lack of a conjunction does not, however, guarantee that ellipsis will be blocked. If the intended contrast is emphasized lexically and intonationally, ellipsis may be possible, as in the following examples:

(26) Он мне **загадку** загадал — я разгадала **∅**.
On mne **zagadku** zagadal — ja razgadala **∅**.
he$_{NOM}$ me$_{DAT}$ **riddle$_{ACC}$** set I_{NOM} guessed **∅$_{ACC}$**
'He set me a riddle and I figured it out.'

(27) Я уложила **ребёнка** — бабушка убаюкала **∅**.
Ja uložila **rebënka** — babuška ubajukala **∅**.
I_{NOM} put-to-bed **baby$_{ACC}$** grandma$_{NOM}$ lulled-to-sleep **∅$_{ACC}$**
'I put the baby to bed and Grandma lulled him to sleep.'

Sentences composed of three coordinated clauses with coreferential direct objects are not very common—a matter of pragmatics, not grammar. However, such examples are possible and permit ellipsis as long as the three conditions listed at the outset of this subsection are met: the conjunction is contrastive, the sentence is utterered with contrastive intonation, and the semantic correlation is typical:

(28) Я выкупала **ребёнка**, мама убаюкала (**его**), а бабушка положила **∅** спать.
Ja vykupala **rebënka**, mama ubajukala (**ego**), a babuška položila **∅** spat'.
I_{NOM} bathed **baby$_{ACC}$** Mom$_{NOM}$ lulled-to-sleep **him$_{ACC}$** and Grandma$_{NOM}$ put-to-bed **∅$_{ACC}$**
'I bathed the baby, Mom lulled him to sleep, and Grandma put him to bed.'

(29) Я выкупала **его**, мама убаюкала **∅**, а бабушка положила **∅** спать.
Ja vykupala **ego**, mama ubajukala **∅**, a babuška položila **∅** spat'.
I_{NOM} bathed **him$_{ACC}$** Mom$_{NOM}$ lulled-to-sleep **∅$_{ACC}$** and$_{CONTR}$ Grandma$_{NOM}$ put-to-bed **∅$_{ACC}$**
'I bathed him, Mom lulled him to sleep, and Grandma put him to bed.'

2. The Assertion and Elaboration Strategy

The Assertion and Elaboration (A and E) Strategy occurs in what I call clause complexes (i.e., juxtaposed clauses), in which the first clause asserts something and

the second clause explains, embellishes, or otherwise comments upon it, as in example (30):

(30) «А я за эти дни прочла **кучу американских детективов** — брала ∅ в библиотеке», — сказала мисс Марпл (Кристи: 199).
«A ja za èti dni pročla **kuču amerikanskix detektivov** — brala ∅ v biblioteke», — skazala miss Marpl (Christie: 199).
and$_{\text{CONTR}}$ I$_{\text{NOM}}$ over these days read$_{\text{3SG.FEM}}$ **heap$_{\text{ACC}}$ American$_{\text{GEN}}$ mysteries$_{\text{GEN}}$** took$_{\text{3SG.FEM}}$ **∅$_{\text{ACC}}$** in library said Miss Marple$_{\text{NOM}}$
'"And during this time I read lots of American mysteries that I borrowed from the library," said Miss Marple.'

The term *clause complex* is more accurate here than *sentence* because A and E structures can be written in various ways: as one sentence with its clauses separated by a comma, colon, or dash or as two sentences.

Unlike coordinate structures, the component parts of an A and E structure are not of equal status—the second is semantically and functionally subordinate to the first. It is tacked on in order to explain, further comment upon, or provide a continuation for what was just said. In most instances, the components of an A and E Strategy are semantically incompatible with a coordinating conjunction: for example, one could not rewrite (30) as **I read lots of American mysteries and I borrowed them from the library.* The A and E Strategy strongly promotes DO ellipsis in Russian because an elaboration, by definition, must be about the same thing as the preceding assertion; therefore, further references to that thing can readily be elided.

Four aspects of the A and E Strategy distinguish it from the coordinate structures discussed earlier:

1. The components of an A and E Strategy will always be analyzed as clauses, never as predicates, because when the second subject is coreferential with the first, it can sometimes be repeated overtly (not so for coordinate structures in Russian).
2. No conjunction ever joins the clauses.
3. The clauses can be uttered by a single speaker or cooperatively by two speakers, which does not affect DO ellipsis potential.
4. The clauses can have the same or different subjects, which does not affect DO ellipsis potential.

A and E Strategies can be categorized based on the properties in the third and fourth aspects: the clauses be uttered by one or two speakers, and they can have the same or different subjects. All four permutations of these factors produce ellipsis-promoting structures, as shown by (31)–(34):

(31) «Я имел подлость убить сегодня **эту чайку**. Кладу ∅ у ваших ног» (Чехов 2: 412).
«Ja imel podlost' ubit' segodnja **ètu čajku**. Kladu ∅ u vašix nog» (Čexov 2: 412).

I_{NOM} had baseness $kill_{INFIN}$ today **this seagull$_{ACC}$** lay_{1SG} **∅$_{ACC}$** at your feet
‘‘“Today I was so base as to kill this seagull. I lay it at your feet.”’

(32) «Я уже говорил, у меня в лаборатории работает приятель, специалист классный, вот **эти фотографии** сделал, я только что получил **∅**»
(Хмелевская 4: 137).
«Ja uže govoril, u menja v laboratorii rabotaet prijatel’, specialist klassnyj, vot **èti fotografii** sdelal, ja tol’ko čto polučil **∅**» (Chmielewska 4: 137).
I_{NOM} already said at me in lab works $friend_{NOM}$ $specialist_{NOM}$ $top\text{-}notch_{NOM}$ here **these photographs$_{ACC}$** $made_{3SG.MASC}$ I_{NOM} just received **∅$_{ACC}$**
‘‘“As I already said, I have a friend who works in a lab, he’s a top-notch expert and he developed these photographs, I just got them.”’

(33) «Я сьела **три пирожка».** — «Сама **∅** испекла?»
«Ja s”ela **tri pirožka**». — «Sama **∅** ispekla?»
I_{NOM} $ate_{SG.FEM}$ **three$_{ACC}$** **pierogis$_{GEN}$** $self_{NOM}$ **∅$_{ACC}$** $baked_{SG.FEM}$
‘‘“I ate three pierogis.” “Did you bake them yourself?”’

(34) «Не знаю, что с вами делать. Я уже **ключи** отдала уборщице». — «Я сейчас **∅** возьму», — сказал Гошка и, не дожидаясь ответа, побежал искать уборщицу (Войнович 1: 32–33).
«Ne znaju, što s vami delat’. Ja uže **ključi** otdala uborščice». — «Ja sejčas **∅** voz’mu», — skazal Goška i, ne dožidajas’ otveta, pobežal iskat’ uborščicu (Vojnovič 1: 32–33).
NEG $know_{1SG}$ what with you_{INSTR} do_{INFIN} I_{NOM} already **keys$_{ACC}$** returned $cleaning\text{-}lady_{DAT}$ I_{NOM} now **∅$_{ACC}$** will-take said Goshka and NEG waiting-for answer ran $search\text{-}for_{INFIN}$ $cleaning\text{-}lady_{ACC}$
‘‘“I don’t know what to do with you. I already gave the keys back to the cleaning lady.” “I’ll run and get them,” said Goshka, and, not waiting for a reply, ran to find the cleaning lady.’

To reiterate, the A and E Strategy consistently promotes DO ellipsis because it first presents some topic of conversation (in the assertion), then further comments upon it (in the elaboration). Since an elaboration, by definition, must have the same topic of conversation as the assertion, that topic (in these examples, the direct object) can readily be elided. The challenge, however, lies in automatically detecting A and E configurations. They must be distinguished, on the one hand, from asyndetic (i.e., conjunctionless) coordination and, on the other, from clauses that are in series with no close semantic or pragmatic tie.

3. A gerund phrase is involved

Russian sentences can be composed of a main clause and a gerund phrase—that is, an indeclinable form of the verb like *читая/čitaja* ‘while reading’ or *посмотрев/*

posmotrev 'having looked at.'[5] The main clause and gerund phrase can occur in either order, but the order crucially affects ellipsis potential of the latter direct object.

3.1. Gerund phrase + main clause

It is not possible to elide a main-clause direct object whose antecedent is located in the preceding gerund phrase. Impressionistically, this can be accounted for by saying that the gerund phrase is the weaker component of the sentence, so an antecedent located there is not prominent enough to license ellipsis in the more powerful main clause:

(35) Разорвав **телеграмму**, он прочёл **её**, догадкой поправляя перевранные, как всегда, слова, и лицо его просияло (Толстой 1: 7).
Razorvav **telegrammu**, on pročël **eë**, dogadkoj popravljaja perevrannye, kak vsegda, slova, i lico ego prosijalo (Tolstoy 1: 7).
having-ripped-open **telegram**$_{\mathbf{ACC}}$ he read **it**$_{\mathbf{ACC}}$ with-a-guess correcting garbled as always words and face his lit-up
'Having ripped open the telegram, he read it, guessing at the words—garbled as always—and his face lit up.'

(36) Подняв **окурок**, он бросил **его** в урну.
Podnjav **okurok**, on brosil **ego** v urnu.
having-picked-up **cigarette-butt**$_{\mathbf{ACC}}$ he$_{\mathrm{NOM}}$ threw **it**$_{\mathbf{ACC}}$ in trash-can
'Having picked up the cigarette butt, he threw it into the trash can.'

3.2. Main clause + gerund phrase

When the sentence is composed of a main clause followed by a gerund phrase, the direct object in the gerund phrase can often be elided. Using the same impressionist reasoning suggested earlier, we can say that an antecedent located in the main clause is prominent enough to license ellipsis in the subsequent weaker gerund phrase:

(37) Я вынул **пистолет** из кармана и положил **его** на кровать, накрыв ∅ одеялом (Пелевин: 177).
Ja vynul **pistolet** iz karmana i položil **ego** na krovat', nakryv ∅ odejalom (Pelevin: 177).
I$_{\mathrm{NOM}}$ took **pistol**$_{\mathbf{ACC}}$ from pocket$_{\mathrm{GEN}}$ and put **it**$_{\mathbf{ACC}}$ on bed having-covered ∅$_{\mathbf{ACC}}$ blanket$_{\mathrm{INSTR}}$
'I took the pistol out of my pocket and put it on the bed, covering it with a blanket.'

(38) . . . Тереса ругала **нас** на чем свет стоит, обзывая ∅ дурами безмозглыми, тупицами безнадежными, идиотками законченными . . . (Хмелевская 5: 124).
. . . Teresa rugala **nas** na čem svet stoit, obzyvaja ∅ durami bezmozglymi, tupicami beznadežnymi, idiotkami zakončennymi . . . (Chmielewska 5: 124).

Teresa$_{NOM}$ balled-out **us$_{ACC}$** like crazy (idiom) calling **∅$_{ACC}$** fools witless dimwits hopeless idiots consummate
'. . . Teresa really socked it to us, calling us witless fools, hopeless dimwits, and consummate idiots . . .'

Some details of these examples deserve comment. First, they are presented without preceding context because if, for example, the given direct object occurred in the preceding sentence, intersentential effects could affect ellipsis judgments. Second, the surface ordering of the antecedent and potentially elided direct object is always the same, with the antecedent occurring first in the sentence. This is because Russian, unlike English, does not permit postcedents. That is, while in English one can say *Having checked **it** over, he gave his professor **the essay***, this is impossible in Russian, as shown by (39a–b). (This is, by the way, another parameter value that must be established for each language.) Example (39c) shows an acceptable elliptical example with the same meaning.

(39) a. * Не проверив **его**, он отдал профессору **эссе**.
* Ne proveriv **ego**, on otdal professoru **èsse**.
NEG having-checked **it$_{ACC}$** he$_{NOM}$ handed-in professor$_{DAT}$ **essay$_{ACC}$**
'Without checking it over, he handed the essay in to his professor.'

b. * Не проверив **∅**, он отдал профессору **эссе**.
* Ne proveriv **∅**, on otdal professoru **èsse**.
NEG having-checked **∅$_{ACC}$** he$_{NOM}$ handed-in professor$_{DAT}$ **essay$_{ACC}$**
'Without checking it over, he handed the essay in to his professor.'

c. Он отдал профессору **эссе**, не проверив **(его).**
On otdal professoru **èsse**, ne proveriv (**ego**).
he$_{NOM}$ handed-in professor$_{DAT}$ **essay$_{ACC}$** NEG having-checked (**it**)$_{ACC}$
'He handed the essay in to his professor, without checking it over.'

4. Subordinate structures

Subordinate structures present challenges for analysis because, depending on the conjunction employed, there can be many different semantic correspondences between the main and subordinate clause, which in turn can affect DO-ellipsis potential. This section underscores several of the most salient factors that affect ellipsis in such structures, making no attempt to exhaust the topic.

As with any "unbalanced" syntactic structure, DO-ellipsis potential in sentences composed of a main clause and a subordinate clause depends in large part upon the ordering of the clauses. For example, several temporal conjunctions in Russian, as well as the conjunctions *чтобы/čtoby* 'to' and *для того*, *чтобы/dlja togo*, *čtoby* 'in order to', permit DO ellipsis when the clause order is [main + subordinate], as in the (a) variants here, but block ellipsis when the clause order is [subordinate + main], as in the (b) variants:

(40) a. Она внимательно перечитала **письмо**, перед тем как <прежде чем> вложить (**его**) в конверт.
Ona vnimatel'no perečitala **pis'mo**, pered tem kak <prežde čem> vložit' (**ego**) v konvert.
she$_{NOM}$ carefully reread **letter**$_{ACC}$ before put$_{INFIN}$ (**it**)$_{ACC}$ in envelope
'She carefully reread the letter before putting it in the envelope.'

b. Перед тем как <прежде чем> вложить **письмо** в конверт, она внимательно перечитала **его.**
Pered tem kak <prežde čem> vložit' **pis'mo** v konvert, ona vnimatel'no perečitala **ego**.
before put$_{INFIN}$ **letter**$_{ACC}$ in envelope she$_{NOM}$ carefully reread **it**$_{ACC}$
'Before putting the letter in the envelope, she carefully reread it.'

(41) a. Отнеси директору **письмо**, как только отредактируешь (**его**).
Otnesi direktoru **pis'mo**, kak tol'ko otredaktirueš' (**ego**)
take$_{IMPER}$ director$_{DAT}$ **letter**$_{ACC}$ as soon as will-edit$_{2SG}$ (**it**)$_{ACC}$
'Take the director the letter as soon as you've edited it.'

b. Как только он отредактировал **письмо**, он отнес **его** директору.
Kak tol'ko on otredaktiroval **pis'mo**, on otnes **ego** direktoru.
as soon as he$_{NOM}$ edited **letter**$_{ACC}$ he$_{NOM}$ took **it**$_{ACC}$ director$_{DAT}$
'As soon as he edited the letter, he took it to the director.'

(42) a. Надо отвезти **детей** к морю, для того, чтобы (**их**) оздоровить.
Nado otvezti **detej** k morju, dlja togo, čtoby (**ix**) ozdorovit'.[6]
necessary$_{IMPERS}$ take$_{INFIN}$ **children**$_{ACC}$ to ocean in-order to (**them**)$_{ACC}$ make-healthy$_{INFIN}$
'You <we, etc.> need to take the children to the ocean to get them healthy.'

b. Для того, чтобы оздоровить **детей**, надо отвезти **их** к морю.
Dlja togo, čtoby ozdorovit' **detej**, nado otvezti **ix** k morju.
in-order to make-healthy$_{INFIN}$ **children**$_{ACC}$ necessary$_{IMPERS}$ take$_{INFIN}$ **them**$_{ACC}$ to ocean
'In order to get the children healthy, you <we, etc.> need to take them to the ocean.'

One conjunction that does not impose this clause-ordering restriction is *esli* 'if', which is discussed separately later. This fact that subordinating conjunctions differ with respect to clause-ordering restrictions on ellipsis supports the hypothesis that the semantics of subordination play a role in determining ellipsis potential.

Apart from clause order, another factor that affects DO-ellipsis potential in subordinate configurations is whether the clauses have the same or different subjects. In Russian sentences that contain a temporal subordinate conjunction, DO ellipsis is possible if the clauses have the same subject, as shown in the (a) variants of (43)–(45), but impossible if they have different subjects, as shown by the (b) variants:

(43) a. Я проверила **эссе**, перед тем как сдать (**его**).
Ja proverila **èsse**, pered tem kak sdat' (**ego**).

I_{NOM} checked-over **essay**$_{ACC}$ before hand-in$_{INFIN}$ (**it**)$_{ACC}$
'I checked over the essay before handing it in.'

b. Я проверила **эссе**, перед тем как Валя сдала **его**.
Ja proverila **èsse**, pered tem kak Valja sdala **ego**.
I_{NOM} checked-over **essay**$_{ACC}$ before Valja$_{NOM}$ handed-in **it**$_{ACC}$
'I checked over the essay before Valja handed it in.'

(44) a. Он несколько раз повторил **стихотворение**, пока не запомнил (**его**).
On neskol'ko raz povtoril **stixotvorenie**, poka ne zapomnil (**ego**).
he_{NOM} several times repeated **poem**$_{ACC}$ before remembered (**it**)$_{ACC}$
'He repeated the poem several times before he could remember it.'

b. Он несколько раз повторил стихотворение, пока девочка не запомнила **его**.
On neskol'ko raz povtoril stixotvorenie, poka devočka ne zapomnila **ego**.
he_{NOM} several times repeated **poem**$_{ACC}$ before girl$_{NOM}$ remembered **it**$_{ACC}$
'He repeated the poem several times before the girl could remember it.'

(45) a. Отнеси директору **письмо**, как только отредактируешь (**его**)**.**
Otnesi direktoru **pis'mo**, kak tol'ko otredaktirueš' (**ego**).
bring$_{IMPER}$ director$_{DAT}$ **letter**$_{ACC}$ as soon as will-edit$_{2SG}$ (**it**)$_{ACC}$
'Bring the letter to the director as soon as you've edited it.'

b. Отнеси директору **письмо**, как только Анна **его** отредактирует.
Otnesi direktoru **pis'mo**, kak tol'ko Anna **ego** otredaktiruet.
bring$_{IMPER}$ director$_{DAT}$ **letter**$_{ACC}$ as soon as Anna$_{NOM}$ will-edit$_{3SG}$ **it**$_{ACC}$
'Bring the letter to the director as soon as Anna has edited it.'

However, here again, semantics plays a role: whereas *temporal* conjunctions impose this same-subject requirement for DO ellipsis, the conjunctions *čtoby* 'so that' and *dlja togo*, *čtoby* 'so that' do not, permitting ellipsis even when the subject shifts:

(46) Я дала брату утюг, чтобы он (**его**) починил.
Ja dala bratu **utjug**, čtoby on (**ego**) počinil.
I_{NOM} gave brother$_{DAT}$ **iron**$_{ACC}$ so-that he_{NOM} (**it**)$_{ACC}$ fix
'I gave my brother the iron to fix.'

(47) Он отдал **письмо** Петру для того, чтобы он отнес (**его**) на почту.
On otdal **pis'mo** Petru dlja togo, čtoby on otnes (**ego**) na počtu.
he_{NOM} gave **letter**$_{ACC}$ Peter$_{DAT}$ so that he_{NOM} took (**it**)$_{ACC}$ to post-office
'He gave the letter to Peter to take to the post office.'

Another factor that can affect DO-ellipsis potential is the mood of the verb that selects the direct object. When the ellipsis-clause verb is in the imperative, DO ellipsis is often possible in configurations that would generally block it on syntactic grounds. This is, it would appear, because imperatives require a close

speaker–interlocutor relationship that promotes ellipsis on pragmatic grounds. A comparison of (41b) and (41c) serves as evidence of this generalization.

(41) c. Как только отредактируешь **письмо**, отнеси (**его**) директору на подпись.
Kak tol'ko otredaktirueš' **pis'mo**, otnesi (**ego**) direktoru na podpis'.
as soon as will-edit$_{2SG}$ **letter$_{ACC}$** bring$_{IMPER}$ (**it**)$_{ACC}$ director$_{DAT}$ for signature
'As soon as you edit the letter, bring it to the director for his signature.'

The subordinating conjunction *если/esli* 'if' deserves special mention because it has a number of properties that make sentences that contain it particularly conducive to DO ellipsis:

1. Semantic Predictability

Sentences that contain an *if*-clause have a higher degree of semantic predictability than many other types of sentences because the clause that presents the *if*-condition must be followed by a clause that presents the outcome if that condition were (or were not, for negatives) met. When the condition refers to some direct object, the result of that condition generally refers to the same direct object. This predictability factor facilitates DO ellipsis in the second clause.

(48) «[Слуги] говорят, что запомнили бы **ее**, если бы видели ∅ в окрестностях» (Кристи: 252).
«[Slugi] govorjat, čto zapomnili by **ee**, esli by videli ∅ v okrestnostjax» (Christie: 252).
[servants]$_{NOM}$ say that remember CONDIT **her$_{ACC}$** if CONDIT saw ∅$_{ACC}$ in area
'"The servants say that they'd recognize her if they saw her around."'

(49) Если бы **его** поймали, то обязательно бы посадили ∅ в тюрьму.
Esli by **ego** pojmali, to objazatel'no by posadili ∅ v tjur'mu.
if CONDIT **him$_{ACC}$** caught$_{3PL}$ then definitely CONDIT throw$_{3PL}$ ∅$_{ACC}$ in jail
'If they caught him, they'd definitely throw him in jail.'

2. Imperative Verb

The verb in the clause that follows the *if*-clause is often in the imperative, which promotes ellipsis based on the requisite momentary common scope of interest of the speaker and hearer:

(50) Если он тебе **пришлет любовное** письмо выбрось ∅ сразу.
Esli on tebe prišlet **ljubovnoe pis'mo** vybros' ∅ srazu.
if he$_{NOM}$ you$_{DAT}$ sends **love letter$_{ACC}$** throw-out ∅$_{ACC}$ immediately
'If he sends you a love letter, throw it out immediately.'

3. Verbal Repetition

Often both clauses in *if*-sentences employ the same verb. This strongly promotes and sometimes practically requires DO ellipsis in the second clause:

(51) Ведь на самом-то деле, если очень хочешь найти внутри учреждения **такие резервы**, всегда найдёшь ∅ (Богуславская: 135).
Ved' na samom-to dele, esli očen' xočeš' najti vnutri učreždenija **takie rezervy**, vsegda najdëš' ∅ (Boguslavskaja: 135).
after-all really (idiom) if very-much want$_{2SG}$ find$_{IMPER}$ inside organization **such reserves$_{ACC}$** always will-find$_{2SG}$ ∅$_{ACC}$
'After all, if you really want to find such reserves within an organization, you can always find them.'

5. More lexico-semantic and pragmatic factors

In the previous subsections, ellipsis potential was considered based primarily on the syntactic structure in which the antecedent and potentially elided category appeared. However, lexico-semantic and pragmatic factors were incorporated as needed to explain different ellipsis potential in minimal pairs. This section explores more such factors that can affect ellipsis potential in any syntactic structure.

5.1. Referent mismatches

In some instances, a direct object and its antecedent do not precisely match in referent: for example, *I like **Ferraris*** [a type of object] *and bought **one*** [a specific instance]*; My neighbor was selling **tomatoes*** [quantity *x*] *and I bought **some*** [some portion of *x*]. The English anaphors *one* and *some* are both lexically specified as nondefinite and differ only with respect to discreteness: *one* is discrete, whereas *some* is nondiscrete.[7] Russian, however, lacks nondefinite anaphors, leading to the question of how to express the second object in sentences like the previous ones. As the following subsections show, eliding the object tends to be the preferred strategy.

The antecedent expresses a class and the following direct object reflects an instance of it

In such contexts, English generally uses *one* to convey the instance whereas Russian requires that it be elided, since Russian has no anaphors specified as [nondefinite, discrete] (*один/odin* 'one' cannot be used this way):

(52) «Может быть, вы купите **календарь**? Нет? Вы уже купили ∅?»
«Možet byt', vy kupite **kalendar**'? Net? Vy uže kupili ∅?»
maybe you$_{NOM}$ will-buy **calendar$_{ACC}$** no you$_{NOM}$ already bought ∅$_{ACC}$
'"Perhaps you'd like to buy a calendar? No? You already bought one?"'

Minimal pair (53a–b) shows how DO-ellipsis potential differs when there is an exact versus an inexact correspondence between the antecedent and the following direct object. Both examples are composed of the same words, but let us assume that contextual cues make it clear that the antecedent in (53a) is some specific covered bridge, while the antecedent in (53b) is a generalized covered bridge.

(53) a. [Assume *the* covered bridge]
« . . . Я ищу **крытый мост**, но никак не могу найти (**его**)» (Уоллер: 42).[8]
« . . . Ja išču **krytyj most**, no nikak ne mogu najti (**ego**)» (Waller: 42).
I_{NOM} look-for **covered bridge**$_{ACC}$ but in-no-way NEG can find$_{INFIN}$ (**it**)$_{ACC}$
‘“. . . I am looking for the covered bridge and just can’t seem to find it.”’

b. [Assume *a* covered bridge]
«Я ищу **крытый мост**, но никак не могу найти ∅».
«Ja išču **krytyj most**, no nikak ne mogu najti ∅».
I look-for **covered bridge**$_{ACC}$ but in-no-way NEG can find$_{INFIN}$ ∅$_{ACC}$
‘“I’m looking for a covered bridge and just can’t seem to find one.”’

In (53a) DO ellipsis is optional because there is an exact correspondence between the antecedent and the subsequent direct object, making possible the overt realization of *его/ego* ‘it’. In (53b) DO ellipsis is mandatory because there is an inexact correspondence between the antecedent and the subsequent direct object, so the direct object cannot be realized by *его/ego* ‘it’.

The antecedent expresses a mass or quantity and the following direct object reflects some portion of it

In such contexts, English generally uses *some* to convey the portion, whereas Russian generally elides it, as shown in (54):

(54) «**Туфельки** ты и без волшебника получишь . . . Поедешь с папой на ярмарку, он и купит ∅ . . . » (Волков: 6).
«**Tufel’ki** ty i bez volšebnika polučiš’ . . . Poedeš’ s papoj na jarmarku, on i kupit ∅ . . . » (Volkov: 6).
shoes$_{ACC}$ you$_{NOM}$ even without wizard will-receive$_{2SG}$ will-go$_{2SG}$ with Dad to market he$_{NOM}$ PARTICLE will-buy ∅$_{ACC}$
‘“You’ll get shoes even without the help of a wizard . . . You’ll go to the market with your dad and he’ll buy you some.”’

Although words for *some* exist in Russian, they are not fully equivalent in usage to English *some* and are not used in contexts like (54).

The antecedent and the following direct object express indefinites that, if overt, would need to be conveyed by different expressions

For example, in (55) the English correspondence is *something/anything*, whereas the preferred Russian correspondence is *koe-čto/* ∅:

(55) «Сдается мне, ты **кое-что** знаешь, Роза?» — «Я? Да Бог с вами, сэр. Честное благородное слово! Откуда мне знать ∅?» (Кристи: 141).
«Sdaetsja mne, ty **koe-čto** znaeš’, Roza?»—«Ja? Da Bog s vami, sèr. Čestnoe blagorodnoe slovo! Otkuda mne znat’ ∅?» (Christie: 141).

it-seems me$_{DAT}$ you$_{NOM}$ **something$_{ACC}$** know$_{2SG}$ Rosa I$_{NOM}$ PARTICLE God$_{NOM}$ with you sir honest to goodness (*idiom*) how me$_{DAT}$ know **∅$_{ACC}$**
'"It seems to me that you know something, Rosa?" "Me? What are you talking about, sir. Honest to goodness! How should I know anything?"'

5.2. Gender agreement quandaries

Direct objects in Russian are often elided when there is a gender-related conflict associated with expressing them overtly. Such conflicts most often occur when the biological gender of a person does not correspond to the grammatical gender of the word used to refer to that person in the context. For example, in the first clause of (56), a biologically female girl is referred to by the grammatically masculine noun *ребёнок* 'child'; the next reference to the girl is as the direct object of *шлёпнула/šlëpnula* 'slapped':

(56) [Assume that the child is a girl]
Мать поймала **ребёнка** и шлёпнула **∅**.
Mat' pojmala **rebënka** i šlëpnula **∅**.
mother$_{NOM}$ caught **child$_{MASC.ACC}$** and slapped **∅$_{ACC}$**
'The mother caught the child and slapped her.'

Biological gender suggests that the object of *slapped* should be expressed using the feminine pronoun *её/eë* 'her'. However, this is grammatically impossible: rules of Russian grammar require that all pronouns agree in gender with their linguistic antecedents, if they have any. Therefore, if the object of *slapped* is to be expressed overtly, the masculine pronoun *его/ego* 'him' must be used. However, since it is strange to refer to a girl using a masculine pronoun, the elliptical variant of this sentence is highly preferred. Ellipsis is preferred when the clauses are uttered by different speakers as well, as in (57):

(57) [Assume that the child is a girl]
«Наконец мать догнала **ребёнка**». — «Но не шлёпнула **∅**, я надеюсь».
«Nakonec mat' dognala **rebënka**».—«No ne šlëpnula **∅**, ja nadejus'».
finally mother$_{NOM}$ caught-up-to **child$_{MASC.ACC}$** but NEG slapped **∅$_{ACC}$** I hope
'"Finally the mother caught up to the child." "But didn't slap her, I hope."'

A Russian joke shows the "convenient" use of ellipsis in contexts with gender-related complications. The English translation, of course, loses the humor.[9]

(58) «Чем отличается заяц от зайчихи?» — «Поднимаешь **∅** за уши и отпускаешь **∅**. Если побежал, то заяц. Если побежала, то зайчиха».
«Čem otličaetsja zajac ot zajčihi?» — «Podnimaeš' **∅** za uši i otpuskaeš' **∅**. Esli pobežal, to zajac. Esli pobežala, to zajčiha».
in-what differs hare$_{MASC}$ from hare$_{FEM}$ pick-up$_{2SG}$ **∅$_{ACC}$** by ears and let-go$_{2SG}$ **∅$_{ACC}$** if ran-away$_{MASC}$ then hare$_{MASC}$ if ran-away$_{FEM}$ then hare$_{FEM}$
'"How do you tell a male hare from a female hare?" "Pick it up by the ears and let it go. If it runs-away$_{MASC}$ it's a male hare. If it runs-away$_{FEM}$ it's a female hare."'

5.3. Repetition Structures for verbal emphasis

Repetition Structures are what I call contexts in which two consecutive conjuncts or clauses contain the same verb selecting the same argument(s).[10] In all Russian Repetition Structures—barring those instances in which overt repetition of arguments serves a stylistic purpose—the second object not only can but should be elided. The preference for object ellipsis derives from the function of Repetition Structures: to focus on the verb and defocus its subject or object(s). To create a Repetition Structure one can: repeat the verb alone (59); repeat it with a modifier (60); repeat it in a different tense or mood (61); repeat it with a change of subject, in which case either both subjects can be overt (62), the first can be elided (63), or the subject can be missing due to the imperative (64) or the indefinite personal construction (65); or repeat it in a clause that is joined to the preceding clause by various types of conjunctions or other binders (66)–(67):

(59) «Красное небо, уже начинает восходить луна, и я гнала **лошадь**, гнала ∅»
(Чехов 2: 396).
«Krasnoe nebo, uže načinaet vosxodit' luna, i ja gnala **lošad'**, gnala ∅» (Čexov 2: 396).
red sky already starts $\text{rise}_{\text{INFIN}}$ moon_{NOM} and I_{NOM} drove $\textbf{horse}_{\textbf{ACC}}$ drove $\varnothing_{\textbf{ACC}}$
'"The sky was red with the moon on the rise, and I drove that horse, drove it hard."'

(60) «Я понимаю, товарищи судьи, перед вами сложная задача: Деточкин нарушал **закон**, но нарушал ∅ из благородных намерений» (Брагинский и Рязанов: 182).
«Ja ponimaju, tovarišči sud'i, pered vami složnaja zadača: Detočkin narušal **zakon**, no narušal ∅ iz blagorodnyx namerenij» (Braginskij i Rjazanov: 182).
{ . . . } $\text{Detochkin}_{\text{NOM}}$ broke $\textbf{law}_{\textbf{ACC}}$ but broke $\varnothing_{\textbf{ACC}}$ from noble intentions
'"I understand, members of the jury, that you are faced with a difficult task: Detochkin broke the law, but he broke it with the best of intentions."'

(61) Он не любил **её** и никогда не смог бы ∅ полюбить (Уоллер: 20).
On ne ljubil **eë** i nikogda ne smog by ∅ poljubit' (Waller: 20).
he_{NOM} NEG loved $\textbf{her}_{\textbf{ACC}}$ and never NEG could CONDIT $\varnothing_{\textbf{ACC}}$ $\text{come-to-love}_{\text{INFIN}}$
'He didn't love her and never could come to love her.'

(62) «Кто впустил в мой дом **этого проходимца**?!» — «Я ∅ впустила»
(Вампилов 1: 35).
«Kto vpustil v moj dom **ètogo proxodimca**?!» — «Ja ∅ vpustila» (Vampilov 1: 35).
who_{NOM} let-in into my house $\textbf{this scoundrel}_{\textbf{ACC}}$ I_{NOM} $\varnothing_{\textbf{ACC}}$ let-in
'"Who let this scoundrel into my house?" "I did."'

(63) «**Баню** сегодня протопишь». — «Пусть инвалид ∅ протопит». —
«Я говорю, ты протопишь ∅. Понял?» (Вампилов 4: 329).
«**Banju** segodnja protopiš'». — «Pust' invalid ∅ protopit». — «Ja govorju, ty protopiš' ∅. Ponjal?» (Vampilov 4: 329).

bathhouse$_{\mathbf{ACC}}$ today will-heat$_{2SG}$ let invalid$_{NOM}$ **Ø**$_{\mathbf{ACC}}$ I$_{NOM}$ say you$_{NOM}$ will-heat$_{2SG}$ **Ø**$_{\mathbf{ACC}}$ understood$_{2SG}$
‘‘“Today you heat up the bathhouse.” “Let the invalid do it.” [lit.: let invalid Ø heat-up] “I said *you’ll* do it. Got it?”’

(64) «**Клапана** проверь, пожалуйста . . . » — «Я ж только что проверял **Ø**» (Токарева: 483).
«**Klapana** prover’, požalujsta . . . » — «Ja ž tol’ko čto proverjal **Ø**» (Tokareva: 483).
valve$_{\mathbf{ACC}}$ check$_{IMPER}$ please I$_{NOM}$ PARTICLE just checked **Ø**$_{\mathbf{ACC}}$
‘‘“Check the valve, please . . .” “I just checked it.”’

(65) «Мама, **Фирса** уже отправили в больницу. Яша отправил **Ø** утром» (Чехов 1: 602).
«Mama, **Firsa** uže otpravili v bol’nicu. Jaša otpravil **Ø** utrom» (Čexov 1: 602).
Mom **Firs**$_{\mathbf{ACC}}$ already sent$_{3PL}$ to hospital Yasha$_{NOM}$ sent **Ø**$_{\mathbf{ACC}}$ morning$_{INSTR}$
‘‘“Mom, Firs has already been taken [lit: they already took Firs] to the hospital. Yasha sent him this morning.”’

(66) Семейство Вежховицких я давно потеряла из виду, слышала только, что вроде бы старший Северин покинул **этот мир**. А раз покинул **Ø**, вряд ли сейчас околачивается у киоска с янтарём, значит, это Северин-младший . . . (Хмелевская 4: 24).
Semejstvo Vežxovickix ja davno poterjala iz vidu, slyšala tol’ko, čto vrode by staršij Severin pokinul **ètot mir**. A raz pokinul **Ø**, vrjad li sejčas okolačivaetsja u kioska s jantarëm, značit, èto Severin-mladšij . . . (Chmielewska 4: 24).
{ . . . } elder Severin left **this world**$_{\mathbf{ACC}}$ and since left$_{3SG}$ **Ø**$_{\mathbf{ACC}}$ unlikely now hangs-around$_{3SG}$ at kiosk with amber so this-is Severin Jr.
‘I had long since lost track of the Vezhxovickys, the only thing I had heard was that Severin Sr. had departed this world. And since he’d departed it, it’s hardly likely that he was, at this moment, hanging around an amber shop, so this must be Severin Jr. . . .’

(67) . . . [Зютек] жаловался на проблемы и неприятности, но жаловался как-то так, абстрактно, надо же хоть перед кем-то облегчить **душу**. Вот он и облегчал **Ø** перед Збиней, не требуя ни сочувствия, ни клятвы верности (Хмелевская 1: 245).
. . . [Zjutek] žalovalsja na problemy i neprijatnosti, no žalovalsja kak-to tak, abstraktno, nado že xot’ pered kem-to oblegčit’ **dušu**. Vot on i oblegčal **Ø** pered Zbinej, ne trebuja ni sočuvstvija, ni kljatvy vernosti (Chmielewska 1: 245).
{ . . . } necessary$_{IMPERS}$ PARTICLE at-least in-front-of someone lighten$_{INFIN}$ **soul**$_{\mathbf{ACC}}$ so he$_{NOM}$ PARTICLE lightened **Ø**$_{\mathbf{ACC}}$ in-front-of Zbinya NEG demanding neither sympathy nor vow secrecy$_{GEN}$
‘. . . Zyutek complained about his problems and troubles, but he complained kind of abstractly—he had to get it off his chest to someone, after all. And he did so to Zbinya, demanding neither sympathy nor a vow of secrecy.’

5.4. The verb has narrow selectional restrictions

Some verbs select a narrowly specified kind of object, which makes their object to some degree predictable; and the more predictable the object, the more readily it may be elided. Consider the following set of examples: all use a verb with narrow selectional restrictions and all permit DO ellipsis, even though the given object was not referred to overtly in the preceding context. The verb *пересдать/peresdat'* 'retake' (68) selects an object like *экзамен/èkzamen* 'exam', *контрольную/kontrol'nuju* 'quiz', and so on; the verb *назвать/nazvat'* 'call' (69) selects a person or thing that can be named; the verb *отслужить/otslužit'* 'serve out' (70) selects an object that refers to service, often military service; and the verb *подписать/podpisat'* 'sign' (71) selects an object that refers to a signable document:

(68) [In an exam-taking context. A student follows the proctor down the hall and asks about retaking the exam. The word *exam* is not used in the immediately preceding context.]
«Может, я пересдам ∅?» — идя за ней, нерешительно попросил Гошка (Войнович 1: 33).
«Možet, ja peresdam ∅?» — idja za nej, nerešitel'no poprosil Goška (Vojnovič 1: 33).
maybe I_{NOM} will-retake $∅_{ACC}$ walking behind her diffidently asked Goshka
'"Maybe I could retake it?" asked Goshka diffidently, walking behind her.'

(69) [About a dog named Tuzik]
«Что за имя — Тузик? Вялое какое-то. Надо было назвать ∅ покрепче» (Золотые: 33).
«Čto za imja — Tuzik? Vjaloe kakoe-to. Nado bylo nazvat' ∅ pokrepče» (Zolotye: 33).
what as name Tuzik wimpy sort-of $necessary_{IMPERS}$ was $name_{INFIN}$ $∅_{ACC}$ more-solidly
'"What kind of a name is that, Tuzik? It's kind of wimpy. You should have given him a better name."'

(70) «Я ж в армию иду. Отслужу ∅ — женюсь» (Токарева: 471).
«Ja ž v armiju idu. Otslužu ∅ — ženjus'» (Tokareva: 471).
I PARTICLE to army go_{1SG} $will\text{-}serve\text{-}out_{1SG}$ $∅_{ACC}$ $will\text{-}get\text{-}married_{1SG}$
'"I'm going into the army. Once I serve out my hitch I'll get married."'

(71) «За подписью надо ехать. Давай освобождай место, другие ждут». — «Слушай. Ну председатель потом приедет, подпишет ∅» (Войнович 1: 32).
«Za podpis'ju nado exat'. Davaj osvoboždaj mesto, drugie ždut». — «Slušaj. Nu predsedatel' potom priedet, podpišet ∅» (Vojnovič 1: 32).
for signature $necessary_{IMPERS}$ go_{INFIN} let's clear way $others_{NOM}$ are-waiting
$listen_{IMPER}$ PARTICLE $chairman_{NOM}$ then will-come will-sign $∅_{ACC}$
'"You have to go get this signed. Clear the way, others are waiting." "Listen. The chairman's coming later, he'll sign it then."'

The narrow selectional restrictions of these verbs make their objects largely predictable and therefore open to ellipsis despite the absence of an antecedent or audible/visible extralinguistic referent.

5.5. Another word in the context suggests the understood object

In some instances, although the verb itself does not have particularly narrow selectional restrictions, some word(s) in the preceding context suggest what object it must select. For example, (72) talks about shortening a last name. The discussion of last names in the preceding sentence makes it clear that *фамилия/familija* 'last name' is the understood object of *сократил/sokratil* 'shortened'.

(72) Зигфрид Фарнон. Странное имя для йоркширского сельского ветеринара. Наверное, немец—учился у нас в Англии и решил обосноваться здесь навсегда. И конечно, по-настоящему он не Фарнон, а, скажем, Фарренен, Сократил ∅ для удобства (Херриот: 13).
Zigfrid Farnon. Strannoe imja dlja jorkširskogo sel'skogo veterinara. Navernoe, nemec—učilsja u nas v Anglii i rešil obosnovat'sja zdes' navsegda. I konečno, po-nastojaščemu on ne Farnon, a, skažem, Farrenen. Sokratil ∅ dlja udobstva (Herriot: 13).
{ . . . } and of-course really he_{NOM} NEG Farnon but say Farrenen $shortened_{MASC.SG}$ $∅_{ACC}$ for convenience
'Siegfried Farnon. A strange name for a country vet in Yorkshire. He's probably a German who studied here in England and decided to put down roots. And, of course, he's not really Farnon but, say, Farrenen. He shortened it for the sake of convenience.'

Similarly, (73) talks about writing, so the object of the verb *прочесть* must be the material written: *то, что мы пишем<написали>/to, čto my pišem<napisali>* 'what we write <wrote>'.

(73) Даже если мы пишем, а не говорим, то собеседник всё равно подразумевается. Мы пишем так, чтобы другие могли прочесть ∅ (Фолсом: 131).
Daže esli my pišem, a ne govorim, to sobesednik vsë ravno podrazumevaetsja. My pišem tak, čtoby drugie mogli pročest' ∅ (Folsom: 131).
even if we write and NEG talk then listener nevertheless is-in-mind we write such that others can read $∅_{ACC}$
'Even if we're writing and not speaking, we still have a listener in mind. We write so that other people can read what we've written.'

Let us consider these examples in terms of Ontological Semantic text processing. In (71), backtracking the text to look for a semantically appropriate complement for *shorten* would yield an acceptable result, albeit not a perfect one. That is, if Farrenen was recognized as a last name and the system accepted a last name as a semantically appropriate object of *shorten*, then *shorten Farrenen* would be the

reconstruction of the ellipsis. This is not precisely correct, since the meaning is *shorten whatever last name he started out with, which was probably something like Farrenen*. However, reaching *that* analysis using the current or near-future state of the art will likely be unattainable. In (73), by contrast, there is no lexical item in the preceding context that is a candidate referent for the elided category. (By the way, this is not nonselection of the direct object, because the verb for *pročest'* 'read' is in the perfective aspect, which requires an object.) Therefore, different processing will be carried out. The verb *pročest'* will be linked to the ontological concept READ, which has as its default theme READING-MATERIAL (whose children are BOOK, NEWSPAPER, etc.). Since there is no contextual information to suggest any specific type of reading material, the generalized concept READING-MATERIAL will be used to fill out the theme slot in the semantic representation.

5.6. The structure is a set phrase

Russian has a number of set phrases in which the direct object is more commonly not expressed than expressed. This is not to say that the object cannot be overt—it can be, especially if having it overt improves the sentence rhythmically. However, such phrases are notable because ellipsis is the norm. They include: *извини*(*me*)*/izvini*(*te*) 'excuse me'; *прости*(*me*)*/prosti*(*te*) 'pardon me'; [of a door] *открой*(*me*)*/otkroj*(*te*) 'open up'; *пусти*(*me*)*/pusti*(*te*) 'let go (of me)'. There are also some clichés that are generally used without an object, like *прошу к столу/prošu k stolu* [lit.: ask_{1SG} to table] 'dinner (lunch, etc.) is served'. Set phrases and clichés that never permit insertion of an object are, of course, not elliptical. One such example is *hands off*—shown in (74):

(74) Илья сам стирал с него [рояля] пыль, а чтобы никто без толку не стучал по клавишам, положил на крышку табличку: «**Руками не трогать**!» (Войнович 1: 40).
Il'ja sam stiral s nego [rojalja] pyl', a čtoby nikto bez tolku ne stučal po klavišam, položil na kryšku tabličku: «**Rukami ne trogat'**!» (Vojnovič 1: 40).
$Ilya_{NOM}$ himself $dusted_{3SG.MASC}$ from it [piano] $dust_{ACC}$ and so-that nobody without reason NEG banged on keyboard $put_{3SG.MASC}$ on keyboard-cover sign **$hands_{INSTR}$ NEG $touch_{INFIN}$**
'Ilya dusted off the piano himself, and so that people wouldn't bang on the keyboard he put a sign on the keyboard cover: "Hands off!"'

A full list of set phrases like these should be added to the computational lexicon so that no syntactic ellipsis processing need be carried out. However, as in this last example, one still must make the inference that the object on which the sign lays is what must not be touched.

6. Using the parameters and values for cross-linguistic comparison

One way of evaluating the usefulness of a proposed inventory of parameters and values for describing a subclass of ellipsis is through cross-linguistic comparison. If com-

binations of parameter values yield different ellipsis judgments in different languages and if the inventory is sufficient to describe cross-linguistic differences, it is on the right track. This section shows the results of a first test of the inventory presented in this chapter, comparing the ellipsis potential of select structures in Russian, Polish, and Czech. All three are Slavic languages—Russian from the East Slavic group and Polish and Czech from the West Slavic group. As reflected by their group membership, Polish and Czech generally have more in common with each other than either has with Russian. In terms of DO ellipsis potential, however, Polish falls between Russian and Czech, patterning closer to Russian, in fact.[11]

In the following examples, when different languages have different DO-ellipsis potential, the direct object in the word-for-word translation is unmarked in order to avoid clutter, so each language example should be checked for DO status. In addition, if one language uses a category that the others do not (e.g., an overt subject or auxiliary) that category is in square brackets in the word-for-word translation.

6.1. Comparison #1

In syndetic VP-coordinate structures that contain an ACC R-expression antecedent,

- Russian and Polish speakers consistently permit optional DO ellipsis.
- Speakers of Czech vary in their judgments: some permit DO ellipsis quite liberally, while others severely restrict the employment of ellipsis. Here and subsequently, ≠ indicates conflicting speaker judgments.

(75) R: On vzjal **mjač** i brosil (**ego**).
P: Wziął **piłkę** i rzucił (**ją**).
Cz: Vzal **míč** a hodil (**ho**).
[he]$_{NOM}$ took **ball**$_{ACC}$ and threw (**it**)$_{ACC}$
'He took the ball and threw it.'

(76) R: Doktor Mortimer složil **gazetu** i položil (**eë**) v karman.
P: Doktor Mortimer złożył **gazetę** i schował (**ją**) do kieszeni (Conan Doyle: 22).[12]
Cz: ≠ Doktor Mortimer složil **noviny** a dal **je** do kapsy.
Doctor Mortimer$_{NOM}$ folded-up **newspaper**$_{ACC}$ and put **it**$_{ACC}$ in pocket
'Doctor Mortimer folded up the newspaper and put it in his pocket.'

6.2. Comparison #2

In multiclause syndetic VP-coordinate structures with an ACC R-expression antecedent,

- Russian speakers generally prefer that at least one direct object be elided; pattern (a) is preferred, (b) is generally acceptable, and (c) is often possible; (d) is considered excessively repetitive.

- Polish speakers differ in their judgments: while some prefer that at least one direct object be elided (with preferences as in Russian), others consider pattern (d) fully acceptable.
- Czech speakers differ radically in their judgments: while some permit ellipsis as liberally as in Russian or Polish, others consistently and emphatically reject patterns (b) and (c).

(77R) a. Mat' vykupala **rebënka**, ubajukala **ego** i položila ∅ spat'.
b. Mat' vykupala **rebënka**, ubajukala ∅ i položila ∅ spat'.
c. Mat' vykupala **rebënka**, ubajukala ∅ i položila **ego** spat'.
d. ⌋ Mat' vykupala **rebënka**, ubajukala **ego** i položila **ego** spat'.
Mother$_{NOM}$ bathed **child$_{ACC}$** lulled-to-sleep **it/∅$_{ACC}$** and put **it/∅$_{ACC}$** to-sleep
'Mother bathed the child, lulled it to sleep, and put it to bed.'

(77P) a. Matka wykąpała **dziecko**, ululała **je** do snu, i położyła ∅ spać.
b. Matka wykąpała **dziecko**, ululała ∅ do snu, i położyła ∅ spać.
c. Matka wykąpała **dziecko**, ululała ∅ do snu, i położyła **je** spać.
d. ≠ Matka wykąpała **dziecko**, ululała **je** do snu, i położyła **je** spać.
Mother$_{NOM}$ bathed **child$_{ACC}$** lulled-to-sleep **it/∅$_{ACC}$** and put **it/∅$_{ACC}$** to-sleep

(77Cz) a. Matka vykoupala **dítě**, uspala **je** a uložila ∅ do postýlky.
b. ≠ Matka vykoupala **dítě**, uspala ∅ a uložila ∅ do postýlky.
c. ≠ Matka vykoupala **dítě**, uspala ∅ a uložila **je** do postýlky.
d. ≠ Matka vykoupala **dítě**, uspala **je** a uložila **je** do postýlky.
Mother$_{NOM}$ bathed **child$_{ACC}$** lulled-to-sleep **it/∅$_{ACC}$** and put **it/∅$_{ACC}$** into bed

6.3. Comparison #3

In syndetic clausal coordinate structures that contain an ACC R-expression antecedent,

- Russian and Polish permit DO ellipsis only if the clauses are joined by a contrastive conjunction and have clearly contrastive semantics and intonation.
- Czech does not permit DO ellipsis.

(78) R: Ja snjala s nego **plašč** i Marija povesila **ego** na vešalku.[13]
P: Zdjęłam mu **płaszcz** i Maria powiesiła **go** na wieszaku.
[I]$_{NOM}$ took-off [from] him/him **raincoat$_{ACC}$** and$_{COORD}$ Mary$_{NOM}$ hung **it$_{ACC}$** on hanger
'I took his raincoat (off of him) and Mary hung it on a hanger.'

(79) R: Ja snjala s nego **plašč**, a Marija povesila (**ego**) na vešalku.
P: Zdjęłam mu **płaszcz**, a Maria powiesiła (**go**) na wieszaku.

Cz: Sundala jsem mu **plášť'** a Marie **ho** povsila na věšák.
I$_{NOM}$ took-off AUX [from] him **raincoat$_{ACC}$** and$_{CONTR}$ Mary$_{NOM}$ hung **it$_{ACC}$** on hanger
'I took his raincoat off of him and Mary hung it on a hanger.'

6.4. Comparison #4

In asyndetic coordinate structures whose conjuncts have coreferential subjects and whose antecedent is an ACC R-expression,

- Russian and Polish regularly permit DO ellipsis, but the elliptical variant may be stylistically marked.
- Czech does not permit DO ellipsis.

(80) R: Stražnik protjagivaet **meč**, vručaet **ego** Klavdiju.
P: Strażnik wyciąga **miecz**, podaje **go** Klaudiuszowi (Broszkiewicz: 283).[14]
guard$_{NOM}$ holds-out **sword**$_{ACC}$ hands **it**$_{ACC}$ Claudius$_{DAT}$
'The guard holds out the sword and hands it to Claudius.'

The elliptical variant of (80) is stylistically marked but can be made stylistically neutral in both languages by adding an overt coordinating conjunction.

6.5. Comparison #5

In the A and E strategy,

- Russian strongly promotes DO ellipsis.
- Polish and Czech do not permit DO ellipsis.

(81) R: Ja s"ela **tri pirožka**. Sama (**ix**) ispekla.
P: Zjadłam **trzy pierogi**. Sama **je** zrobiłam.
Cz: Snědla jsem **tři pirohy**. Upekla jsem **je** sama.
[I]$_{NOM}$ ate [AUX] **three$_{ACC}$ pierogis** self AUX **them$_{ACC}$** baked/made
'I ate three pierogis. I made them myself.'

6.6. Comparison #6

In sentences composed of a matrix clause and a gerund phrase, each of which contains an ACC direct object and the first of which realizes that direct object as an R-expression,

- Russian and Polish permit DO ellipsis only with element order [matrix + GP].
- Czech does not permit DO ellipsis.

(82) R: On otdal professoru **èsse**, ne proveriv (**ego**).
P: On oddał profesorowi **wypracowanie**, nie przeczytawszy (**je**).[15]

he$_{NOM}$ gave professor$_{DAT}$ **essay$_{ACC}$** NEG having-checked (**it**)$_{ACC}$
'He gave his professor the essay, not having checked it over.'

(83) R: Podnjav **okurok**, on brosil **ego** v urnu.
P: Podniósłszy **niedopałek**, wrzucił **go** do śmietnika.
Cz: Zvednuv **nedopalek**, hodil **ho** do popelnice.
having-picked-up **cigarette-butt$_{ACC}$** [he]$_{NOM}$ threw **it$_{ACC}$** into trash-can
'Having picked up the cigarette butt, he threw it into the trash can.'

6.7. Comparison #7

In Polish and Czech, as in Russian, pronominal antecedents have stronger ellipsis-promoting properties than R-expression antecedents, as evidenced by comparing (75) and (84). However, in Polish and Czech, unlike Russian, the strong preference for the elliptical variant does not consistently obtain, as shown by the optionality of the direct object in (85):

(84) [In reference to a ball]
R: On vzjal **ego** i brosil **Ø/ego.**
P: Wziął **ją** i rzucił **Ø/ją**.
Cz: Vzal **ho** a hodil **Ø/ho.**
[he]$_{NOM}$ took **it**$_{ACC}$ and threw **Ø/it**$_{ACC}$
'He took it and threw it.'

(85) R: Ona [. . .] celuet **menja** v lob i kladet **Ø/menja** k sebe na koleni (Tolstoj 2: 49).
P: Ona całuje **mnie** w czoło i sadza (**mnie**) na kolanie.
Cz: Líbá **mě** na čelo a bere (**mě**) na klín.
[she]$_{NOM}$ kisses **me**$_{ACC}$ on forehead and puts **me**$_{ACC}$ to self on knees
'She kisses me on the forehead and puts me on her lap.'

6.8. Comparison #8

Referent mismatches (generic specific or whole-part) produce the following effects:

- In Russian, the second direct object is unexpressed or, in the case of "some," represented using an appropriate referential expression.
- In Polish, in contexts that mean "one," *jeden* 'one' can be overt or elided. In contexts that mean "some," the category is unexpressed.
- In Czech, the direct object is always realized by an overt anaphor.

(86) R: Včera v universitete prodavali **komp'jutery** i moj brat **Ø/*odin** kupil.
P: Wczoraj na uniwersytecie sprzedawali **komputery** i mój brat (**jeden**) kupił.
Cz: Včera na univerzitě prodávali **počítače** a můj bratr **jeden** koupil.
yesterday at university were-selling$_{3PL}$ **computers$_{ACC}$** and my brother$_{NOM}$ **one$_{ACC}$** bought
'Yesterday they were selling computers at the university and my brother bought one.'

(87) R: Na uglu prodavali **apel'siny**, i ja kupila (**neskol'ko štuk**).
P: Na rogu sprzedawali **pomarańcze**, więc kupiłam (**parę/trochę**).[16]
Cz: Na rohu prodávali **pomeranče** a já jsem **nějaké** koupila.
on corner were-selling$_{3PL}$ **oranges$_{ACC}$** and/so [I]$_{NOM}$ AUX bought **a-few/some$_{ACC}$**
'They were selling oranges on the corner and/so I bought some.'

6.9. Comparison #9

Gender agreement quandaries in configurations that contain an ACC R-expression antecedent produce the following effects:

- In Russian and Polish, they render DO ellipsis virtually mandatory.
- In Czech, some speakers permit optional DO ellipsis while others do not.

(88) R: [Assume that the child is a girl]
Mat' pojmala **rebënka** i šlëpnula ∅/**lego**.
mother$_{NOM}$ caught **child$_{ACC.MASC}$** and slapped ∅/**him$_{ACC}$** [here: her]
'The mother caught the child and slapped her.'

(89) P: [Assume that the person in question is a man]
Policjant schwycił **jedną osobę** i uderzył ∅/**ją**.
policeman$_{NOM}$ grabbed **one$_{ACC.FEM}$ person$_{ACC.FEM}$** and hit ∅/**her$_{ACC}$** [here: him]
'The policeman grabbed one person and hit him.'

In Czech, some speakers permit DO ellipsis in contexts like these, while others do not. Those that do not must tolerate a lack of correspondence between grammatical gender and biological gender (however, in reality, such a gender mismatch would likely not arise because a problematic antecedent like the one in [90] would be avoided).

(90) Cz: [Assume that the person in question is a man]
Policista chytil **jednu osobu** a uhodil ≠**ji**.
policeman$_{NOM}$ grabbed **one$_{ACC.FEM}$ person$_{ACC.FEM}$** and struck ≠**her$_{FEM}$** [here: him]
'The policeman grabbed one person and struck him.'

7. Applying the description

A description of this grain size is clearly required if one wants to *generate* ellipsis with results similar to that of a native speaker. The benefits for *analysis* relate largely to the degree of confidence one can attribute to one or another analysis. As will be shown in chapter 7, unexpressed objects are not always due to syntactic ellipsis: objects can be unselected if they are lexically specified as optional, or they can be unexpressed due to modality, a generalized-human referent, or their belonging to

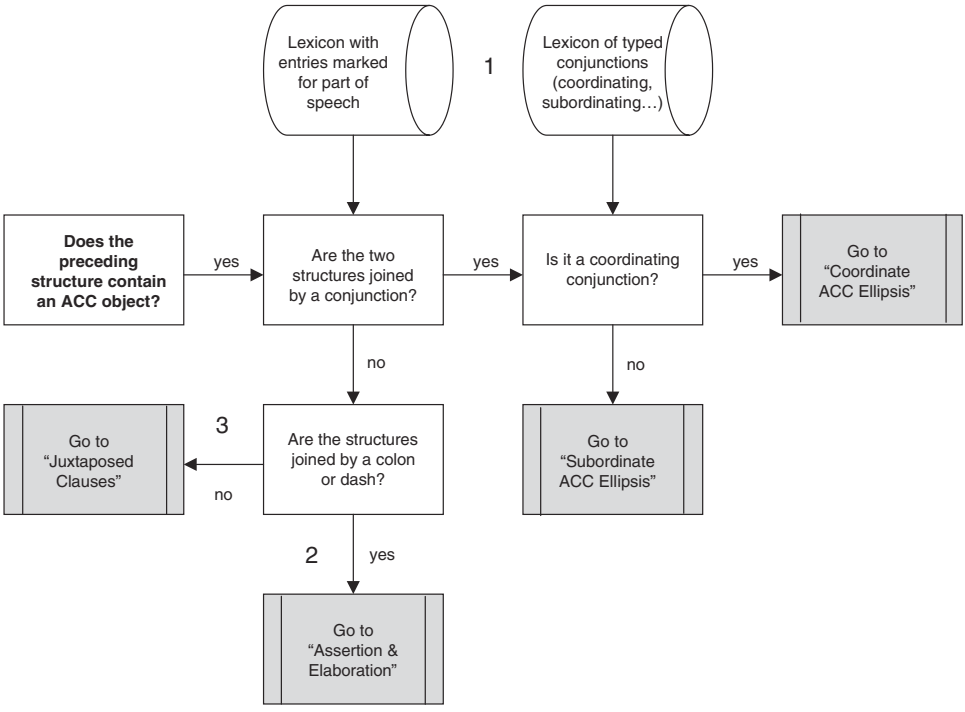
a series of actions. Therefore, every time an analysis system detects an unfilled object slot, all of these options are initially open, and programs must first determine if any can be excluded. Then the remaining valid analyses must be ranked using heuristics. It is descriptive details of the type shown earlier that permit a level of confidence to be applied to each analysis. For example, analyses of configurations that practically require syntactic ellipsis (e.g., the ellipsis of an ACC direct object with a semantically valid pronominal ACC antecedent in the preceding, coordinated verb phrase) carry an extremely high degree of confidence, so if they are detected they should practically exclude all other analyses. By contrast, elliptical configurations that are less frequent (e.g., the ellipsis of an ACC direct object with a semantically valid R-expression ACC antecedent in the preceding, coordinated clause) carry a lower degree of confidence, meaning that other potential sources of the missing object should be more rigorously explored. An implemented system would require a program to coordinate the search for and subsequent evaluation of all possible analyses of each unexpressed object.

Three sample algorithms that employ rules from various parts of this chapter follow. Although they are actually part of one large algorithm, they are split for reasons of space. In the algorithms, the traditional diamonds that indicate decisions are replaced by rectangles for two reasons. First, the diamonds provide too little space for text. Second, determining the answer to each question actually is a process, so indicating it by a rectangle—the traditional indicator of process—is appropriate. Another space-saving convention is use of the shorthand *semantically appropriate* when determining if the referent of the potential antecedent falls within the selectional restrictions of the verb whose object has been elided; obviously, if this is not the case there can be no coreference relationship between the potential antecedent and the elided category and that potential antecedent is excluded from consideration. I will abstract away from the preprocessing steps because the static resources and processing engines could take many forms.

Algorithm 3.1, "ACC Ellipsis with an ACC Antecedent," is called when a verb whose argument structure can or must include an ACC object is used in a text without that object overtly expressed *and* the preceding "structure" (nonparenthetical verb phrase or clause) contains an ACC noun phrase that could potentially be its antecedent. These conditions are detected by syntactic analysis, which includes determining phrase boundaries and seeking fillers for the lexically specified argument structure of each verb. This algorithm determines the relationship (e.g., coordination, subordination) between the two relevant structures and, in turn, calls the appropriate follow-up algorithm.

The three numbered points in the algorithm correspond to these three points of commentary:

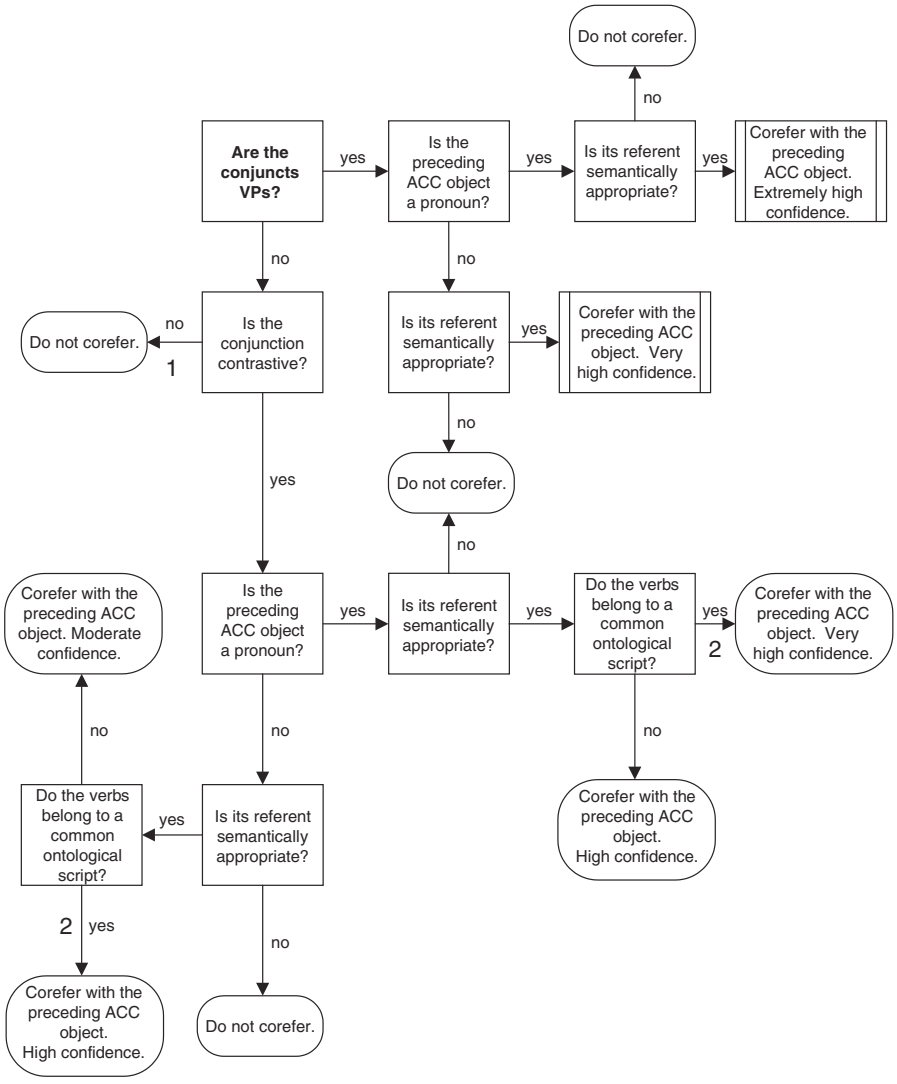
1. This algorithm, like those to follow, is not application specific and therefore leaves open the possibility for using various types of static resources. For example, the word list with parts of speech indicated, which is required for one decision, might be separate from the list of typed conjunctions, which is required for another decision.



ALGORITHM 3.1 ACC ellipsis with an ACC antecedent

2. The structures' being joined by a colon or dash is a strong indicator that the relationship between them is Assertion and Elaboration (an ellipsis-promoting structure for semantic and pragmatic reasons). It also excludes the possibility that the structure is asyndetic (i.e., conjunctionless) coordination.
3. In this eventuality, the structures can be joined by a comma (which can indicate coordination or Assertion and Elaboration) or by a semicolon or period (both of which can indicate many types of semantic and functional relationships).

Algorithm 3.2, "Coordinate Accusative Ellipsis," continues work on coordinate structures in the case when algorithm 3.1 has determined that the elliptical structure and the structure that contains the potential antecedent are in a syndetic coordinate relationship. A number of ellipsis-promoting and ellipsis-impeding factors are tested for here in order to assign a level of confidence to this analysis—that is, the analysis according to which this potential ACC antecedent is the actual antecedent. Factors that increase confidence that this is the correct analysis include the realization of the antecedent as a pronoun, use of both conjuncts' verbs in the same ontological script, and—if the subjects of the conjuncts are different—the conjunction being contrastive. In all cases, the meaning of the antecedent must be within the selectional



ALGORITHM 3.2 Coordinate Accusative Ellipsis, called when algorithm 3.1 has determined that the elliptical structure and the structure that contains the potential antecedent are in a syndetic coordinate relationship.

restrictions of the ellipsis-clause verb. If not, this algorithm returns "Do not corefer," which, depending on how traces of candidate analyses are stored, is equivalent to coreferring with zero confidence.

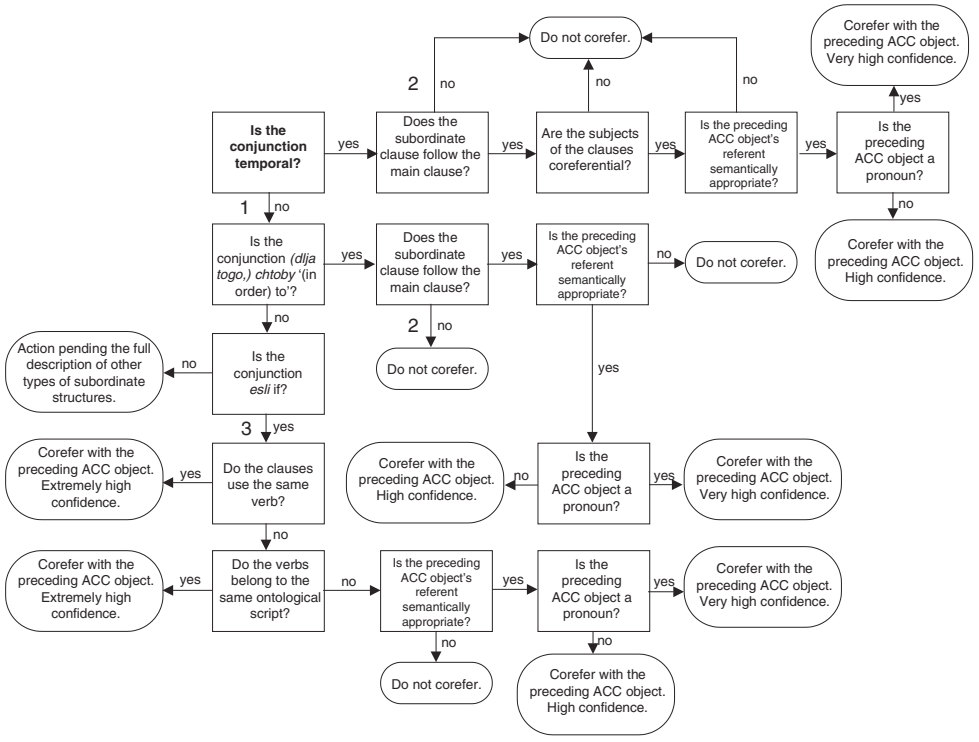
The numbers in algorithm 3.2 correspond to the following comments:

1. If the conjuncts are clauses and the coordinating conjunction is not contrastive, the ACC noun phrase in the first conjunct cannot

license ellipsis of a coreferential noun phrase in the second conjunct.

2. If ellipsis is to be possible in coordinate structures whose conjuncts are clauses (which, according to the classification of coordinate structures here, must have different subjects), the context must be common or typical. In Ontological Semantics, many such contexts will be recorded in ontological scripts. Therefore, tracing the verbs in the conjuncts back to a common ontological script increases the likelihood that this elliptical analysis is correct.

Algorithm 3.3, "Subordinate Accusative Ellipsis," continues work on subordinate structures in the case when algorithm 3.1 has determined that the elliptical structure and the structure that contains the potential antecedent are in a subordinate relationship. A number of ellipsis-promoting and ellipsis-impeding factors are tested for here in order to assign a level of confidence to this analysis—that is, the analysis according to which this potential ACC antecedent is the actual antecedent. Factors that increase confidence that this is the correct analysis include the realization of the antecedent as a pronoun and use of both conjuncts' verbs in the same ontological



ALGORITHM 3.3 Subordinate Accusative Ellipsis, called when algorithm 3.1 has determined that the elliptical structure and the structure that contains the potential antecedent are in a subordinate relationship.

script. Different classes of conjunctions require testing for different ellipsis-affecting factors. In all cases, the meaning of the antecedent must be within the selectional restrictions of the ellipsis-clause verb. If not, this algorithm returns "Do not corefer," which, depending on how traces of candidate analyses are stored, is equivalent to coreferring with zero confidence.

The numbers in algorithm 3.3 are associated with the following comments:

1. All of the temporal conjunctions studied in depth here had similar ellipsis-affecting properties. The nontemporal conjuctions split into several groups, based on their ellipsis-affecting properties.
2. In structures with temporal conjunctions or the conjunction (*dlja togo*,) *chtoby* '(in order) to' only the clause order [main + subordinate] supports ellipsis.
3. The conjunction *esli* 'if' has semantic and pragmatic ellipsis-promoting properties. Ellipsis is possible with either clause order and whether or not the clauses have the same subject.

This sampling of algorithms should suffice for illustration, so the algorithms for the other outcomes are not presented.

4

Direct Object Ellipsis with a Nominative Antecedent

As the last chapter showed, direct objects are readily elided when there is an accessible syntactic antecedent that is also a direct object. When the antecedent is syntactically accessible but not a direct object, DO-ellipsis potential is significantly restricted but not ruled out altogether. This chapter discusses what factors play a role in determining DO-ellipsis potential when the antecedent is overt in the immediately preceding context but is case-marked Nominative (NOM).

NOM noun phrases in Russian can, among other things, function as subjects and predicate nominals, as in (1), or as independent discourse topics, as in (2):

(1) Этот скрипач — гений.
Ètot skripač — genij.
that violinist$_{NOM}$ genius$_{NOM}$
'That violinist is a genius.'

(2) Цветы. Как я люблю цветы.
Cvety. Kak ja ljublju cvety.
flowers$_{NOM}$ how I$_{NOM}$ love flowers$_{ACC}$
'Flowers. How I love flowers.'

NOM subjects and independent topics have different potentials to support direct object ellipsis and will be considered separately.

1. The antecedent is a nominative subject

NOM subjects are not all created equal, nor do they function alike as ellipsis antecedents. For purposes of ellipsis, it seems sufficient to delineate two classes of NOM subjects:

1. subjects of lexical verbs, which are verbs that carry a full semantic load: e.g., *читать/čitat'* 'to read', *рисовать/risovat'* 'to draw'; and
2. subjects of existential and quasi-existential verbs. The main existential verb in Russian is *быть/byt'* 'to be'; another is *существовать/suščestvovat'* 'to exist.' Quasi-existential verbs have full lexical meanings but can be used in a semantically impoverished way as well. For example, *висеть/viset'* means "to hang," but in a sentence like *Картина висит на стене/Kartina visit na stene* 'The painting is [lit.: hangs] on the wall,' *висит/visit* 'hang' has basically existential function.

Subjects of lexical verbs cannot antecede DO ellipsis, no matter how logically obvious the object may appear:

(3) Занимают меня тоже **букашки и жучки**, я **их** сбираю, есть очень нарядные . . . (Достоевский 2: 55).
Zanimajut menja tože **bukaški i žučki**, ja **ix** sbiraju, est' očen' narjadnye . . . (Dostoevskij 2: 55).
interest$_{3PL}$ me$_{ACC}$ also **bugs$_{NOM}$ and beetles$_{NOM}$** I$_{NOM}$ **them$_{ACC}$** collect there-are very fancy$_{NOM}$
'Bugs and beetles also interest me, I collect them, there are some really fancy ones . . .'

(4) **Эти воспоминания** вставали сами, я редко вызывал **их** по своей воле (Достоевский 2: 54).
Èti vospominanija vstavali sami, ja redko vyzyval **ix** po svoej voli (Dostoevskij 2: 54).
these memories$_{NOM}$ arose selves$_{NOM}$ I$_{NOM}$ rarely summoned **them$_{ACC}$** by self's$_{DAT}$ will$_{DAT}$
'These memories would arise themselves, I rarely summoned them intentionally.'

By contrast, subjects of existential and quasi-existential verbs often *can* antecede DO ellipsis. For example, the NOM subject of *быть/byt'* 'to be' can antecede DO ellipsis in sentences that express existence, location, and possession, as shown by examples (5)–(7). Note that *быть/byt'* generally has null realization in the present tense, as in (5) and (6), but in certain types of existential contexts it is rendered by the fixed form *есть/est'*, as in (7):

(5) [After finding part of a valuable stamp collection]
«Это только половина успеха. И вряд ли мы добьёмся полного, ибо **остальное** рассеяно по свету. Собрать ∅ не удастся» (Хмелевская 1: 388).
«Èto tol'ko polovina uspeha. I vrjad li my dob'ëmsja polnogo, ibo **ostal'noe** rassejano po svetu. Sobrat' ∅ ne udastsja» (Chmielewska 1: 388).

this-is only half$_{NOM}$ success$_{GEN}$ and hardly PARTICLE we$_{NOM}$ will-achieve full$_{GEN}$ since **the-rest$_{NOM}$** scattered around earth collect$_{INFIN}$ **∅$_{ACC}$** NEG will-be-possible
‘‘This is only half a success. And it’s unlikely that we’ll ever achieve full success because the rest of the collection is scattered around the globe. We’ll never be able to gather it all together.’’

(6) «**Еда** на плите, разогреете ∅» (Вампилов 2: 120).
«**Eda** na plite, razogreete ∅» (Vampilov 2: 120).
food$_{NOM}$ on stove heat-up$_{IMPER}$ **∅$_{ACC}$**
‘‘The food’s on the stove, heat it up.’’

(7) «У меня есть **ливийские марки** . . . Задаром тебе отдам ∅» (Хмелевская 1: 228).[1]
«U menja est’ **livijskie marki** . . . Zadarom tebe otdam ∅» (Chmielewska 1: 228).
at me is **Libyan stamps$_{NOM}$** for-free you$_{DAT}$ will-give **∅$_{ACC}$**
‘‘I have some Libyan stamps . . . I’ll give them to you for free.’’

The NOM subjects of *остаться/ostat’sja* ‘stay’ and *лежать/ležat’*, among others, can antecede DO ellipsis when used quasi-existentially, as shown by (8)–(10).

(8) «Сейчас заскочим в бригаду. Там **подборщик** с осени остался, захватить ∅ надо. Мы бегом» (Войнович 1: 29).
«Sejčas zaskočim v brigadu. Tam **podborščik** s oseni ostalsja, zaxvatit’ ∅ nado. My begom» (Vojnovič 1: 29).
now will-run$_{1PL}$ to brigade-site there **harvester$_{NOM}$** since fall remained$_{3SG}$ fetch$_{INFIN}$ **∅$_{ACC}$** necessary$_{IMPERS}$ we$_{NOM}$ quickly
‘‘We’ll run right over to the brigade site. The harvester has been there since the fall, we need to get it. We’ll be quick about it.’’

(9) «А где щётка?» . . . — «Вот она». — «Ведь как удобно, когда **вещь** лежит на своём месте. Каждый подошёл, почистил сапоги, положил ∅ обратно. Другой подошёл, почистил, положил ∅ обратно. Никто времени не теряет» (Токарева: 505).
«A gde ščëtka?» . . . — «Vot ona». — «Ved’ kak udobno, kogda **veščˊ** ležit na svoëm meste. Každyj podošël, počistil sapogi, položil ∅ obratno. Drugoj podošël, počistil, položil ∅ obratno. Nikto vremeni ne terjaet» (Tokareva: 505).
so where(-is) brush$_{NOM}$ here-is it$_{NOM}$ after-all so convenient when **thing$_{NOM}$** lies in self’s place each(-person)$_{NOM}$ walked-up cleaned boots put **∅$_{ACC}$** back another(-person)$_{NOM}$ walked-up cleaned put **∅$_{ACC}$** back nobody$_{NOM}$ time$_{GEN}$ NEG wastes
‘‘Where’s the brush?” . . . “Here it is.” “You know, it’s so nice when a thing is where it’s supposed to be. Somebody comes, brushes off his boots, and puts the brush back. Then somebody else comes, brushes off his boots, and puts the brush back. And nobody wastes any time.’’

(10) «Открыли дверь, чтоб завтрак внести, а **он** лежит на полу в луже крови. Я велел не трогать ∅, пусть пока полежит» (Акунин: 191).

«Otkryli dver', čtob zavtrak vnesti, a **on** ležit na polu v luže krovi. Ja velel ne trogat ∅, pust' poka poležit» (Akunin: 191).
opened$_{3PL}$ door$_{ACC}$ in-order-to breakfast$_{ACC}$ bring-in and **he$_{NOM}$** lies on floor in pool blood$_{GEN}$ I$_{NOM}$ ordered NEG touch$_{INFIN}$ **∅$_{ACC}$** let-it-be for-now lies-for-a-while$_{3SG}$
'"They opened the door to bring in breakfast, and he was lying on the floor in a pool of blood. I ordered that he not be touched, that he be left lying there awhile."'

Quasi-minimal pair (11a-b) further illustrates how the nature of a NOM antecedent affects the ellipsis potential of a coreferential direct object. When the antecedent is the subject of the lexical verb *играть/igrat'* 'play', ellipsis is impossible, but when the antecedent is the subject of locative *быть/byt'* 'be', ellipsis is possible:

(11) a. **Мои внуки** играют в подвале. Приведи **их**, пожалуйста.
Moi vnuki igrajut v podvale. Privedi **ix**, požalujsta.
my grandchildren$_{NOM}$ play$_{3PL}$ in basement bring$_{IMPER}$ **them$_{ACC}$** please
'My grandchildren are playing in the basement. Bring them here, please.'
b. **Мои сапоги** в подвале. Принеси (**их**), пожалуйста.
Moi sapogi v podvale. Prinesi (**ix**), požalujsta.
my boots$_{NOM}$ in basement bring (**them**)$_{ACC}$ please
'My boots are in the basement. Bring them here, please.'

The obvious question is, why do the subjects of lexical and (quasi-)existential verbs behave so differently as ellipsis antecedents? Linguistic theory suggests an answer, which I will just briefly describe here.[2]

It has been argued that the subjects of existential verbs do not originate in subject position; instead, they originate in DO position. Accordingly, the English verb *to be* would be abstractly represented as follows:

[to be ___ ___]

where the slots represent the internal arguments of the verb. If we are constructing the sentence *Lou is brilliant*, *Lou* will go in the first slot and *brilliant* will go in the second one:

[to be Lou brilliant]

Then, according to rules of English grammar, *Lou* gets moved to sentence-initial position and *to be* gets changed to *is*, to agree with *Lou*. So, although *Lou* ends up as the subject, it does not begin in subject position; it begins in DO position—the perfect position for an antecedent for DO ellipsis in Russian.

Lexical verbs receive a different abstract representation: their subjects originate in subject position and remain there throughout sentence formation. The verb *to play*, for example, would be abstractly represented as follows (the parentheses indicate that it has an optional object):

[___ to play (__)]

To construct the sentence *The children are playing*, we insert *the children* into the slot to the left of *to play* and change *to play* to *are playing*, to agree with *the children*:

[the children to play (—)]

Since *the children* starts out in subject position and remains there, it has no objectlike properties and is not a good antecedent for DO ellipsis.

To summarize, subjects of (quasi-)existential verbs originate in DO position and become subjects only after abstract movement has occurred; such subjects retain some of their direct object–like properties and, therefore, can antecede DO ellipsis. Subjects of lexical verbs, by contrast, originate in subject position and remain there. Their lack of direct object–like properties makes them poor antecedents for DO ellipsis.

All of the elliptical examples shown earlier have one property in common: the antecedent clause and ellipsis clause are not in a coordinate relationship—that is, they are not joined by a coordinating conjunction like *and* or *or*. This is no accident. The single most defining property of coordinate structures is the parallel nature of their components. NOM subjects are not at all parallel, in function or case marking, to direct objects. Accordingly, if they are to be related as antecedent and elided category, it is better for them to be in a syntactic structure that does not tend toward parallelism.

Examples (12) and (13) show the difference in DO ellipsis potential between coordinate and noncoordinate structures that contain a NOM antecedent. (The pairs are not strictly minimal because different sentence structures demand different semantic relationships between clauses.[3]) The (a) variants are coordinate structures that block ellipsis; the (b) variants are A and E Strategies that permit ellipsis. The verbs *lie* and *stay* are used in these sentences in their quasi-existential function:

(12) a. Под дверью лежало **письмо**, и я **его** поднял.
Pod dver'ju ležalo **pis'mo**, i ja **ego** podnjal.
under door lay **letter**$_{\mathbf{NOM}}$ and I$_{\mathrm{NOM}}$ **it**$_{\mathbf{ACC}}$ picked-up
'There was a letter under the door and I picked it up.'

b. Почему **подушка** лежит на полу? Подними (**её**), пожалуйста.
Počemu **poduška** ležit na polu? Podnimi (**eë**), požalujsta.
why **pillow**$_{\mathbf{NOM}}$ lies on floor pick-up (**it**)$_{\mathbf{ACC}}$ please
'Why is the pillow on the floor? Pick it up, please.'

(13) a. У них остался **наш компьютер**, и мы забрали **его.**
U nix ostalsja **naš komp'juter**, i my zabrali **ego**.
at them remained **our computer**$_{\mathbf{NOM}}$ and we$_{\mathrm{NOM}}$ picked-up **it**$_{\mathbf{ACC}}$
'Our computer was still at their house and we picked it up.'

b. У них остался **наш комьютер**. Заберём (**его**) на следующей неделе.
U nix ostalsja **naš kom'juter**. Zaberëm (**ego**) na sledujuščej nedele.
at them remained **our computer**$_{\mathbf{NOM}}$ pick-up$_{1\mathrm{PL.FUT}}$ (**it**)$_{\mathbf{ACC}}$ on next week
'Our computer is still at their house. We'll pick it up next week.'

2. The antecedent is an independent nominative topic

Apart from functioning as subjects, NOM noun phrases in Russian also commonly function as (relatively) independent discourse topics upon which the following utterance comments. (These are not, however, grammaticalized discourse topics of the type found in Chinese.) Topiclike noun phrases have prominent status in the discourse and therefore can often support DO ellipsis. Some configurations in which (relatively) independent NOM topics can support DO ellipsis in Russian follow.

The NOM antecedent presents a discourse topic that is commented upon by the same speaker's following statement

(14) «**Вечный студент**! Уже два раза увольняли ∅ из университета» (Чехов 1: 587).
«**Večnyj student**! Uže dva raza uvol'njali ∅ iz universiteta» (Čexov 1: 587).
eternal student$_{NOM}$ already two times expelled$_{3PL}$ ∅$_{ACC}$ from university
'"An eternal student! They've already expelled him from the university twice."'

(15) «**Старые ненужные бумажонки**! Наверняка кто-то от немцев ∅ прятал, теперь это никому не нужно» (Хмелевская 5: 26).
«**Starye nenužnye bumažonki**! Navernjaka kto-to ot nemcev ∅ prjatal, teper' èto nikomu ne nužno» (Chmielewska 5: 26).
old useless papers$_{NOM}$ probably someone from Germans ∅$_{ACC}$ hid now this-is nobody$_{DAT}$ NEG necessary
'"Old useless papers! Probably someone hid them from the Germans and now they're of no use to anybody."'

The NOM antecedent presents a discourse topic that is commented upon by the listener's following statement

(16) «Какой **сильный мороз**!» — «Правда? А я совсем не чувствую ∅!».
«Kakoj **sil'nyj moroz**!» — «Pravda? A ja sovsem ne čuvstvuju ∅!».
what **hard frost$_{NOM}$** really but I$_{NOM}$ totally NEG feel ∅$_{ACC}$
'"What a hard frost!" "Really? I don't feel it at all!"'

The NOM antecedent is introduced by a word or phrase that points it out or draws attention to it: вот/vot *'here is',* вон/von *'there is',* смотри/smotri *'look', and so on.*

(17) Вот **бумажник** <Смотри, **бумажник**>. Должно быть, кто-то ∅ потерял.
Vot **bumažnik** <Smotri, **bumažnik**>. Dolžno byt', kto-to ∅ poterjal.
there-is **wallet$_{NOM}$** <look **wallet$_{NOM}$**> must be someone$_{NOM}$ ∅$_{ACC}$ lost
'There's a wallet <Look, a wallet>. Someone must have lost it.'

(18) Вон **наш дом**. Муж сам ∅ построил.
Von **naš dom**. Muž sam ∅ postroil.

there-is **our house$_{NOM}$** husband$_{NOM}$ self **Ø$_{ACC}$** built
'There's our house. My husband built it himself.'

The NOM antecedent is the response to a question like Что это?/Čto èto*?* (*word order free*) *'What's that?'*

(19) «А это что?» — «**Чайка**. Константин Гаврилыч убил **Ø**» (Чехов 2: 417).
«A èto čto?» — «**Čajka**. Konstantin Gavrilyč ubil **Ø**» (Čexov 2: 417).
and that what **seagull$_{NOM}$** Konstantin Gavrilych$_{NOM}$ killed **Ø$_{ACC}$**
'"And what's that?" "A seagull. Konstantin Gavrilych killed it."'

The NOM antecedent is in the first remark of a dialogue that begins with a question of the sort На что тебе X?/Na čto tebe X*? 'What do you want with X?'*

(20) «На что тебе **кассета**?» — Слушать **Ø** буду».
«Na čto tebe **kasseta**?» — «Slušat' **Ø** budu».
for what you$_{DAT}$ **tape$_{NOM}$** listen$_{INFIN}$ **Ø$_{ACC}$** will$_{1SG}$
'"What do you want with this tape?" "I'm going to listen to it."'

3. The inventory of parameters and values

The inventory of parameters and values suggested by the Russian data to cover DO ellipsis with a NOM antecedent is listed in table 4.1.

In addition, for a full description of this type of ellipsis, inventories of categories will be necessary, including the existential, quasi-existential, and lexical verbs in the language as well as words and phrases that can introduce an "independent topic" (e.g., *von* 'there is' in Russian). Also, in languages in which subjects can have case marking other than Nominative, the effects of case marking will need to be explored.

TABLE 4.1 Parameters and values for DO ellipsis with a Nominative NP antecedent

Parameters	*Values*	*Scope of applicability*
Function of the NP$_{NOM}$ antecedent	Subject Independent topic	All
Nature of the selecting verb	Lexical Existential Quasi-existential	The NP$_{NOM}$ is a subject
Syntactic relation between clauses	Coordination Assertion and Elaboration Dialogue Subordination Etc.	The NP$_{NOM}$ is a subject
How the topic is presented	Alone as a nominal sentence With an introductory phrase (like *There is . . .*)	The NP$_{NOM}$ is an independent topic

4. Using the parameters and values for cross-linguistic comparison

As in chapter 3, I will do a preliminary validation of the inventory of parameters and values, using Polish and Czech as test cases. Not unexpectedly, since Czech restricts DO ellipsis even with an "ideal" ACC antecedent, it completely blocks it with a NOM antecedent. (This suggests a variation on the theme of a grammatical role hierarchy, as posited by Keenan and Comrie [1977] and used, in identical or slightly amended form, by many others, e.g., Lappin and Leass 1994). The results from Polish are similarly expected: it blocks ellipsis in the same instances as Russian (when the verb selecting the NOM argument is semantically full) and permits it in only some of the contexts where Russian permits it. Specifically, Polish requires that the context contain ellipsis-promoting lexico-semantic or discourse factors that are not required in Russian. For example, Polish informants explain that ellipsis is possible in (21) because the context is so typical. In (22), ellipsis is possible because the direct object in question must be visible to both the speaker and the interlocutor—a discourse factor that strongly promotes DO ellipsis. Without this kind of lexico-semantic or discourse support, DO ellipsis with a NOM subject antecedent is blocked in Polish, as shown by (23)–(24).

(21) R: **Moj učebnik** v klasse; sejčas (**ego**) prinesu.
P: **Mój podręcznik** jest w klasie; zaraz (**go**) przyniosę.
Cz: **Moje učebnice** je ve třídě. Hned **ji** přinesu.
my textbook$_{\text{NOM}}$ [is] in classroom right-away **it**$_{\text{ACC}}$ will-bring$_{\text{1SG}}$
'My textbook is in the classroom; I'll bring it right away.'

(22) R: Počemu **poduška** ležit na polu? Podnimi (**ee**), požalujsta.
P: Dlaczego **poduszka** leży na podłodze? Podnieś (**ją**), proszę.
Cz: Proč **ten polštář** leží na podlaze? Zvedni **ho**, prosím
why **this pillow**$_{\textbf{NOM}}$ lies on floor? Pick-up **it**$_{\textbf{ACC}}$ please
'Why is this pillow (lying) on the floor? Pick it up, please.'

(23) R: **Moi sapogi** v podvale. Prinesi mne (**ix**), požalujsta.
P: **Moje kalosze** są w piwnicy. Przynieś mi **je** proszę.
Cz: **Moje boty** jsou ve sklepě. Přines mi **je**, prosím.
my boots$_{\textbf{NOM}}$ [are] in basement, bring$_{\text{IMPER}}$ me$_{\text{DAT}}$ (**them**)$_{\textbf{ACC}}$ please
'My boots are in the basement. Bring them to me, please.'

(24) R: U nix ostalsja **naš komp'juter**. Zaberem (**ego**) na sledujuščej nedele.
P: **Nasz komputer** został u nich. Przyniesiemy **go** w następnym tygodniu.
Cz: **Náš počítač** zůstal u nich. Přineseme **ho** příští týden.
our computer$_{\textbf{NOM}}$ remained at them will-bring$_{\text{1PL}}$ **it**$_{\textbf{ACC}}$ on next week
'Our computer remained at their house. We'll pick it up next week.'

When the antecedent is a (relatively) independent NOM discourse theme, Polish, like Russian, often permits DO ellipsis; again, Czech never does. Naturally, the inventory of "leading" words and phrases for discourse topics differs in Russian and Polish.

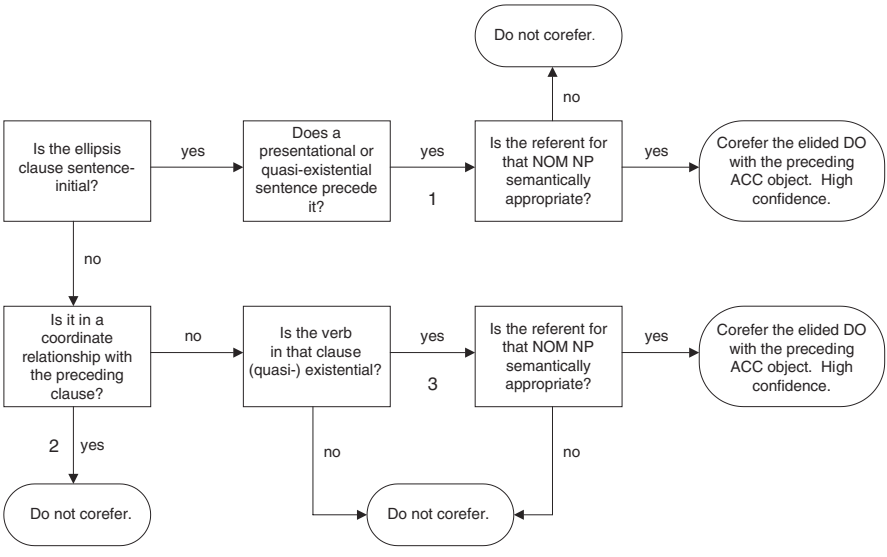
(25) R: «Kakoj **sil'nyj moroz**!» — «Pravda? A ja sovsem ne čuvstvuju (**ego**)»!
what **hard frost$_{NOM}$** really but I$_{NOM}$ totally NEG feel$_{1SG}$ (**it**)$_{ACC}$
P: „Ale **ostry mróz**." „Naprawdę? Wcale (**go**) nie poczułam."
what **hard frost$_{NOM}$** really totally (**go**)$_{GEN}$[4] NEG feel$_{1SG}$[5]
'"What a hard frost!" "Really? I don't feel it at all <I didn't notice it>."'

(26) R: Vot **bumažnik**. Dolžno byt', kto-to (**ego**) poterjal.
here **wallet$_{NOM}$** must be someone$_{NOM}$ (**it**)$_{ACC}$ lost
P: O! **Portfel.** Pewnie ktoś (**go**) zgubił.
oh **wallet$_{NOM}$** must-be someone$_{NOM}$ (**it**)$_{ACC}$ lost
Cz: Jé, **peněženka**! Někdo **ji** musel ztratit.
here's **wallet$_{NOM}$** someone$_{NOM}$ **it$_{ACC}$** must-have lost
'Here's/Oh, a wallet. Someone must have lost it.'

(27) R: «Čto èto?» — «**Zajac**. Moj brat (**ego**) pojmal».
P: „Co to?" „**Zając**. Mój brat (**go**) złapał."
Cz: „Co je to?" „**Zajíc**. Bratr **ho/*∅** chytil."
what [is] that **hare$_{NOM}$** [my] brother$_{NOM}$ **it$_{ACC}$** caught
'"What's that?" "A hare. My brother caught it."'

5. Applying the description

Algorithm 4.1 is a sample algorithm for an analysis system that exploits the description of Russian developed in this chapter.



ALGORITHM 4.1 ACC Ellipsis with NOM Antecedent, called in cases when it has been determined (e.g., by the failure to find a strong candidate antecedent with ACC case marking) that the NOM-case category in the preceding structure should be tested as a possible antecedent for the unexpressed ACC noun phrase.

The numbers in the algorithm correspond to the following comments:

1. A presentational sentence is composed of a bare NOM noun phrase or *vot/von/eto* 'here is, there is, this is' + NOM NP. A (quasi-)existential sentence contains an existential verb or a verb that can be used with diminished semantics (like *hang* in *The picture is hanging on the wall*,' which really means that the picture "is" on the wall).
2. The NOM subject of the first conjunct of a coordinate structure can never antecede ellipsis of the ACC direct object in the next conjunct. If ellipsis is to be possible, the clauses must be in a more "unbalanced" relationship, like an A and E Strategy.
3. Existential verbs and verbs used quasi-existentially are unaccusative, meaning that their subjects have some direct object–like properties. One such property is that they support DO ellipsis, whereas subjects of other verbs do not.

5

Direct Object Ellipsis with an Oblique Antecedent

In configurations in which the potentially elided direct object has a syntactically accessible antecedent with oblique case marking, ellipsis potential is subject to significant restrictions but is not ruled out, just as with the NOM antecedents discussed in chapter 4. Oblique antecedents for Russian include complements of verbs with GEN, DAT, or INSTR case marking and complements of prepositions with GEN, DAT, INSTR, PREP, or ACC case marking. Although it is terminologically inexact to consider ACC complements of prepositions oblique, I will do so in order to avoid the cumbersomeness of a more precise formulation, like "oblique objects and ACC complements of prepositions." The idea is that ACC complements of prepositions differ in structural position and behavior from ACC complements of verbs, putting them into the category that I am loosely calling oblique.

One might question whether it is justified to bunch together lexical obliques, configurational obliques, and objects of prepositions, since the literature on case in Russian reveals many differences between them. It *is* justified because many of the factors that affect DO ellipsis with an oblique antecedent apply uniformly to all kinds of oblique antecedents. Therefore, we will begin by considering them as a group, then proceed to differentiating properties.

All of the examples to be discussed here have the following features:

- the potentially elided category is a direct object with configurational ACC case marking;
- the antecedent is a syntactically accessible object oblique (including ACC assigned by a preposition) case marking;

- the antecedent is located in the clause that precedes the potentially elided category;
- the clauses/conjuncts have the same subject (I have been unable to find any examples where ellipsis was permitted despite a shift in subject, making this restriction practically supported, at least for Russian); and
- the clauses/conjuncts are either in a coordinate configuration, the A and E configuration, or a configuration that first presents a topic in relative isolation and then comments on it (subordinate configurations, gerund phrases, etc., are postponed since no valid elliptical examples of them have been found, either).

The real question for this subtype of ellipsis is, how does the case marking of NP antecedents and—in the case of noun phrases embedded in a prepositional phrase, their structural position—affect their ability to antecede DO ellipsis? The problem is finding enough examples of each case marking to support confident conclusions, especially for instances of lexical (quirky) case marking, since relatively few verbs impose each type; thus, it can be difficult to distinguish lexical effects from grammatical generalizations.

Because of the small number of examples found in texts (despite much searching!), many of the examples in this chapter were invented, then tested on half a dozen native speakers. One of the challenges encountered when inventing examples to test this class of ellipsis was that all elliptical utterances must be absolutely logical and the small number of verbs to choose from made inventing examples that were both logical and structurally appropriate quite difficult. That is, while it is possible to create illogical but grammatical nonelliptical utterances (à la Chomsky's "colorless green ideas sleep furiously"), the same cannot be said of elliptical utterances. This is because ellipsis represents a type of contract between the speaker and the listener: the speaker may elide only those categories that are readily recoverable. Ready recoverability cannot be achieved in the absence of semantic congruity. For example, a sentence like (1) is an invalid diagnostic for ellipsis potential with a PP_{PREP} antecedent because in our world boys generally do not eat the ponies they've just ridden:

(1) Мальчик катался на **пони** три часа, потом съел **его.**
Mal'čik katalsja na **poni** tri časa, potom s"el **ego.**
boy_{NOM} rode on $\textbf{pony}_{PREP}$ three hours then ate $\textbf{it}_{ACC}$
'The boy rode the pony for three hours, then ate it.'

Without ellipsis, the sentence is highly unusual in Russian and English, but it is nevertheless grammatical. With ellipsis, semantic oddity invalidates the structure.

For this class of ellipsis, informant surveys revealed a considerable lack of consensus, which is not surprising, since ellipsis potential is only in part determined by rules of grammar. Ellipsis judgments are further affected by style, register, intonation, the broader linguistic and extralinguistic context, the informant's personal language experience, his or her idiolect, and perhaps much more. Furthermore, the more infrequent the type of ellipsis, the more inconclusive judgments tend to be. For these reasons, I must make clear how I arrived at the judgments for each invented example

and what they are intended to represent. My informants were asked to evaluate examples of ellipsis on a four-point scale: perfect, somewhat less than perfect but acceptable, quite bad, impossible. Those examples that drew radically different responses were excluded. Those that received relatively consistent judgments were retained, but the rating was simplified to a two-point scale: the direct object may be elided (most speakers accept ellipsis, fully or with only slight degradation) or the direct object may not be elided (most speakers reject ellipsis or consider it highly degraded). The conclusions render tendencies in the employment of ellipsis as a step toward understanding the parameters and values that might affect ellipsis potential cross-linguistically.

The remainder of the chapter is organized as follows. Each of subsections 1–8 isolates some factor that promotes or impedes ellipsis in the configurations of interest. Minimal or quasi-minimal pairs are provided as evidence where possible. Section 9 shows that in most "real" (i.e., not invented) elliptical examples more than one ellipsis-promoting factor is at work. Section 10 discusses implications of these findings.

1. R-expression versus pronominal antecedent

In configurations with an oblique antecedent, perhaps the strongest ellipsis-promoting factor is for the antecedent to be a pronoun rather than an R-expression. To be expressed as such, it must either be the established topic of discourse or be readily recoverable from the extralinguistic context. In the following set of examples, which show antecedents with each type of case marking, the (a) variants have an R-expression antecedent and block ellipsis, whereas the (b) variants have a pronominal antecedent and permit ellipsis. In order to reduce the size of examples, the (a) and (b) variants are presented together, with a single gloss and translation that uses slash notation to show the alternative antecedents. The direct object is left unmarked for elidability in the word-for-word translation since its ellipsis potential is different for the (a) and (b) variants.[1] As a reminder, PP_{case} indicates a noun phrase with "case" case marking located within a prepositional phrase.

An aside: Recall that when the antecedent was itself an ACC direct object, its R-expression versus pronominal status determined whether ellipsis was optional or strongly preferred:

Он поднял **мяч** и бросил (**его**).	Он поднял **его** и бросил ∅/⊥⊥ **его**.
On podnjal **mjač** i brosil (**ego**).	On podnjal **ego** i brosil ∅/⊥⊥**ego**.
he_{NOM} picked-up $\textbf{ball}_{ACC}$ and threw $(\textbf{it})_{ACC}$	he_{NOM} picked-up $\textbf{it}_{ACC}$ and threw ∅/⊥⊥$\textbf{it}_{ACC}$
'He picked up the ball and threw it.'	'He picked it up and threw it.'

The antecedent is NP_{DAT}

(2) a. Он льстит **начальнику**, всячески старается ублажить **его.**
On l'stit **načal'niku**, vsjačeski staraetsja ublažit' **ego**.

b. On l'stit **emu**, vsjačeski staraetsja ublažit' (**ego**).
Он льстит **ему**, всячески старается ублажить (**его**).

he$_{NOM}$ flatters **boss/him$_{DAT}$** in-every-possible-way tries to-cater-to **him$_{ACC}$**
'He flatters his boss/him; he tries to cater to him in every possible way.'

The antecedent is NP$_{INSTR}$

(3) a. Он командует **женой и детьми**, держит **их** в кулаке.
On komanduet **ženoj i det'mi,** deržit **ix** v kulake.
b. Он командует **ими**, держит (**их**) в кулаке.
On komanduet **imi,** deržit (**ix**) v kulake.
he$_{NOM}$ rules-over **wife-and-kids/them$_{INSTR}$** keeps **them$_{ACC}$** in fist
'He rules over his wife and children/them; he keeps them on a tight rein.'

The antecedent is NP$_{GEN}$

(4) a. Он придерживается **своих прежних взглядов** и менять **их** не собирается.
On priderživaetsja **svoix prežnix vzgljadov** i menjat' **ix** ne sobiraetsja.
b. Он придерживается **их** и менять (**их**) не собирается.
On priderživaetsja **ix** i menjat' (**ix**) ne sobiraetsja
he$_{NOM}$ adheres-to **self's-former-views/them$_{GEN}$** and change **them$_{ACC}$** NEG plans-to
'He adheres to his former views/them and has no intention of changing them.'

The antecedent is PP$_{ACC}$

(5) a. Он согласился на **операцию**, но в последний момент отложил **её.**
On soglasilsja na **operaciju**, no v poslednij moment otložil **eë**.
b. Он согласился на **нее**, но в последний момент отложил **(её).**
On soglasilsja na **nee**, no v poslednij moment otložil (**eë**).
he$_{NOM}$ agreed to **operation/it$_{ACC}$** but at last moment postponed **it$_{ACC}$**
'He agreed to the operation/it but at the last moment postponed it.'

The antecedent is PP$_{DAT}$

(6) a. Он привык **к нашим методам** и стал **их** использовать.
On privyk k našim **metodam** i stal **ix** ispol'zovat'
b. Он привык **к ним** и стал (**их**) использовать.
On privyk k **nim** i stal (**ix**) ispol'zovat'.
he$_{NOM}$ became-accustomed to **our-methods/them$_{DAT}$** and began **them$_{ACC}$**
to-use$_{INFIN}$
'He became accustomed to our methods/them and began using them.'

The antecedent is PP$_{INSTR}$

(7) a. Девочка ходила по парку **с куклой** и не выпускала **её** из рук.
Devočka xodila po parku s **kukloj** i ne vypuskala **eë** iz ruk.
b. Девочка ходила по парку **с ней** и не выпускала (**её**) из рук.
Devočka xodila po parku s **nej** i ne vypuskala **(eë)** is ruk.
girl$_{NOM}$ walked around park with **doll/it$_{INSTR}$** and NEG let-go-of **it$_{ACC}$** from hands
'The girl walked around the park with her doll/it and wouldn't let go of it for anything.'

Undoubtedly, pronominal antecedents promote ellipsis on the level of discourse, but they might promote it on the level of form as well, since there is some degree of case syncretism among pronouns, as shown in table 5.1.

This table should be interpreted as follows, taking the first row as an example: Assume that the potentially elided direct object is a masculine singular ACC object: *ego/ego*. Four types of pronominal antecedents match it morpho-phonemically and therefore *look* and *sound* like ideal antecedents: an ACC or GEN verbal complement and an ACC or GEN prepositional complement (if it is the object of a preposition, an epenthetic *n* will be inserted, but the form is still basically the same). So, if there exists a parallelism constraint that correlates the morpho-phonemic form of coreferential elements and ellipsis, antecedents with the aforementioned types of case marking should have some ellipsis-promoting power.[2]

Although pronominal antecedents promote ellipsis, they do not guarantee that it will be licit. Two factors that block ellipsis even when the antecedent is pronominal are: (1) the subject of the second clause, albeit coreferential with that of the first clause, being overt; and (2) the ellipsis-clause verb having wide selectional restrictions. All of the following examples have pronominal antecedents and are preceded by an indication of what noun phrase is being referred to. The R-expression variant, naturally, resists ellipsis even more strongly than the pronominal variant.

Ellipsis is blocked by an overt subject

As mentioned earlier, I have constrained the scope of phenomena such that all of the clause complexes have coreferential subjects. In most of them, the second subject is unexpressed, which is the normal state of affairs for Russian, which is a *pro*-drop language.[3] If, however, for some reason the coreferential second subject is overt—as for emphasis or because of theme-rheme relations in the discourse—object ellipsis is blocked, as shown in (8):

(8) [About his wife and children]
  a. Он командует **ими**, держит (**их**) в кулаке.
     On komanduet **imi**, deržit (**ix**) v kulake.
     he_{NOM} rules-over $\mathbf{them}_{\mathbf{INSTR}}$ keeps $(\mathbf{them})_{\mathbf{ACC}}$ in fist
  b. Он командует **ими**, он держит **их** в кулаке.
     On komanduet **imi**, on deržit **ix** v kulake.
     he_{NOM} rules-over $\mathbf{them}_{\mathbf{INSTR}}$ he keeps $\mathbf{them}_{\mathbf{ACC}}$ in fist
     'He rules over them; he keeps them on a tight rein.'

TABLE 5.1 Case syncretism among Russian pronouns

Potentially elided category	*Features*	*Antecedents that match in form*
ero/ego	MASC.SG	NP_{ACC} NP_{GEN} PP_{ACC} PP_{GEN}
ero/ego	NEUT.SG	NP_{ACC} NP_{GEN} PP_{ACC} PP_{GEN}
eë/ee	FEM.SG	NP_{ACC} NP_{GEN} PP_{ACC} PP_{GEN}
ix/ix	PL	NP_{ACC} NP_{GEN} PP_{ACC} PP_{GEN} PP_{PREP}

Whereas in (8) the second subject is optionally overt, in (9) the second subject must be overt. As expected, DO ellipsis is blocked:

(9) [About an institute: институт/institut]
Им руководит профессор Петров, и он значительно расширил **его** за последние несколько лет.
Im rukovodit Professor Petrov, i on značitel'no rasširil **ego** za poslednie neskol'ko let.
it$_{\textbf{INSTR}}$ runs Professor Petrov and he$_{\text{NOM}}$ significantly expanded **it**$_{\textbf{ACC}}$ over last few years
'Professor Petrov runs it, and he has significantly expanded it over the past few years.'

The impossibility of ellipsis in (8b) and (9) derives from a combination of syntax and theme-rheme relations. In Russian coordinate structures, the subject of the second conjunct generally must be elided if coreferential with the subject of the first conjunct. Moreover, the subject, in a sense, creates the elliptical climate for the sentence. If something about the context causes coreferential *subjects* to be repeated, then that same something will cause coreferential objects to be repeated. This is an instance of dependencies of ellipsis, which is the topic of chapter 12.

Ellipsis is blocked by wide subcategorization and selectional restrictions

Some verbs, like English *cook*, have a single subcategorization pattern (subject, direct object) and narrow selectional restrictions (the agent must be a human and the object must be edible). Other verbs, like English *love*, permit several subcategorization frames and more types of complements:

- a noun phrase that indicates a person or animal: *I love my goldfish*;
- a noun phrase that indicates a concrete object: *I love my car*;
- a noun phrase that indicates an abstract object: *I love life*;
- an infinitival clause: *I love to sleep late on Sunday mornings*; and
- a gerund phrase: *I love sleeping late on Sunday mornings*.

When the ellipsis-clause verb in Russian has wide selectional restrictions, DO ellipsis is often blocked due to potential ambiguity. Although ambiguity is permissible in language in general, ambiguity-producing ellipsis tends to be avoided.[4] Ambiguity based on selectional restrictions is particularly common when the ellipsis-clause verbs can take either a nominal or a clausal complement, as in (10).[5]

(10) [About painting: živopis']
Он увлекается **ею**, понимает **её**.
On uvlekaetsja **eju**, ponimaet **eë**.
he$_{\text{NOM}}$ is-very-interested-in **it**$_{\textbf{INSTR}}$ understands **it**$_{\textbf{ACC}}$
'He's very interested in it; he understands it.'

Here the ellipsis-clause verb is *ponimaet* 'understands'. Whereas one can understand painting itself (a noun phrase that is coreferential with the antecedent-clause noun phrase), one can also understand how it is created (a clause), who the first-tier modern-day painters are (a clause), and numerous other things that can be expressed by noun phrases or clauses. Thus, if the object of *ponimaet* is not overtly expressed, the interlocutor is left in limbo, not sure whether ellipsis is intended or whether an overt complement will be supplied. A similar analysis applies to (11): one can value scholars themselves or one can value their accomplishments, what they do, how they think, and so on:

(11) [About some scholars: učenye]
Мы гордимся **ими**, высоко ценим **их**.
My gordimsja **imi** vysoko cenim **ix**.
we$_{NOM}$ are-proud-of **them$_{INSTR}$** highly value **them$_{ACC}$**
'We are proud of them; we highly value them.'

While the preceding examples met with speaker concensus, other examples of this type did not, suggesting two things: (1) speakers have considerably differing levels of tolerance for ambiguity in elliptical structures that contain verbs with wide selectional restrictions and (2) narrow versus wide selectional restrictions is not a binary choice but, rather, a continuum.

2. The nature of the clause complex

As was shown in chapter 4 with respect to NOM antecedents, when the antecedent and potentially elided direct object do not match in case marking, coordinate structures are resistant to ellipsis. When ellipsis is permitted, it tends to be in a configuration that highlights the given category on the level of discourse, like the A and E Strategy. In the following minimal pairs, all of which have referential-expression antecedents (which are less ellipsis promoting than pronouns), the coordinate (a) variants block ellipsis while the A and E (b) variants permit it:

(12) a. Он больше не катается на **велосипеде** и, наверно, скоро **его** продаст.
On bol'še ne kataetsja na **velosipede** i, naverno, skoro **ego** prodast.
he$_{NOM}$ anymore NEG rides on **bicycle$_{PREP}$** and probably soon **it$_{ACC}$** will-sell
'He doesn't ride his bike anymore and will probably sell it soon.'
b. Он больше не катается на **велосипеде**. Наверно, скоро (**его**) продаст.
On bol'še ne kataetsja na **velosipede** i, naverno, skoro (**ego**) prodast.
he$_{NOM}$ anymore NEG rides on **bicycle$_{PREP}$** and probably soon (**it**)$_{ACC}$ will-sell
'He doesn't ride his bike anymore. He'll probably sell it soon.'

(13) a. Папа дорожил **своими импортными подтяжками** и надевал **их** только по праздничным дням.
Papa dorožil **svoimi importnymi podtjažkami** i nadeval **ix** tol'ko po prazdničnym dnjam.

Dad$_{NOM}$ loved **self's imported suspenders$_{INSTR}$** and would-put-on **them$_{ACC}$** only on holiday days
'Dad loved his imported suspenders and wore them only on holidays.'

b. Папа дорожил **своими импортными подтяжками**. Надевал (**их**) только по праздничным дням.
Papa dorožil **svoimi importnymi podtjažkami**. Nadeval (**ix**) tol'ko po prazdničnym dnjam.
Dad$_{NOM}$ loved **self's imported suspenders$_{INSTR}$** would-put-on (**them**)$_{ACC}$ only on holiday days
'Dad loved his imported suspenders. He wore them only on holidays.'

Two factors appear to be jointly responsible for the differing ellipsis potential in these minimal pairs. First, the A and E strategy in the (b) variants gives the noun phrases in question more prominent thematic status than the coordinate structures in the (a) variants. Second, in the A and E strategy the structural (and case) mismatch between the antecedent and the potentially elided category is not as strongly felt as it is in a coordinate structure, since the single most defining property of coordinate structures is their tendency toward strict parallelism.[6]

Another way to make an antecedent highly thematic and therefore enhance its ability to support ellipsis is to present it in a question like "What about *x*?" (which is a theme with many variations). Since the response to such a question must refer to *x*, *x* can often be elided. (Cf. NOM examples in chapter 4, section 2):

(14) «Кстати, о **сигаретах**. Привезли ∅?» (Вампилов 1: 45).
«Kstati, o **sigaretax**. Privezli ∅?» (Vampilov 1: 45).
by-the-way about **cigarettes$_{PREP}$** delivered$_{3PL}$ ∅$_{ACC}$
'"By the way, about the cigarettes. Have they been delivered?"'

(15) «А что тебе сказали про **твой реферат**?» — «Ничего особенного, — сказал Тишка. — Сказали, что опубликуют ∅ в учёных записках» (Войнович 2: 116).
«A čto tebe skazali pro **tvoj referat**?» — «Ničego osobenno, — skazal Tiška. — Skazali, čto opublikujut ∅ v učenyx zapiskax» (Vojnovič 2: 116).
and what$_{ACC}$ you$_{DAT}$ said$_{3PL}$ about **your$_{ACC}$ paper$_{ACC}$** nothing special said Tiška said$_{3PL}$ that will-publish$_{3PL}$ ∅$_{ACC}$ in academic papers
'"And what did they tell you about your paper?" "Nothing special," said Tishka. "They said they'd publish it in the working papers."'

(16) «Что мне делать с **этими панировочными сухарями**?» — Положи ∅ на противень» (Фрэнсис: 92).
«Čto mne delat' s **ètimi panirovočnymi suxarjami**?» — «Položi ∅ na protiven'» (Francis: 92).
what$_{ACC}$ me$_{DAT}$ do$_{INFIN}$ with **these$_{INSTR}$ for-breading$_{INSTR}$ crumbs$_{INSTR}$** put$_{INFIN}$ ∅$_{ACC}$ on griddle
'"What should I do with these bread crumbs?" "Put them on the griddle."'

3. Case-marking effects

In Russian, the particular type of oblique case marking of the antecedent deserves investigation, since different types of case marking have different structural and semantic properties that could affect ellipsis (cf. chapter 2, section 2). Data suggest that three types of oblique antecedents have increased potential to support ellipsis: Genitive of negation (GEN-Neg), Partitive Genitive (GEN-Part), and lexical ACC (i.e., ACC within a prepositional phrase).

3.1 The antecedent is configurational genitive

Of all oblique antecedents, ones with configurational GEN case marking—which include Genitive of negation and Partitive Genitive—appear to support DO ellipsis the best, based on the number of collected examples and the consistency with which informants accept ellipsis in them. There is nothing surprising in this, considering that configurational GEN objects show both semantic and structural affinity to ACC direct objects (see chapter 2). However, the felicity of DO ellipsis in structures that contain a configurational GEN antecedent does not derive solely from case marking as such. Instead, contexts that contain GEN-Neg and GEN-Part antecedents tend to have other ellipsis-promoting properties as well. The relative effects of case marking and these other factors, however, cannot be completely teased apart.

All of the following examples contain R-expression antecedents since they are more resistant to ellipsis than pronominal antecedents and therefore more clearly isolate the ellipsis-promoting properties of configurational GEN case marking.

The antecedent is GEN-Neg

In some instances, like (17), the GEN-Neg case marking of the antecedent, with the associated ellipsis-promoting syntactic and semantic factors discussed in chapter 2, appears to be solely responsible for permitting ellipsis:

(17) «Не натягивайте **струны** и попробуйте перервать ∅ — очень трудо; но натяните ∅ до последней возможности и наляжьте тяжестью пальца на натянутую струну — она лопнет» (Толстой 1: 262).
«Ne natjagivajte **struny** i poprobujte perervat' ∅ — očen' trudno; no natjanite ∅ do poslednej vozmožnosti i naljaž'te tjažest'ju pal'ca na natjanutuju strunu — ona lopnet» (Tolstoj 1: 262).
NEG stretch$_{IMPER}$ **string$_{GEN}$** and try$_{IMPER}$ break$_{INFIN}$ **∅$_{ACC}$** very hard but stretch$_{IMPER}$ **∅$_{ACC}$** to last possibility and set$_{IMPER}$ weight$_{INSTR}$ finger$_{GEN}$ on stretched string it$_{NOM}$ will-break
'"Try to break a string without stretching it—it's very hard; but stretch it to the maximum and set a finger to it and the stretched string will break."'

However, in other instances, other factors concurrently promote ellipsis. In (18) and (19), the direct object has a nonspecific referent (*one*, *some*), and Russian has no

pronouns with the feature 'nondefinite' (cf. chapter 3, section 5.1). Eliding the pronoun is a convenient way to work around this lexical gap.

(18) [About dancing]
Серёжа предложил мне быть с ним vis-à-vis. «Хорошо, — сказал я, — хотя у меня нет **дамы,** я найду ∅» (Толстой 2: 78).
Serëža predložil mne byt' s nim vis-à-vis. «Xorošo, — skazal ja, — xotja u menja net **damy**, ja najdu ∅» (Tolstoj 2: 78).
Serezha offered me_{DAT} be_{INFIN} with him vis-à-vis OK said I although at me no **lady**$_{GEN}$ I will-find $\emptyset_{ACC}$
'Serezha said I could dance with him vis-à-vis. "OK," I said. "I don't have a partner yet, but I'll find one."'

(19) «Во всём доме не найдётся ни **одного шприца**?» — «Да, собирался купить ∅ и забыл» (Хмелевская 1: 347).
«Vo vsem dome ne najdëtsja ni **odnogo šprica**?» — «Da, sobiralsja kupit' ∅ i zabyl» (Chmielewska 1: 347).
in whole house NEG will-be-found not **single syringe**$_{GEN}$ yes $planned_{1SG.MASC}$ buy_{INFIN} $\emptyset_{ACC}$ and $forgot_{1SG.MASC}$
'"There's not **a single syringe** in the whole house?" "Yeah, I was planning to buy **some** and forgot."'

In (20), the direct object has a nonspecific referent of a different sort—it refers to some unspecified portion of some unspecified sum of money that the speaker doesn't have to begin with. The highly unspecific nature of the referent promotes ellipsis on semantic grounds:

(20) «Как же, разбежался, не столько у меня **денег**, чтобы выбрасывать ∅ на ненужную вещь» (Хмелевская 4: 165).
«Kak že, razbežalsja, ne stol'ko u menja **deneg,** čtoby vybrasyvat' ∅ na nenužnuju veščʹ» (Chmielewska 4: 165).
what-is-he-crazy (idiom) NEG so-much at me **money**$_{GEN}$ in-order-to throw-away $\emptyset_{ACC}$ at unnecessary thing
'"What is he, crazy? As if I had so much money that I could blow it on something frivilous."'

In (21), the direct object refers to *such things*, which is the positive correlate for *anything like that* in the first clause. While *such things* can, of course, be expressed by the noun phrase *takie veščʹi*, it is preferably elided:

(21) «У нас я **такого** не видел, а за границей встречал» (Хмелевская 1997: 111).
«U nas ja **takogo** ne videl, a za granicej ∅ vsrečal» (Chmielewska 1: 111).
at us I **such-a-thing**$_{GEN}$ NEG saw but beyond border $\emptyset_{ACC}$ encountered
'"I've never seen anything like that here, but I have encountered such things abroad."'

And in (22), both clauses employ the same verb, which is negated in the first and positive in the second. The repetition of the verb strongly suggests that the object will also be repeated, and this promotes (even practically necessitates, on stylistic grounds) object ellipsis (cf. chapter 3, section 5.3):

(22) В прошлый раз я не нашла **музея**, но в этот раз непременно найду ∅.
V prošlyj raz ja ne našla **muzeja**, no v ètot raz nepremenno najdu ∅.
at last time I_{NOM} NEG found **museum**$_{GEN}$ but at this time definitely will-find **it**$_{ACC}$
'Last time I didn't find the museum, but this time I definitely will (find it).'

We conclude that antecedents with GEN-Neg case marking in and of themselves have ellipsis-promoting power, but this tends to combine with other ellipsis-promoting properties in the types of contexts where one typically employs GEN-Neg.

3.2. The antecedent is GEN-part

Noun phrases with Partitive Genitive case marking also have certain direct object–like properties and, as such, support DO ellipsis quite freely.

(23) Она налила себе **молока** и выпила (**его**).
Ona nalila sebe **moloka** i vypila (**ego**).
she_{NOM} poured $self_{DAT}$ **milk**$_{GEN}$ and drank **(it)**$_{ACC}$
'She poured herself some milk and drank it.'

One factor that can enhance ellipsis potential with a GEN-Part antecedent is the direct object's referring to an unspecific quantity, *some*, for which Russian lacks a pronoun. Rather than specify some quantity using an R-expression, ellipsis tends to be employed, as in (23).

(24) Не забыть завезти ей **лампадного масла**, в Москве можно купить ∅ (Московские новости #24, 1999).[7]
Ne zabyt' zavezti ej **lampadnogo masla**, v Moskve možno kupit' ∅ (Moskovskie novosti, #24, 1999).
NEG $forget_{INFIN}$ $bring_{INFIN}$ her_{DAT} **icon-lamp oil**$_{GEN}$ in Moscow $possible_{IMPERS}$ buy_{INFIN} $∅_{ACC}$
'I must remember to bring her some icon-lamp oil. I can pick some up in Moscow.'

A second factor that promotes ellipsis in certain partitive contexts is the partitive object's being located in a container such that the container and the object are acted upon simultaneously in the ellipsis clause. Eliding the object frees the speaker from having to choose which of the two parts to refer to, which is especially difficult if the partitive substance and container do not match in grammatical gender.

Note that English *it*, which has no gender feature, can be used with generalizing effect, to refer to both the container and the substance contained therein:

(25) Максим налил в стакан **воды** и протянул ∅ немому (Брагинский и Рязанов: 130).
Maksim nalil v stakan **vody** i protjanul ∅ nemomu (Braginskij i Rjazanov: 130).
Maxim$_{NOM}$ poured in glass$_{ACC.MASC}$ **water$_{GEN\cdot FEM}$** and held-out **∅$_{ACC}$** mute$_{DAT}$
'Maxim poured some water into a glass and held it out to the man who was stupefied into silence.'

(26) Потом [Тихоновна] встала, вынула из печи закопчённый казанок, налила в тарелку **борща**, поставила ∅ перед изкой: «Ешь» (Войнович 1: 70–71).
Potom [Tixonovna] vstala, vynula iz peči zakopčennyj kazanok, nalila v tarelku **boršča**, postavila ∅ pered Lizkoj: «Eš'» (Vojnovič 1: 70–71).
then [Tixonovna]$_{NOM}$ stood-up took from oven$_{GEN}$ blackened pot$_{ACC}$ poured$_{FEM.SG}$ into bowl$_{ACC.FEM}$ **borscht$_{GEN.MASC}$** put$_{FEM.SG}$ **∅$_{ACC}$** in-front-of Lizka eat$_{IMPER}$
'Then Tikhonovna got up, took a blackened pot from the oven, ladled out a bowl of borscht, and set it in front of Lizka: "Eat."'

3.3. The antecedent is PP$_{ACC}$

ACC complements of prepositions have the ideal case marking to antecede DO ellipsis, but they are located in a nonideal structural position: embedded in a prepositional phrase. The question is, does their case marking make them better antecedents for DO ellipsis than other complements of prepositions? The answer appears to be yes, since significantly more examples of DO ellipsis with a PP$_{ACC}$ R-expression antecedent were found than examples with other case markings. There appear to be no other ellipsis-promoting factors at work to explain the possibility of ellipsis in (27)–(28):

(27) Войдя в швейцарскую, Алексей Александрович взглянул на **письма и бумаги, принесённые из министерства**, и велел внести ∅ за собой в кабинет (Толстой 1: 364).
Vojdja v švejcarskuju, Aleksej Aleksandrovič vzgljanul na **pis'ma i bumagi, prinesennye iz ministerstva**, i velel vnesti ∅ za soboj v kabinet (Tolstoj 1: 364).
walking into foyer Aleksej Aleksandrovič$_{NOM}$ glanced at **letters$_{ACC}$ and papers$_{ACC}$ brought from ministry** and ordered bring$_{INFIN}$ **∅$_{ACC}$** after self into study
'Entering the foyer, Aleksei Aleksandrovich glanced at the letters and papers sent from the ministry and had them brought into his study after him.'

(28) Кто-то налетел на **Машу** сзади и изо всей силы толкнул ∅, так что она чуть не упала.
Kto-to naletel na **Mašu** szadi i izo vsej sily tolknul **∅,** tak čto ona čut' ne upala.
somebody$_{NOM}$ plowed into **Masha$_{ACC}$** from-behind and with all strength pushed **∅$_{ACC}$** so that she almost PARTICLE fell
'Someone plowed into Masha from behind and pushed her so hard that she almost fell down.'

4. Narrow selectional restrictions of the ellipsis-clause verb

As shown in chapter 3, section 5.4, when the ellipsis-clause verb has narrow selectional restrictions, this can increase the predictability of the object and thereby facilitate its elidability. This is true regardless of the nature of the antecedent. Example (29) is a particularly relevant case in point:

(29) Папа с мамой разошлись во мнениях: мама дала **сыну** подзатыльник, а папа похвалил (**его**).
Papa s mamoj razošlis' vo mnenijax: mama dala **synu** podzatyl'nik a papa poxvalil (**ego**).
Dad with Mom differed in opionions Mom gave **son**$_{\textbf{DAT}}$ smack-on-the-back-of-the-head$_{\text{ACC}}$ whereas Dad praised (**him**)$_{\textbf{ACC}}$
'Dad and Mom had a difference of opinion: Mom gave their son a smack on the back of the head whereas Dad praised him.'

The notable aspect of this example is that the antecedent clause contains both a DAT object (*son*) and an ACC one (*smack on the back of the head*). Chapter 3 showed that the most common antecedent for an elided direct object is the direct object in the preceding chunk of text. Here, however, a check of selectional restrictions for 'praise' should exclude that analysis, leading to the search for a semantically appropriate antecedent.

5. Paired adverbs

Certain adverbs form natural semantic pairs, like *first . . . then . . .* and *before . . . now. . . .* Using such pairs of adverbs in a clause complex focuses attention on the contrasted categories (here: the verbs), consequently defocusing the noncontrasted categories (here: the direct objects) and making them more open to ellipsis. In (30), the paired adverbs appear to be the only factor responsible for permitting ellipsis despite the fact that the antecedent is an NP$_{\text{PREP}}$ R-expression:

(30) Я сперва внимательно читаю, что написано на **лекарстве**, а потом уже начинаю принимать (**его**).
Ja sperva vnimatel'no čitaju, čto napisano na **lekarstve**, a potom uže načinaju prinimat' (**ego**).
I first carefully read what is-written on **medication**$_{\textbf{PREP}}$ and then already start take$_{\text{INFIN}}$ (**it**)$_{\textbf{ACC}}$
'First I carefully read what is written on medication, and only then start taking it.'

(31) Сначала разберитесь в **своих ошибках**, а потом исправляйте (**их**).
Snačala razberites' v **svoix ošibkax**, a potom ispravljajte (**ix**).
first figure-out PREPOSITION **your mistakes**$_{\textbf{PREP}}$ and then correct (**them**)$_{\textbf{ACC}}$
'First figure out your mistakes, then correct them.'

6. The A and E Strategy is signaled by a lexical entity

In all of the examples of A and E configurations presented so far, the clauses were juxtaposed without any lexical entity to overtly mark their functional and semantic relationship. This is perfectly common in Russian, and such structures sometimes permit DO ellipsis with an oblique antecedent. However, ellipsis potential is increased when the second clause is introduced by a lexical entity that anticipates the upcoming elaboration. This matter of discourse accounts for the difference in ellipsis potential in minimal pair (32a)-(32b):

(32) a. Она боится **за мужа**; любит **его**.
Ona boitsja za **muža**; ljubit **ego**.
she$_{NOM}$ fears for **husband$_{ACC}$** loves **him$_{ACC}$**
'She fears for her husband; she loves him.'
b. Она боится **за мужа**—значит, любит (**его**).
Ona boitsja za **muža**—značit, ljubit (**ego**).
she$_{NOM}$ fears for **husband$_{ACC}$** that-means loves (**him**)$_{ACC}$
'She fears for her husband—that means she loves him.'

7. Common semantic or pragmatic context

Just as adverbs can form natural semantic pairs, so, too, can verbs. This is particularly true if verbs are considered in conjunction with a given complement: for example, *make tea ~ drink it*, *buy an ice-cream cone ~ eat it*. In fact, the second action in such contexts is largely predictable from the first based upon our real-world knowledge. Having a highly predictable combination of verbs increases the ellipsis potential of direct objects, permitting their ellipsis despite having an antecedent that is an oblique R-expression:

(33) Бей **по стеклу**, пока не разобъёшь (**его**)!
Bej po **steklu**, poka ne razob"eš' (**ego**)!
beat$_{IMPER}$ at **glass$_{DAT}$** until PARTICLE will-break (**it**)$_{ACC}$
'Beat at the glass until you break it!'

(34) Ефим, мысленно чертыхаясь, пошёл **к дверям**, открыл ∅ и отпрянул (Войнович 2: 103).
Efim, myslenno čertyxajas', pošel k **dverjam**, otkryl ∅ i otprjanul (Voinovič 2: 103).
Efim$_{NOM}$ mentally cursing walked to **door$_{DAT}$** opened ∅$_{ACC}$ and recoiled
'Efim, cursing to himself, walked up to the door, opened it and recoiled.'

8. Rhythm, prosody

Some factors that affect ellipsis are difficult to formalize. One such factor is the perceived rhythmic balance of the utterance. In the following minimal pairs the (b) variants differ from the (a) variants only insomuch as their second clause contains an

extra modifier. The mere addition of this modifier makes the ellipsis of the direct object more viable for many speakers, the contrast here being important rather than the absolute judgments. Native speakers say that the added modifier "softens," "smoothes out," or "balances" the elliptical variant. The added modifiers are presented in braces in the combined glosses and translations.

(35) a. Он отнёсся к **тому делу** легкомысленно и отложил **его**.
On otnessja k **ètomu delu** legkomyslenno i otložil **ego**.
b. Он отнёсся к **этому делу** легкомысленно и отложил (**его**) на завтра.
On otnessja k **ètomu delu** legkomyslenno i otložil (**ego**) na zavtra.
he$_{NOM}$ related to **this matter$_{DAT}$** frivolously and put-off **it$_{ACC}$** {until tomorrow}
'He didn't take this matter seriously and put it off {until tomorrow}.'

(36) [About a bad system of political economics: *nerazumnaja èkonomičeskaja politika*]
a. Правительство придерживается **её** и навязывает **её** своим союзникам.
Pravitel'stvo priderživaetsja **ee** i navjazyvaet **ee** svoim sojuznikam.
b. Правительство придерживается **её** и даже навязывает (**её**) своим союзникам.
Pravitel'stvo priderživaetsja **ee** i daže navjazyvaet (**ee**) svoim sojuznikam.
government$_{NOM}$ adheres-to **it$_{GEN}$** and {even} imposes **it$_{ACC}$** self's allies$_{DAT}$
'The government adheres to it and {even} imposes it on its allies.'

9. The interaction of factors in typical examples

In the previous sections, an attempt was made to tease apart factors that promote and detract from ellipsis by contrasting similar examples with and without a given property. However, in typical examples that show DO ellipsis with an oblique antecedent, more than one ellipsis-promoting factor is at work. A small sampling drawn from literature is presented here, with the relevant factors named.

The antecedent is pronominal and the ellipsis-clause verb has narrow selectional restrictions

(37) В заключение режиссёр сунул **ему** в руки длинную шпагу и погнал ∅ на сцену биться с первым попавшимся (Брагинский и Рязанов: 143).
V zaključenie režisser sunul **emu** v ruki dlinnuju špagu i pognal ∅ na scenu bit'sja s pervym popavšimsja (Braginskij i Rjazanov: 143).
in conclusion director$_{NOM}$ shoved **him$_{DAT}$** into hands long sword$_{ACC}$ and drove **∅$_{ACC}$** onto stage fight$_{INFIN}$ with first person-one-runs-into
'In the end, the director shoved a sword into his hands and shoved him onto the stage to fight with whoever he ran into.'

The antecedent is pronominal and has ACC case marking (*as the complement of a preposition*)

(38) «Да, это очень дурно», — сказала Анна и, взяв сына за плечо . . .
посмотрела на **него** и поцеловала ∅ (Толстой 1: 373).

«Da, èto očen' durno», — skazala Anna i, vzjav syna za plečo . . . posmotrela na **nego** i pocelovala ∅ (Tolstoj 1: 373).
yes this-is very bad said Anna_{NOM} and having-taken son_{ACC} by shoulder looked at $\textbf{him}_{\textbf{ACC}}$ and kissed $\varnothing_{\textbf{ACC}}$
'"Yes, this is really bad," said Anna and, taking her son by shoulder, looked at him and kissed him.'

The antecedent is pronominal and the configuration is A and E

(39) [In response to being asked if she had carried out a request]
«Нет. Сам видишь, у меня нет условий, все чего-то хотят от **меня**, отвлекают ∅ . . . » (Хмелевская 3: 142).
«Net. Sam vidiš', u menja net uslovij, vse čego-to xotjat ot **menja**, otvlekajut ∅ . . . » (Chmielewska 3: 142).
no self see_{2SG} at me no conditions $\text{everybody}_{\text{NOM}}$ $\text{something}_{\text{ACC}}$ want_{3PL} from $\textbf{me}_{\text{GEN}}$ $\text{distract}_{\text{3PL}}$ $\varnothing_{\textbf{ACC}}$
'"No. As you see, I'm not in a position to do that, everyone wants something from me, everyone keeps distracting me . . ."'

The antecedent is pronominal, it has ACC case marking (as the complement of a preposition), and the configuration is A and E

(40) [About a woman being harrassed by criminals]
«Похоже, на **неё** действительно напали, надо спасать ∅» (Хмелевская 1: 200).
«Pохože, na **nee** dejstvitel'no napali, nado spasat' ∅» (Chmielewska 1: 200).
it-looks-like PREPOSITION $\textbf{her}_{\textbf{ACC}}$ really attacked $\text{necessary}_{\text{IMPERS}}$ $\text{save}_{\text{INFIN}}$ $\varnothing_{\textbf{ACC}}$
'"It looks like they really have assaulted her, we need to save her."'

10. Discussion

Unlike DO ellipsis with an "ideal" DO antecedent, ellipsis with an oblique antecedent is rarely required to produce a stylistically normal utterance. Therefore, one could write (for a computer) or learn (as a non-native speaker) a *production* grammar of Russian that lacked these elliptical patterns. However, an *analysis* grammar would require knowledge of these elliptical patterns because they will be met with in practice.

The observations in this chapter were presented in an order that can roughly be described as more easily formalizable to more difficult to formalize. For example, whether an antecedent is a pronoun or an R-expression is a fact accessible to any person, theory, or computational application, whereas the rhythmic balance of an utterance is accessible only to people—more specifically, only to native speakers.

In this analysis, one factor that strongly affects ellipsis potential was used as a control mechanism for the testing of other factors: the pronominal versus R-expression nature of the antecedent. Since oblique pronominal antecedents support ellipsis quite

well, I sought out contexts where ellipsis was impossible despite a pronominal antecedent. And since oblique R-expression antecedents block ellipsis quite strongly, I sought out contexts where ellipsis was possible despite an R-expression antecedent. The delineation of control factors to test other factors is a crucial aspect of the methodology that underlies this theory.

The properties that affect ellipsis judgments and the supporting evidence for them are summarized in table 5.2.

The rating of ellipsis-affecting factors using terms like *moderate*, *strong*, and *virtually absolute* is too general to be converted directly into a predictive algorithm. However, creating algorithms for the interaction of factors will depend in large part upon the given application. On the one hand, an analysis system needs to know when ellipsis can occur, not necessarily the factors that make it sound more or less natural, whereas a generation system needs the latter knowledge. On the other hand, not every application will have access to all the types of information needed to exploit all of these generalizations.

TABLE 5.2 Properties, ellipsis effects, and evidence for them for DO ellipsis with an oblique antecedent. [– ellipsis] indicates a factor that impedes ellipsis; [+ ellipsis] indicates one that favors ellipsis

Property	*+/– Ellipsis*	*Evidence*
Overt (coreferential) subject in second clause	Virtually absolute [– ellipsis]	Ellipsis is impossible despite a pronominal antecedent
Wide selectional restrictions of potentially elided category's verb	Very strong [–ellipsis]	Ellipsis is impossible despite a pronominal antecedent
A and E configuration (as constrasted with coordination)	Moderate [+ ellipsis]	Ellipsis is possible despite a referential-expression antecedent
Antecedent has GEN-Neg, GEN-Part, PP_{ACC}, or $NP_{DAT\text{-}HUMAN}$ case marking	Strong [+ ellipsis]	Ellipsis is possible despite a referential-expression antecedent
The clauses contain paired adverbs	Moderate [+ ellipsis]	Ellipsis is possible despite a referential-expression antecedent
A and E is signaled by a lexical category	Moderate [+ ellipsis]	Ellipsis is possible despite a referential-expression antecedent
Common and predictable semantic relationship between the clause's verbs	Moderate [+ ellipsis]	Ellipsis is possible despite a referential-expression antecedent
Good rhythm and prosodic features	Moderate [+ ellipsis]	Minimal pairs in which some rhythmic/prosodic element makes ellipsis possible

6

Elided Lexically Case-Marked Objects

Lexical (quirky) case marking is semantically vacuous, oblique case marking assigned to an object by the category that selects it. Objects assigned lexical case marking have an important property in common with objects that bear configurational case marking: they express objects directly affected by the verbal action. The question is, do they have the same ellipsis potential as configurationally case-marked direct objects? This chapter explores the ellipsis potential of lexically case-marked objects of verbs, which can have GEN, INSTR, and DAT case marking.[1]

Studying lexically case-marked objects is a challenge because, in languages that use lexical case marking, there tend to be relatively few words that impose each lexical case marking. This can significantly limit resources from which to build a database of examples—and the fewer the examples, the less convincing the analysis (cf. chapter 5). In addition, as will be described in chapter 7, there are a number of nonelliptical sources of missing objects that must not be confused with ellipsis as such: the nonselection of optional objects, the nonexpression of generalized human objects, the nonexpression of objects under modality, and the nonexpression of objects in a series of actions. Collected examples show that many of the verbs that impose lexical case marking in Russian are optionally transitive or select a human object; moreover, virtually all of them permit object nonexpression under modality or in a series of actions. This makes it difficult to find clear-cut instances of elided lexically case-marked objects. While this state of affairs poses certain challenges for the study of ellipsis, it also reminds us of the nature of natural language: it is not carefully constructed to meet the needs of linguistic analysis; it develops idiosyncratically to meet the needs and reflect the whims of

human speakers. So, although in studying the ellipsis potential of lexically case-marked noun phrases one should seek the grain size of description used for configurationally case-marked direct objects, this might not be possible. However, *some* elliptical facts should be discernible nevertheless, even from a small inventory of examples, and it is this goal that I put forth in analyzing this subtype of ellipsis.

1. Ellipsis in verbal repetition structures

Repetition Structures (cf. chapter 3, section 5.3) strongly favor object ellipsis, regardless of the object's case marking. This preference for ellipsis derives from the function of Repetition Structures: to focus on the verb and defocus its arguments. In Repetition Structures, one can repeat the verb alone for emphasis (1), repeat it with a modifier (2), or repeat it in a different tense or mood (3). In addition, the clauses may have the same or different (4) subjects. The examples show the ellipsis of objects with all three types of lexical case marking for verbal complements in Russian: GEN, DAT, and INSTR:

(1) Скажи ребёнку, что он молодец! Он заслуживает **похвалы**, заслуживает ∅.
Skaži rebënku, čto on molodec! On zasluživaet **poxvaly**, zasluživaet ∅.
$\text{tell}_{\text{IMPER}}$ kid_{DAT} that he_{NOM} great he_{NOM} deserves $\textbf{praise}_{\textbf{GEN}}$ deserves $\textbf{∅}_{\textbf{GEN}}$
'Tell the kid he did a great job! He deserves a pat on the back, he really deserves it.'

(2) Было видно, что громкая музыка за стеной мешала **бабушке**, очень даже мешала ∅.
Bylo vidno, čto gromkaja muzyka za stenoj mešala **babuške**, očen' daže mešala ∅.
was obvious that loud $\text{music}_{\text{NOM}}$ behind wall bothered $\textbf{Grandma}_{\textbf{DAT}}$ very-much really bothered $\textbf{∅}_{\textbf{DAT}}$
'It was clear that the loud music on the other side of the wall was bothering Grandma, *really* bothering her.'

(3) «Я должен вам объяснить свои чувства, те, которые руководили **мной** и будут руководить ∅ чтобы вы не заблуждались относительно меня» (Толстой 1: 531).
«Ja dolžen vam ob''jasnit' svoi čuvstva, te, kotorye rukovodili **mnoj** i budut rukovodit' ∅, čtoby vy ne zabluždalis' otnositel'no menja» (Tolstoj 1: 531).
I_{NOM} must you_{DAT} explain self's $\text{feelings}_{\text{ACC}}$ $\text{those}_{\text{NOM}}$ that guided $\textbf{me}_{\textbf{INSTR}}$ and will guide $\textbf{∅}_{\textbf{INSTR}}$ so-that you_{NOM} $_{\text{NEG}}$ be-mistaken about me
'"I must explain my feelings to you, those that guided me and will continue to guide me, so that you don't harbor any misconceptions about me."'

(4) Не обращая внимания на баб, я попросила Марека поделиться **новостями**. Марек **поделился** ∅ (Хмелевская 5: 126).
Ne obraščaja vnimajija na bab, ja poprosila Mareka podelit'sja **novostjami**. Marek podelilsja ∅ (Chmielewska 5: 126).

NEG paying attention to women ja$_{NOM}$ asked Marek$_{ACC}$ share$_{INFIN}$ **news$_{INSTR}$** Marek$_{NOM}$ shared **∅$_{INSTR}$**
'Not paying attention to the women, I asked Marek to tell me his news. He did.'

The next two examples show variations on the repetition theme. In (5), the verbs have the same stem but different prefixes that convey semantic nuances as well as aspectual differences. In (6), if the second object were overt it would be the negative indefinite pronoun *ничего/ničego*, which has a surface form different from its positive antecedent, *чего/čego*. In both examples, however, the preference for object ellipsis derives from the verbal repetition:

(5) Художник размахивал руками и бубнил: «Нельзя, чтобы детская книжка вышла без картинок! . . . » Я сначала не размахивал **руками**, но потом рассердился и тоже замахал **∅**: «Я совсем не хочу, чтобы книжка осталась без картинок! Просто ребёнок сам должен нарисовать себе картинки!» (Шинов: 12).
Xudožnik razmaxival rukami i bubnil: «Nel'zja, čtoby detskaja knižka vyšla bez kartinok! . . . » Ja snačala ne razmaxival **rukami**, no potom rasserdilsja i tože zamaxal **∅**: «Ja sovsem ne xoču, čtoby knižka ostalas' bez kartinok! Prosto rebënok sam dolžen narisovat' sebe kartinki!» (Šinov: 12).
{ . . . } I$_{NOM}$ at-first NEG waved **arms$_{INSTR}$** but then got-mad and also started-waving **∅$_{INSTR}$** { . . . }
'The artist waved his arms and muttered: "You can't publish a children's book without pictures! . . ." At first I didn't wave my arms, but then I got mad and started waving them too: "I don't want my book to be without pictures! I just want each kid to draw them for himself!"'

(6) «Он так метался, сидя за столом, что я боялся, как бы [он] **чего** не разбил. Представьте, не разбил **∅**, даже удивительно» (Хмелевская 4: 56).
«On tak metalsja, sidja za stolom, čto ja bojalsja, kak by [on] **čego** ne razbil. Predstav'te, ne razbil **∅**, daže udivitel'no» (Chmielewska 4: 56).
he so-much$_{NOM}$ thrashed-around sitting at table that I$_{NOM}$ feared as if [he]$_{NOM}$ **something$_{GEN}$** NEG broke imagine$_{IMPER}$ NEG broke **∅$_{GEN}$** even amazing
'"He was thrashing around so much at the table that I was afraid he'd break something. But would you believe it, he didn't break anything, it's really quite amazing."'

2. Ellipsis in other structures

In the next set of examples that show the ellipsis of lexically case-marked objects, the antecedent is still syntactically accessible, but the clauses have different verbs. Here, as elsewhere, the available examples drive the investigation. They leave a striking gap in data, however: they show only elided Datives. I have not found a single example of this basic profile in which a missing lexical GEN or INSTR object unambiguously represents ellipsis. While this state of affairs might initially seem provocative, it more likely derives from the mundane considerations discussed earlier:

the small inventory of lexical case-assigning verbs and the ready availability of nonelliptical sources of missing objects.

The most typical antecedent for an elided DAT object appears to be an object with ACC case marking, the most common type of object in Russian:

(7) Странно, что это её откровенное рассуждение и обидело **его**, и чем-то понравилось ∅ (Д. Гранин).
Stranno, čto èto eë otkrovennoe rassuždenie i obidelo **ego**, i čem-to ponravilos' ∅ (D. Granin).
strange that this her forthright reasoning$_{NOM}$ both offended **him$_{ACC}$** and somehow pleased ∅$_{DAT}$
'It was strange how her forthright reasoning both offended and somehow pleased him.'

(8) Она никогда не корила **его**, не досаждала ∅ поучениями, всё понимала, всё принимала . . . (В. Васильев).
Ona nikogda ne korila **ego**, ne dosaždala ∅ poučenijami, vsë ponimala, vsë prinimala . . . (V. Vasil'ev).
she$_{NOM}$ never NEG upbraided **him$_{ACC}$** NEG vexed ∅$_{DAT}$ sermons$_{INSTR}$ everything$_{ACC}$ understood$_{3SG.FEM}$ everything$_{ACC}$ accepted$_{3SG.FEM}$
'She never upbraided him, never vexed him with sermons. She understood and accepted everything.'

(9) Теперь Вы понимаете, как радует **меня** наша переписка. Не даёт ∅ потерять квалификацию (Рачко: 30).
Teper' Vy ponimaete, kak raduet **menja** naša perepiska. Ne daët ∅ poterjat' kvalifikaciju (Račko: 30).
now you$_{NOM}$ understand how pleases **me$_{ACC}$** our correspondence$_{NOM}$ NEG allows ∅$_{DAT}$ lose$_{INFIN}$ skills$_{ACC}$
'Now you understand how happy our correspondence makes me. It keeps me from losing my skills.'

(10) [In reference to a horse being killed by a snapped electrical wire]
Потянули к ответу **паренька-конюха**, стали грозить ∅ судом (В. Тендряков).
Potjanuli k otvetu **paren'ka-konjuxa**, stali grozit' ∅ sudom (V. Tendrjakov).
brought$_{3PL}$ to answer **stable-boy$_{ACC}$** began$_{3PL}$ threaten ∅$_{DAT}$ lawsuit$_{INSTR}$
'They accused a stable boy and began to threaten him with a lawsuit.'

These examples show that case matching between antecedent and elided object need not obtain. The reason that ACC antecedents readily support ellipis might be their unmarked (default) DO status. However, this postulation is premature, because all of the preceding examples show other ellipsis-promoting factors that could contribute to these ellipsis results: in (7)–(9) the antecedent is a pronoun; in (7) the *u . . . u/ i . . . i* 'both . . . and' construction increases the predictability and "balance" factors that promote ellipsis; and in (9)–(10) the A and E Strategy is employed. Thus, the specific role played by the ACC case marking of the antecedent is impossible to determine.

In example (11), the antecedent, like the elided element, has lexical DAT case marking. Ellipsis may or may not be promoted by matching case marking, but it is certainly promoted by the antecedent being realized as a pronoun:

(11) Минф показывает **мне** папку с документами, но советует ∅ узнать подробности об организации у Тим Лим, консультантки на первом этаже (Богуславская: 148).
Minf pokazyvaet **mne** papku s dokumentami, no sovetuet ∅ uznat' podrobnosti ob organizacii u Tim Lim, konsul'tantki na pervom ètaže (Boguslavskaja: 148).
Minf_{NOM} shows $\textbf{me}_{\textbf{DAT}}$ $\text{folder}_{\text{ACC}}$ with documents but advises $\textbf{∅}_{\textbf{DAT}}$ $\text{learn}_{\text{INFIN}}$ details of organization from Tim Lim consultant on first floor
'Minf shows me a folder filled with documents but advises me to learn about the details of the organization from Tim Lim, a consultant on the first floor.'

Ideally, one would test each case marking of the antecedent with each case marking of the potentially elided object in an attempt to tease apart case-marking effects on ellipsis potential—a procedure that is, however, unrealistic for Russian due to the scarcity of data.

The ellipsis of lexically case-marked objects can also be licensed by an extralinguistic antecedent, as in (12), but I leave further discussion of this type of ellipsis until chapter 13.

(12) [The stepmother, after looking over the dresses Zoluška made, says:]
«У нас нет оснований отвергать твою работу. Помоги ∅ одеться» (Шварц 3: 541).
«U nas net osnovanij otvergat' tvoju rabotu. Pomogi ∅ odet'sja» (Švarc 3: 541).
$\text{at us no grounds reject}_{\text{INFIN}}$ $\text{your work}_{\text{ACC}}$ $\text{help}_{\text{IMPER}}$ $\textbf{∅}_{\textbf{DAT}}$ $\text{get-dressed}_{\text{INFIN}}$
'"We have no grounds to reject your work. Help us get dressed."'

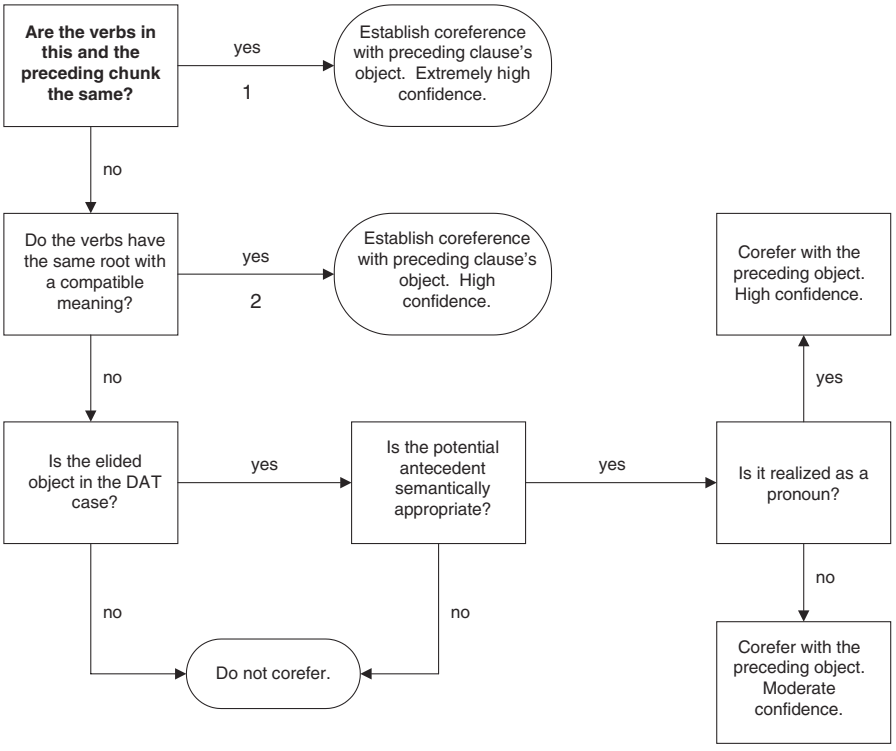
It would be difficult to delineate rules for generating the ellipsis of lexically case-marked objects in Russian, except in Repetition Structures, where they can predictably be elided. However, from the point of view of analysis, DAT objects certainly can be elided when they have an ACC, DAT, or extralinguistic antecedent, and there is insufficient evidence to exclude the possibility of GEN or INSTR ones being elided as well. Unexpressed lexically case-marked objects will further figure into the discussion of nonelliptical and hybrid sources of missing objects in chapter 7.

3. Applying the description

Algorithm 6.1, "Oblique Ellipsis," is a sample algorithm for an analysis system that exploits the description of Russian developed in this chapter.

The following numbered comments correspond to the numbers in the algorithm:

1. This is a Repetition Structure, which practically requires object ellipsis, regardless of the object's case marking.



ALGORITHM 6.1 Oblique Ellipsis. An unexpressed oblique object has been detected using syntactic analysis that relies on lexically specified argument structure. This algorithm determines if the object (with any case marking) in the preceding structure is a potential antecedent.

2. This is similar to a Repetition Structure, but the verbs have different prefixes, which convey different semantic nuances. Morphological or Ontological Semantic analysis is required to determine if verbs with the same root have similar meanings. Two common instances are: when affixation on one indicates a different aspect and when different affixes indicate different semantic nuances, which can be determined by checking if they both point to the same or related ontological concepts.

7

Unexpressed Objects That Do Not or May Not Represent Syntactic Ellipsis

Syntactic ellipsis—that is, the non-expression of a syntactically obligatory category with a contextually determined referent—is only one of several sources of unexpressed objects. Others include:

1. The Nonselection of Optional Objects. Some verbs, like *sing*, are optionally transitive, meaning that they may or may not select an object. Nonselection occurs in sentences like *The girl on the balcony is singing*.
2. Object Nonexpression Triggered by Modality. In some cases, verbs that would generally require an object can be used without one if the clause has certain types of modality. For example, *love* in English is obligatorily transitive, but one *can* say *I know that you know how to love*. It is the modality of *know how to* that shifts the focus to the verbal process and licenses the nonexpression of the object (cf. **You love/*You are loving/*You always love*).
3. The Nonexpression of Generalized-Human Referents. This type of object nonrealization is not found in English but is found in languages such as Russian and Polish. Unexpressed generalized-human objects in Russian can refer to all of humanity, some contextually implied subset of humanity (e.g., women, children), or some specific person viewed as a representative of all of humanity (similar to the generalized use of *you* in English).
4. The Nonexpression of Objects in Series. When numerous verbs are presented in an action-focused series, their objects can often be

unexpressed, even if they are lexically specified as being mandatory. That is, the syntactic/pragmatic nature of the utterance can override the subcategorization requirements of the verbs in question. For example: *People are really very much alike: they love, hate, marry, retire, have regrets and maybe some fun, then die.* Although the English verbs *love* and *hate* have obligatory direct objects (and *marry* may or may not), in a series like this the focus is on actions, not their arguments, so object nonexpression is permitted.

Depending on what one considers a full syntactic structure, these phenomena may or may not represent syntactic ellipsis. But for semantically rich NLP, structures like the preceding ones certainly represent semantic ellipsis, since some type of object is always implied. Prerequisites for analyzing and generating such elliptical structures include the following:

1. The computational lexicon must carefully and consistently specify which objects are optional.
2. The system must include lexico-syntactic rules to the effect that objects become optional if used with a modal, and the computational lexicon must note the relevant modals as well as any verbs for which such object nonexpression is semantically impossible: for example, *undergo*.
3. The computational lexicon or ontology must specify selectional restrictions for verbs, indicating which ones can or must take a human object, and the computational grammar for a language like Russian must contain a rule that permits nonspecification of human objects.
4. The system must include lexico-syntactic rules to the effect that objects become optional if verbs are used in series, with system trials used to determine the cutoff point for what represents a series (perhaps three actions, perhaps four . . .) and what other constraints might obtain.

For text analysis, each time a potential object gap is encountered in the syntax, several processes for resolving it must be launched. In the best case, all but one analysis will be excluded based on textual clues; however, more often multiple competing analyses will be valid, requiring some procedure for ranking outcomes. For text generation, a knowledge of the rules that govern the use of these types of missing objects and an understanding of their expressive power broadens the inventory of expressive means available to a language generator.

Missing-object phenomena can be described as five distinct linguistic processes: syntactic ellipsis plus the four strategies mentioned earlier. However, there is actually much gray area between them, meaning that the source of a missing object in a given context can defy clear analysis. For example, in a given context one might wonder: “Was the object selected, then elided, or not selected at all?” “Was it not selected, or was it selected but not expressed because it refers to generalized humans?”—and so on.

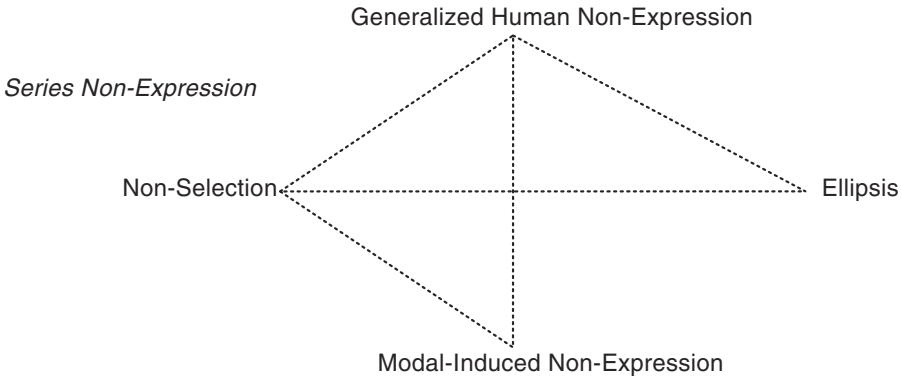


FIGURE 7.1 The space of missing-object phenomena

We might think of missing-object phenomena as a broad plane with five distinct nodes whose spheres of influence radiate out with unclear borders. Each node independently accounts for some missing objects, but there are expansive areas of overlap between them, as shown in figure 7.1. Each line represents hybrids—missing objects that cannot be unambiguously attributed to either of the sources at its end points. (Although I have found no examples that are ambiguous between modal nonexpression and ellipsis, I do not exclude the possibility that such examples exist. In addition, examples whose elliptical sources are three-ways ambiguous could exist as well).

Figure 7.1 is an approximation of the real interaction of phenomena, which would require a far more complex graphic. The reason that Series Nonexpression does not fit neatly into the schematic is because a series of verbs can contain verbs with different lexical properties: some may have optional objects and others obligatory objects; some select humans and some do not; and so on. The placement of this phenomenon in the upper left corner is arbitrary.

The following subsections present examples and discussion of nine of the ten missing-object phenomena—all clear-cut and hybrid phenomena except syntactic ellipsis, which was covered in chapters 3 through 6. Since the missing objects in the following examples do not or may not represent syntactic ellipsis, they are not indicated by ∅. Instead, the verbs whose objects are in question are in boldface for emphasis.

1. Clear-cut nonselection

Clear-cut nonselection of an object occurs when an optionally transitive verb is used without an object either to convey a process (1a) or to convey a general state of affairs (1b):

(1) a. Таня **читает** на крыльце.
Tanja **čitaet** na kryl'ce.
Tanja$_{\text{NOM}}$ **reads$_{\text{3SG.PRES}}$** on porch
'Tanya is reading on the porch.'

b. Таня любит **читать** на крыльце.
Tanja ljubit **čitat'** na kryl'ce.
Tanja$_{NOM}$ likes **read$_{INFIN}$** on porch
'Tanya likes reading on the porch.'

Even some verbs that are less purely process oriented appear to optionally select an object in both Russian and English—for example, *искать/iskat'* 'to seek' and *противоречить/protivorečit'* 'to object'.

(2) Я вообще считаю, что творческие люди — это люди, которые всегда **ищут**, находятся в движении, а значит, и переходят с одного места на другое (Домашний очаг, 1996).
Ja voobšče sčitaju, čto tvorčeskie ljudi — èto ljudi, kotorye vsegda **iščut**, naxodjatsja v dviženii, a značit, i perexodjat s odnogo mesta na drugoe (Domašnij očag, 1996).
I$_{NOM}$ basically think that creative people$_{NOM}$ that-is people$_{NOM}$ who always **seek$_{3PL}$** find-selves in motion and therefore also move from one place to another
'I basically think that creative people are people who are always seeking, who are always in motion, and therefore are always moving from one place to another.'

(3) [Они] примирились. Они не сопротивляются, не **противоречат**, не выдвигают никаких своих встречных предложений, они делают то, что им велят . . . (М. Кочнев).
[Oni] primirilis'. Oni ne soprotivljajutsja, ne **protivorečat**, ne vydvigajut nikakix svoix vstrečnyx predloženij, oni delajut to, čto im veljat . . . (M. Kočnev).
[they] succumbed$_{3PL}$ they NEG resist$_{3PL.PRES}$ NEG **object$_{3PL}$** NEG put-forward$_{3PL}$ none self's counterproposals they do that which them$_{DAT}$ order$_{3PL}$
'They have succumbed. They're not resisting or objecting or putting forward counterproposals. They're just doing what they're told . . .'

Naturally, there is a semantically oriented overlap between the verbs that optionally select objects in different languages, but this overlap is not complete. For example, the Russian verb *мешать/mešat'* 'bother', does not require an object in the meaning shown in (4), but the default English equivalent, 'bother', does. Since virtually any meaning can be expressed in any language, there *is* a way to retain the verbal focus in English and avoid drawing focus to an overtly expressed object—that is, by using the phrase *get in the way*:

(4) Наверное, кое-что из названного можно было бы решить на месте. Однако **мешают** стереотипы мышления, сказывается дефицит самостоятельности (Правда, 1988).
Navernoe, koe-čto iz nazvannogo možno bylo by rešit' na meste. Odnako **mešajut** stereotipy myšlenija, skazyvaetsja deficit samostojatel'nosti (Pravda, 1988).
probably something$_{ACC}$ of that-named possible$_{IMPERS}$ was CONDIT decide$_{INFIN}$ on spot but **get-in-the-way$_{3PL}$** stereotypes$_{NOM}$ thinking$_{GEN}$ is-evident deficit$_{NOM}$ independence$_{GEN}$
'Some of the abovementioned things could probably have been decided on the spot. But stereotyped thinking got in the way, a lack of independence was evident.'

This example has different implications for different applications. If one were building a knowledge-lean machine translation system, optional-object verbs in each language should be compared. If an object is optional for some source-language verb but not optional for its target-language equivalent, an objectless equivalent (like 'get in the way' for *mešat'*) should be added to the target-language lexicon, along with a rule that favors this translation in structures where the source has no overt object. If one were building a knowledge-rich, semantically oriented analysis system, the fact that optional-object verbs *do* imply some object (often of a given semantic class) must be accounted for in the semantic representation. That is, a person *eats* something edible, *sings* some song, *smokes* a cigarette, pipe, cigar, or joint, and *reads* some reading material. So, whether the lexicon, the ontology, or the surrounding context is used to restore the object's referent and whether that referent is very general or quite specific, it must somehow be represented.

2. Clear-cut modal-induced nonexpression

Clear-cut modal-induced nonexpression occurs when a verb that would normally require an object is used in a modal clause and the modality of the clause permits the object to go unexpressed. Semantically, this defocuses the object and draws focus to the modality and the action itself. Modality in Russian can be carried by a verb (5), an adverb (6), or a noun (7):

(5) Джинджи был здоровый и сильный, **умел хотеть** и точно знал, чего хочет (Токарева: 34).
Džindži byl zdorovyj i sil'nyj, **umel xotet'** i točno znal, čego xočet (Tokareva: 34).
Džindži was healthy and strong **knew-how$_{3SG}$ want$_{INFIN}$** and exactly knew$_{3SG}$ what wants$_{3SG}$
'Dzhindzhi was healthy and strong, he knew how to want and knew exactly what he wanted.'

(6) . . . Этому другу он, правда, отдал много, всё, что имел. Когда сам не имел ничего. **Просто отдавать**, когда есть, трудно — когда нет (Минчин: 135).
. . . Ètomu drugu on, pravda, otdal mnogo, vsë, čto imel. Kogda sam ne imel ničego. **Prosto otdavat'**, kogda est', trudno — kogda net (Minčin: 135).
this friend$_{DAT}$ he$_{NOM}$ it-is-true gave a-lot$_{ACC}$ everything$_{ACC}$ that had$_{3SG}$ when self$_{NOM}$ NEG had$_{3SG}$ nothing$_{ACC}$ **easy give-away$_{INFIN}$** when exists hard when not
'. . . As for this friend, he really gave him a lot—everything he had. Even when he himself didn't have anything. It's easy to give things away when you've got plenty, it's hard when you've got nothing.'

(7) «Он вам не нравится?» — «Нет, почему же? У него особый **талант очаровывать**» (Кристи: 309).
«On vam ne nravitsja?» — «Net, počemu že? U nego osobyj **talant očarovyvat'**» (Christie: 309).

he$_{NOM}$ you$_{ACC}$ NEG pleases no why PARTICLE at him$_{GEN}$ special **talent$_{NOM}$ charm$_{INFIN}$**
'"You don't like him?" "Why would you say that? He's got the knack for charming a girl."'

Missing objects can be attributed to modal-induced nonexpression when pairs like the following produce one ungrammatical and one grammatical entity: *He knows how to want* versus **He wants/is wanting*.

The fact that a language permits modal-induced nonexpression of objects does not mean that it can be applied to any verb: there are semantic restrictions. For example, highly object-oriented verbs like *prefer* and *undergo* resist such usage, as can be confirmed by collocating them with modals: **He knows how to undergo/prefer, *It's easy to undergo/prefer.* Moreover, the verbs that are excluded vary from language to language, as shown by the impossibility of English **It's easy to give away* in (6).

Marking all verbs for their ability to have object nonexpression under modality could serve text analysis by excluding for the relevant subset of verbs the modal-induced nonexpression analysis from competing analyses of object gaps. The necessity of recovering the implied objects for a full semantic representation holds here as elsewhere.

3. Clear-cut nonexpression of generalized-human objects

Language is anthropocentric, affording people a special status. In some languages, *generalized* people have yet another status that is reflected grammatically. In Russian, unspecified people as agents must be unexpressed (cf. chapter 13, section 2) and unspecified people as objects are often preferably unexpressed:

(8) «Вам понадобились великаны . . . Они только в сказках хороши, а так они **пугают**» (Чехов 1: 580).
«Vam ponadobilis' velikany . . . Oni tol'ko v skazkax xoroši, a tak oni **pugajut**» (Čexov 1: 580).
you$_{DAT}$ were-necessary giants$_{NOM}$ they$_{NOM}$ only in fairytales good but otherwise they$_{NOM}$ **frighten$_{3PL}$**
'"You needed giants . . . It's only in fairy tales that they're good; in real life they scare people <are scary>."'

(9) . . . Дома всегда было много спиртного. Стояло открытым и не запрещалось. А не запретное — не **манит** (Минчин: 29).
. . . Doma vsegda bylo mnogo spirtnogo. Stojalo otkrytym i ne zapreščalos'. A ne zapretnoe — ne **manit** (Minčin: 29).
at-home always was a-lot-of$_{NOM}$ liquor$_{GEN}$ stood open and NEG was-forbidden but NEG forbidden NEG **tempts**
'. . . There was always a lot of liquor at home. The bottles were open and no one said it was off-limits. And what's not forbidden doesn't tempt you.'

(10) «Я останусь здесь, я буду бороться, драться, кусаться, но унижать моё человеческое достоинство не **позволю**» (Войнович 2: 94).
«Ja ostanus' zdes', ja budu borot'sja, drat'sja, kusat'sja, no unižat' moë čelovečeskoe dostoinstvo ne **pozvolju**» (Vojnovič 2: 94).
I_{NOM} will-remain here I_{NOM} will fight claw bit but belittle$_{INFIN}$ my human dignity NEG **will-allow$_{1SG}$**
'"I'm staying here. I'll fight, claw, bite, but I won't let anyone strip me of my human dignity <won't let my human dignity be stripped>."'

(11) Она никогда не раздражалась . . . Она **раздражала** (Рачко: 44).
Ona nikogda ne razdražalas' . . . Ona **razdražala** (Račko: 44).
she$_{NOM}$ never NEG got-irritated she$_{NOM}$ **irritated**
'She never got irritated . . . She did the irritating.'

As shown by the English equivalents for these examples, English does not permit the nonexpression of unspecified people as objects, using various other strategies instead: using *people* (8), generalized *you* (9), or (in negated clauses) *anyone* (10) as the object; passivizing the clause (10); or using a context-specific equivalent (11).

The diagnostics for generalized-human nonexpression are: the given verb requires an object; that object can or must be human; there is no context-specific referent that strongly suggests an elliptical analysis (based on all relevant ellipsis-resolution algorithms); and the clause is not modal.

4. Nonselection or generalized-human nonexpression?

The question of nonselection or generalized-human nonexpression arises when the following four conditions hold: the verb optionally selects an object; that object can or must be human; there is no context-specific referent that strongly suggests an elliptical analysis; and the clause is not modal. See figure 7.2 for the graphical representation of this choice space.

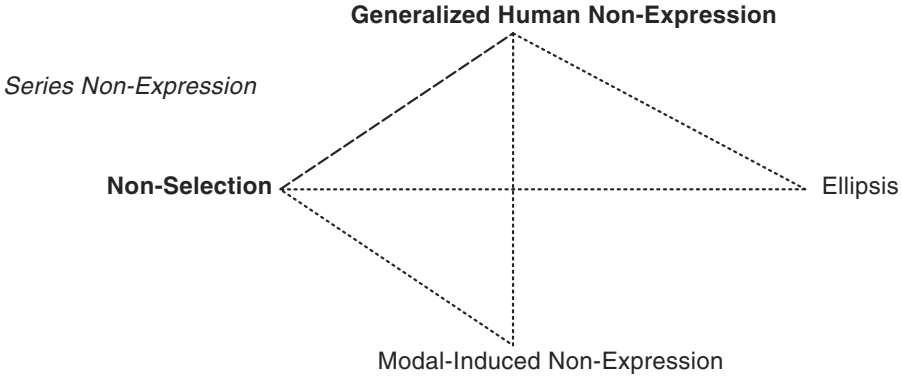


FIGURE 7.2 Nonselection or Generalized-Human Nonexpression?

In such contexts, either the object is not selected (implying stronger verbal focus) or it *is* selected but not expressed because it refers to generalized humans. In Russian contexts like (12)–(13), both analyses are available with only slightly different semantic focuses, which can be reflected by different English equivalents:

(12) «Что ж, деревня, говорят, **успокаивает**, он правильно сделал, что туда уехал» (Вампилов 4: 327).
«Čto ž, derevnja, govorjat, **uspokaivaet**, on pravil'no sdelal, čto tuda uexal» (Vampilov 4: 327).
well PARTICLE say_{3PL} $country_{NOM}$ **soothes** he_{NOM} correctly did that $there_{DIRECTIONAL}$ went
'"Well, they say the country has a soothing effect [or: soothes people]. He did right in going there."'

(13) Там наверху **командовал** Афоня Бронников (В. Распутин).
Tam naverxu **komandoval** Afonja Bronnikov (V. Rasputin).
there upstairs **commanded** $Afonja\ Bronnikov_{NOM}$
'There, upstairs, Afonja Bronnikov was giving orders [or: was ordering people around].'

Since one analysis cannot be selected over another, even by people, a practical approach for NLP would be to opt for the simpler one, generalized humans. Alternatively, one could create an internal representation that covers both options, selecting one over the other only if necessary—as in machine translation if the target language did not have a nonexpression strategy that matched the object-status ambiguity of the source-language original.

5. Modal-induced or generalized-human nonexpression?

The question of modal-induced or generalized-human nonexpression arises when the following four conditions hold: the verb requires an object; that object can or must be human; there is no context-specific referent that strongly suggests an elliptical analysis; and the clause is modal. Figure 7.3 shows the relevant decision space.

Since both modality and having a generalized-human referent permit the nonexpression of an otherwise obligatory object in Russian, it is impossible to select one or the other as the unambiguous licensor of the missing objects in examples like (15)–(17):

(15) Он знал за собой: умел **нравиться**. Но — людям определённого склада (Ю. Трифонов).
On znal za soboj: umel **nravit'sja**. No — ljudjam opredelënnogo sklada (Ju. Trifonov).
he_{NOM} knew about self knew-how $\mathbf{please_{INFIN}}$ but $people_{DAT}$ $certain_{GEN}$ $type_{GEN}$
'He was well aware that he knew how to make a good impression. But only on people of a certain type. [Or: He was well aware that he knew how to make people like him. But only people of a certain type.]'

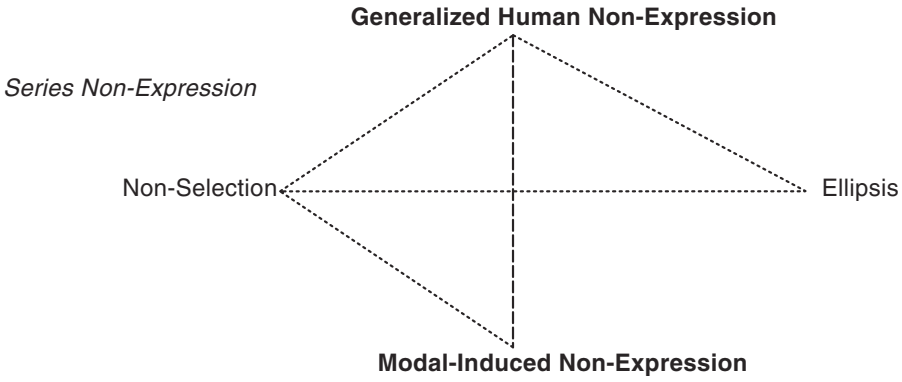


FIGURE 7.3 Modal-Induced or Generalized-Human Nonexpression?

(16) Рок-н-ролл может **освободить**, но он же может и **убить** (Огонёк, 1996).
Rok-n-roll možet **osvobodit'**, no on že možet i **ubit'** (Ogonëk 1996).
rock-n-roll$_{NOM}$ can **free** but it PARTICLE can also **kill**
'Rock'n'roll can free, but it can also kill [or: Rock'n'roll can free a person, but it can also kill him].'

(17) [About acting]
«Как трудно стало **смешить**! Это дело эмоции» (Огонёк 1996).
«Kak trudno stalo **smešit'**! Èto delo èmocii» (Ogonek 1996).
how difficult has-become$_{IMPERS}$ **make-laugh$_{INFIN}$** it-is matter$_{NOM}$ emotions$_{GEN}$
'"It's become so hard to make people laugh! It's all about emotions."'

Again, for processing purposes the generalized-human interpretation is more convenient if, of course, a hybrid semantic representation is infeasible.

6. Generalized-human nonexpression or ellipsis?

The area of overlap between generalized-human nonexpression and ellipsis (as shown in fig. 7.4) occurs when a specific human referent is available in the context (based on algorithms for resolving ellipsis, which permit both linguistic and extralinguistic referents), but the statement can be understood to apply to all of humanity as well.

For example, in (18) the first speaker asks how the listener feels when he reads about himself in newspapers. The objectless reply can be understood to convey both how he personally feels and how anyone else in that situation would feel. The generalizing effect is reinforced in this context by the second-person singular verb form in the last clause: *чувствуешь себя/čuvstvueš' sebja* 'you feel'.

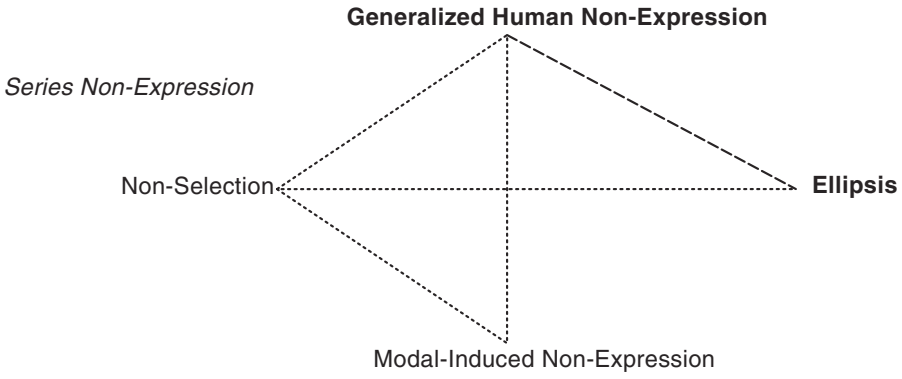


FIGURE 7.4 Generalized-Human Nonexpression or Ellipsis?

(18) «А если читаете про себя в газетах?» — «Когда **хвалят**, приятно, а когда **бранят**, то потом два дня чувствуешь себя не в духе» (Чехов 2: 413).
«A esli čitaete pro sebja v gazetax?» — «Kogda **xvaljat**, prijatno, a kogda **branjat**, to potom dva dnja čuvstvueš' sebja ne v duxe» (Čexov 2: 413).
and if read$_{2PL}$ about self in newspapers when **praise$_{3PL}$** nice$_{IMPERS}$ but when **criticize$_{3PL}$** then two days feel$_{2SG}$ self NEG in good-mood
'"And when you read about yourself in the newspaper?" "When they praise you, it's nice, but when they criticize you, it makes you out of sorts for a few days."'

(19) «Ну, а личная жизнь? **Даёт** ли истинную радость и удовлетворение власть над столькими людьми?» (Богуславская: 230).
«Nu, a ličnaja žizn'? **Daët** li istinnuju radost' i udovletvorenie vlast' nad stol'kimi ljud'mi?» (Boguslavskaja: 230).
well and private life$_{NOM}$ **give$_{3SG}$** PARTICLE real pleasure$_{ACC}$ and satisfaction$_{ACC}$ power$_{NOM}$ over so-many people
'"And what about your private life? Does power over so many people bring (you) true joy and fulfillment?"'

(20) «Было бы понятно, если бы вы вообще потеряли память, а такая странная избирательность **удивляет**» (Пелевин: 142).
«Bylo by ponjatno, esli by vy voobšče poterjali pamjat', a takaja strannaja izbiratel'nost' **udivljaet**» (Pelevin: 142).
was CONDITIONAL understood if CONDITIONAL you$_{NOM}$ in-general lost memory$_{ACC}$ but such strange selectiveness$_{NOM}$ **surprises$_{3SG}$**
'"It would be understandable if you totally lost your memory, but this strange selective memory loss is really shocking."'

(21) [Of people participating in amateur theater]
«Товарищи! — продолжал режиссёр. — Звание Народного театра ко многому **обязывает**» (Брагинский и Рязанов: 65).

«Tovarišči! — prodolžal režisser. — Zvanie Narodnogo teatra ko mnogomu **objazyvaet**» (Braginskij i Rjazanov: 65).
comrades$_{NOM}$ continued director$_{NOM}$ name$_{NOM}$ Community Theater$_{GEN}$ to a-lot **holds-responsible$_{3SG}$**
‘“Comrades!” continued the director. “The name Community Theater holds us [or: a person, one] to high standards.”’

Processing examples like these will be difficult, not only because there is ambiguity in the interpretation but also because locating the referent for an elliptical reading is difficult, since the referent is often recoverable only by extralinguistic means.

7. Nonselection or ellipsis?

Ambiguity between object nonselection and ellipsis (fig. 7.5) arises when a specific referent for the missing object is contextually available (based on ellipsis diagnostics), but the verb is not required to select an object.

In such contexts, if one were pressed to select a reading—as when translating into English—a more process-oriented reading (nonselection) or a more specific reading (ellipsis) could be selected, as shown by the variants in (22)–(23):

(22) Дядюшка моей жены сильно страдал печенью. Никакие медицинские средства уже не **помогали** (Вю Солоухин).
Djadjuška moej ženy sil’no stradal pečen’ju. Nikakie medicinskie sredstva uže ne **pomogali** (V. Solouxin).
uncle$_{NOM}$ my$_{GEN}$ wife$_{GEN}$ terribly suffered liver$_{INSTR}$ no$_{NOM}$ medicinal$_{NOM}$ means$_{NOM}$ already NEG **helped$_{3PL}$**
‘My wife’s uncle suffered terribly from a diseased liver. No medication helped [or: helped him] anymore.’

(23) «Пожалейте меня, хороший, добрый человек». — «Вы знаете, я **сочувствую** всей душой» (Чехов 1: 588).

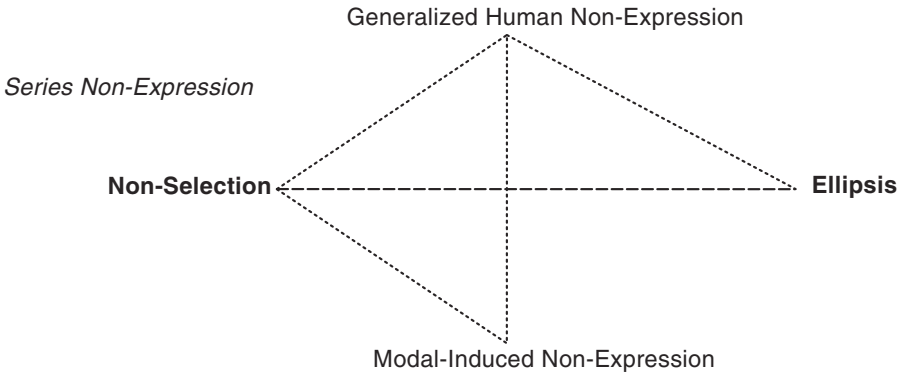


FIGURE 7.5 Nonselection or Ellipsis?

«Požalejte menja, xorošij, dobryj čelovek». — «Vy znaete, ja **sočuvstvuju** vsej dušoj» (Čexov 1: 588).
pity$_{IMPER}$ me$_{ACC}$ good kind person$_{NOM}$ you$_{NOM}$ know I$_{NOM}$ **sympathize** all$_{INSTR}$ soul$_{INSTR}$
'"Pity me, good, kind person." "You know, I wholeheartedly sympathize [or: sympathize with you]."'

The same text-processing issues described earlier apply here as well.

8. Nonselection or modal-induced nonexpression?

Object nonselection and modal-induced nonexpression apply to different classes of verbs. Nonselection applies to verbs that optionally select an object, whereas modal-induced nonexpression applies to verbs that would normally require an object but relax that necessity in modal clauses. Therefore, in theory, the question of nonselection versus modal-induced nonexpression should not arise and the boldface section of figure 7.6 should not exist: the lexically stipulated status of the object should make things clear.

The problem, however, is that for some verbs—like *ask* and *demand*—it is difficult to determine the base status of an object:

(24) Он не умеет **просить**, он умеет только **требовать**.
On ne umeet **prosit'**, on umeet tol'ko **trebovat'**.
he$_{NOM}$ NEG knows-how **ask$_{INFIN}$** he$_{NOM}$ knows-how only **demand$_{INFIN}$**
'He doesn't know how to ask; he only knows how to demand.'

In English, in contrastive sentences, like (24), these verbs can be used without an object in a nonmodal clause, but in noncontrastive contexts they cannot (**He demands every day, *He is demanding*). Thus, object necessity/optionality is more than a binary opposition.

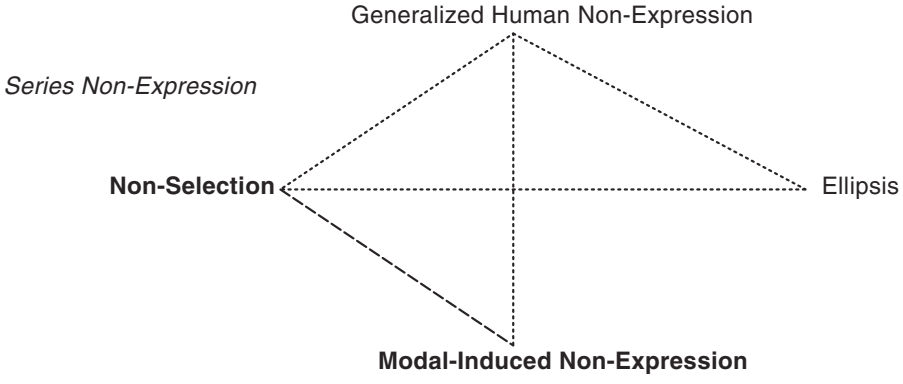


FIGURE 7.6 Nonselection or Modal-Induced Nonexpression?

Although this is an interesting lexicographic problem and although it might require special measures for knowledge-poor machine translation, it has little consequence for semantically rich text processing, since either analysis would trigger the search for the implied object in order to complete the semantic interpretation and representation.

9. Nonexpression of objects in a series of verbs

When several actions are presented in series with the focus on the verbs themselves, the objects of those verbs can often go unexpressed. While this is sometimes possible in English, especially in stylized literary texts, it is much more common in Russian:

(25) Она [Гингема] опустила в котёл большое помело и стала расплёскивать вокруг своё варево. «Разразись, ураган! Лети по свету, как бешеный зверь! **Рви, ломай, круши**!» (Волков: 7).
Ona [Gingema] opustila v kotël bol'šoe pomelo i stala rasplëskivat' vokrug svoë varevo. «Razrazis', uragan! Leti po svetu, kak bešenyj zver'! **Rvi, lomaj, kruši**!» (Volkov: 7).
{ . . . } burst-forth$_{IMPER}$ tornado fly$_{IMPER}$ across earth like wild beast **rip$_{IMPER}$ shatter$_{IMPER}$ destroy$_{IMPER}$**
'She [Gingema, a witch] stuck a big broom into the cauldron and started splashing her brew all over. "Burst forth, tornado! Fly across the land like a wild beast! Rip, shatter, destroy!"'

(26) «Я работаю как лошадь. Я **бегаю, хлопочу, очаровываю, ходатайствую, требую, настаиваю**» (Шварц 3: 540).
«Ja rabotaju kak lošad'. Ja **begaju, xlopoču, očarovyvaju, xodatajstvuju, trebuju, nastaivaju**» (Švarc 3: 540).
I$_{NOM}$ work like horse I$_{NOM}$ **run$_{1SG}$ bustle-about$_{1SG}$ charm$_{1SG}$ petition$_{1SG}$ demand$_{1SG}$ insist$_{1SG}$**
'"I work like a horse. I run, bustle about, charm people, petition, demand, insist."'

(27) Я вообще люблю участвовать в жизни других людей: **сватать, советовать, лечить**. **Лечить** мне пока не приходилось, а **сватать** и **советовать** — довольно часто (Токарева: 67).
Ja voobšče ljublju učastvovat' v žizni drugix ljudej: **svatat', sovetovat', lečit'. Lečit'** mne poka ne prixodilos', a **svatat'** i **sovetovat'** — dovol'no často (Tokareva: 67).
I$_{NOM}$ generally like participate$_{INFIN}$ in life other$_{GEN}$ people$_{GEN}$ **match-up$_{INFIN}$ advise$_{INFIN}$ treat$_{INFIN}$ treat$_{INFIN}$** me$_{DAT}$ so-far NEG necessary$_{IMPERS}$ but **match-up$_{INFIN}$** and **advise$_{INFIN}$** quite often
'In general I like to take part in the lives of other people: match them up, give them advice, nurse them back to health. I haven't had to nurse anyone yet, but I've quite often had occasion to match them up and give them advice.'

A full semantic respresentation would require that the understood objects of all of these verbs be accounted for, even if only in very general class-oriented terms (e.g., *give advice to HUMANs*).

10. Discussion

Because missing objects can be due to syntactic ellipsis or semantic ellipsis, all instances of missing objects in text must be handled not just by ellipsis-resolution algorithms, which seek *co*reference relations, but also by *reference* resolution algorithms, which amounts to specifying the generalized or specific object that people understand to be referred to in all such contexts. There will certainly be overlaps in analysis, and the one selected should best serve the application. In most cases, this will be the analysis that directly indicates a real-world referent rather than postponing that work for later analysis. Whether errors might result from this least-complexity approach remains to be seen.

Lambrecht and Lemoine (1996) independently developed an analysis of French null complements that intersects at certain points with this analysis of English and Russian. They posit three types of referents for null complements in French: *indefinite*, in which the referent is independent of the context and completely vague; *free*, in which the referent may be definite or indefinite depending on the context; and *definite*, in which the referent is contextually specified. Each of these meanings, they show, can be determined lexically or by given grammatical constructions (they work within Fillmore's Construction Grammar). Important for our purposes is the fact that a similar inventory of key issues derived from work on two typologically very different languages—for example, the necessity to exploit context for the interpretation of null objects, the reliance on world knowledge (frames or scripts) for such interpretation, the sometimes inconclusive results of such analysis, and the different object-omission properties of different lexical items. Moreover, they propose preliminary results of research into what factors underlie the choice between overt and elided realization of an object, hypothesizing (as do I) that relevant factors derive from morpho-syntax, semantics, and pragmatics. Although one can never hope to reach full agreement among researchers on any one framework or formalism—as can be seen, for example, by the number of computational dictionaries of English developed for different applications—some consensus on the "what" and "how" of description (for instance, using parameters and value sets) could ultimately help to foster exchange between similar lines of research applied to different languages.

8

Head Noun Ellipsis . . . or Not?

1. The facts

The ellipsis of head nouns is permitted in English when the head noun follows certain closed-class items, like numbers, quantifiers, and *one's own*.[1]

(1) It takes three **hours** to get to Albuquerque and four ∅ to get to Santa Fe.

(2) If you've got **Irish Cream**, I'd love some ∅.

(3) I don't need to rent **skis**; I've brought my own ∅.

Following adjectives, however, English *one*(*s*) must generally stand in for the head noun:

(4) I already have blue **shorts**, I need some green **ones/*∅**.

This use of *one*(*s*) does not represent syntactic ellipsis, since there is no gap in the syntax, but it does represent semantic ellipsis, since *one*(*s*) must be linked to a real-world referent for a full semantic representation.

In some languages, however, like Russian, adjectives are valid licensors of head noun ellipsis.[2] A comparison of usage shows much in common between head noun ellipsis licensed by a number/quantifier in English and head noun ellipsis licensed by an adjective in Russian.[3] The following are among the salient properties of head noun ellipsis licensed in either language by either type of category.

The inflection of the elided head noun need not match that of its antecedent

(5) He has one **small chore to do**, whereas she has many **Ø**.

(6) «Помимо экономических **проблем**, о которых теперь думает каждый, есть ещё одна **Ø**» (Богуславская: 6).
«Pomimo èkonomičeskix **problem**, o kotoryx teper' dumaet každyj, est' eščë odna **Ø**» (Boguslavskaja: 6).
apart-from economic$_{GEN}$ **problems$_{GEN.PL}$** about which$_{PREP}$ now thinks each-person$_{NOM}$ there-exists yet another$_{NOM}$ **Ø$_{NOM.SG}$**
'"Apart from economic problems, which everyone thinks about now, there is yet another problem."'

The antecedent can be syntactically overt or understood from the context

(7) [Pushing forward a bowl of candies] Take two ∅!

(8) [Choosing a puppy] Выбери этого большого ∅!
Vyberi ètogo bol'šogo ∅!
choose$_{IMPER}$ that$_{ACC}$ big$_{ACC}$ **Ø$_{ACC}$**
'Choose that big one!'

Ellipsis can be carried out cooperatively among speakers

(9) '"Have a little **ice cream**." "I'd rather have a lot **Ø**."'

(10) « . . . Я тебе просто хотел загадать **шараду**». — «**Шараду**?» — «Очень интересную **Ø**» (Войнович 2: 23–24).
« . . . Ja tebe prosto xotel zagadat' **šaradu**». — «**Šaradu**?» — «Očen' interesnuju **Ø**» (Vojnovič 2: 23–24).
I$_{NOM}$ you$_{DAT}$ simply wanted set$_{INFIN}$ **riddle$_{ACC}$** **riddle$_{ACC}$** very interesting$_{ACC}$ **Ø$_{ACC}$**
'". . . I just wanted to set you a riddle." "A riddle?" "A really interesting one."'

The structure of the noun phrases in question need not be parallel

In the following examples, the antecedent noun phrases have post nominal modifiers, whereas the elliptical noun phrases have prenominal modifiers:

(11) The **political convictions** of her in-laws were highly conservative, nothing like her **own Ø**.

(12) Чемодан сразу же по прибытии был отнесён в **комнату** Павлика, а не Яночкину **Ø** (Хмелевская 1: 37).
Čemodan srazu že po pribytii byl otnesën v **komnatu** Pavlika, a ne Janočkinu **Ø** (Chmielewska 1: 37).

suitcase$_{ACC}$ immediately PARTICLE after arrival was taken to **room$_{ACC}$** Pavlik$_{GEN}$ and NEG Yanochka's$_{ACC.ADJ}$ **∅$_{ACC}$**
'Right after they arrived home, they took the suitcase to Pavlik's room, not to Yanochka's ∅.'

Although there are no apparent syntactic restrictions regarding the placement of the antecedent with respect to the elided head noun, there are syntactic structures in which ellipsis is generally (often highly) preferred or, at least, always permitted. These include all types of coordinate structures, comparative structures, and instances when the antecedent and elliptical noun phrases are arguments of the same verb. All of these structures ensure close proximity of the antecedent to the elided category and thereby ensure recoverability of the latter's referent.

The two noun phrases are selected by the same verb

(13) He recognized the **voice** as his father's ∅.

(14) Я бы с удовольствием работал двадцать четыре часа в год, но тогда моя годовая **зарплата** равнялась бы недельной ∅ (Токарева: 127).
Ja by s udovol'stviem rabotal dvadcat' četyre časa v god, no togda moja godovaja **zarplata** ravnjalas' by nedel'noj ∅ (Tokareva: 127).
I$_{NOM}$ CONDIT with pleasure work twenty-four hours in year but then my yearly **salary** equal CONDIT weekly ∅
'I would be happy to work twenty-four hours a year, but then my yearly salary would equal my weekly one.'

The two noun phrases are coordinated

(15) It makes little difference if you read four **books** or five ∅.

(16) А шофёр, тот так и остался сидеть с открытым ртом, отчего были видны все его белые **зубы** и один золотой ∅ (Шинов: 73).
A šofër, tot tak i ostalsja sidet' s otkrytym rtom, otčego byli vidny vse ego belye **zuby** i odin zolotoj ∅ (Šinov: 73).
but driver$_{NOM}$ he$_{NOM}$ COMPOUND PARTICLE continued sit$_{INFIN}$ with open mouth from which were visible all his white$_{NOM}$ **teeth$_{NOM}$** and one gold$_{NOM}$ **∅$_{NOM}$**
'But the driver kept sitting there with his mouth open, from which you could see all his white teeth and one gold one.'

The two noun phrases are in coordinated verb phrases

(17) In that amount of time you can either race through many **novels** or carefully read a few ∅.

(18) А Ленка промолчала. Ей требовалось время, чтобы выйти из одного **состояния** и переместиться в другое ∅ (Токарева: 340).
A Lenka promolčala. Ej trebovalos' vremja, čtoby vyjti iz odnogo **sostojanija** i peremestit'sja v drugoe ∅ (Tokareva: 340).

but Lenka$_{NOM}$ remained-silent her$_{DAT}$ was-necessary time$_{NOM}$ in-order-to leave from one **state** and shift into another ∅
'But Lenka remained silent. She needed time to leave one state and shift to another.'

The two noun phrases are in coordinated clauses

(19) If you ask nicely, Dad will probably lend you 100 **dollars** and Mom should be good for another 50 ∅.

(20) Знаете, у теноров бывают такие тоненькие голоса, что когда услышишь их по радио, невольно возникает вопрос: «Это женщина поёт толстым **голосом** или же мужчина тоненьким ∅?» (Шинов: 37).[4]
Znaete, u tenorov byvajut takie tonen'kie golosa, čto kogda uslyšiš' ix po radio, nevol'no voznikaet vopros: «Èto ženščina poët tolstym **golosom** ili že mužčina tonen'kim ∅?» (Šinov: 37).
know$_{2PL.PRES}$ at tenors$_{GEN}$ can-be such thin voices$_{NOM}$ that when hear$_{2SG.PRES}$ them$_{ACC}$ on radio automatically arises question that-is woman$_{NOM}$ singing thick$_{INSTR}$ **voice**$_{INSTR}$ or PARTICLE man$_{NOM}$ thin$_{INSTR}$ **∅$_{INSTR}$**
'You know, some tenors have such thin voices that when you hear them on the radio you unwittingly wonder: "Is that a woman singing in a thick voice or a man in a thin one?"'

The two noun phrases are in a comparative structure

(21) A little **help** from a professional is better than a lot ∅ from a nonprofessional.

(22) Их **привычка** к стабильности так же прочна, как наша ∅ к нестабильности (Богуславская: 261).
Ix **privyčka** k stabil'nosti tak že pročna, kak naša ∅ k nestabil'nosti (Boguslavskaja: 261).
their **accustomedness$_{NOM}$** to stability$_{DAT}$ just as strong as our$_{NOM}$ **∅$_{NOM}$** to instability$_{DAT}$
'They were as accustomed to stability as we were to instability.'

The latter noun phrase further specifies the former one

(23) There's only two **weeks** before the show closes. Two and a half ∅, to be exact.

(24) «Господи, до чего жрать хочется! . . . Жизнь отдала бы за бутерброд с **колбасой**. С копчёной ∅, конечно . . . » (Антология: 44).
«Gospodi, do čego žrat' xočetsja! . . . Žizn' otdala by za buterbrod s **kolbasoj**. S kopčënoj ∅, konečno . . . » (*Antologija*: 44).
God to what-point eat$_{INFIN}$ is-desired life$_{ACC}$ give CONDIT for sandwich$_{ACC}$ with **sausage$_{INSTR}$** with smoked$_{INSTR}$ **∅$_{INSTR}$** of-course
'"Boy, am I ravenous! . . . I'd give anything for a sausage sandwich. Smoked sausage, of course . . ."'

2. Discussion

Since all of the aforementioned properties of head noun ellipsis hold for both English and Russian, even though these languages have different inventories of lexical

licensors and very different typological profiles, we can provisionally posit head noun ellipsis as a superparameter. This means that if a language employs head noun ellipsis, most of the basic properties would be expected to hold. Cross-linguistic parameterization comes, therefore, in the inventory of licensors and any necessary amendments to the basic properties. There are two justifications for introducing the idea of superparameter: it represents a simplifying shortcut when describing ellipsis patterns, and as progress on the processing of these elliptical patterns is made, it can be applied to other languages more readily when the core versus parameterizable properties are clearly delineated.

There is no simple computational algorithm for recovery of the referent for an elided head noun since there may or may not be a syntactic antecedent and, if there is one, it can be located anywhere in the preceding context (cf. VP Ellipsis, discussed in chapter 9, section 4). However, the generalizations listed earlier can guide the creation of an algorithm. For example, syntactic antecedents tend to be in the same sentence or, if not, in the preceding sentence. The antecedents themselves tend to be modified and, if so, their modifiers are semantically comparable to those of the elided category (e.g., both are numbers, quantifiers, or colors). In languages with morphological gender or other morphologically expressed inherent features of nouns, the gender, and so on, of the overt element(s) in the noun phrase can act as a clue for finding the antecedent.

Although in English the triggers for head noun ellipsis are relatively unambiguous, not so for all languages.[5] In Russian, for example, there can be ambiguity between two different types of entities that look like bare adjectives (i.e., adjectives without a head noun): real adjectives and diachronically substantivized adjectives. For example, the Russian for *blind* (*слепой/slepoj* [with inflectional variants]) can be an adjective that can modify any blind person or animal, or it can function as a full-fledged noun that means "blind man, woman, men, women." When using it as a noun, one does not start with collocations like *blind man* and then elide the head noun (although this might be what happened historically); one simply retrieves the noun from the lexicon. Therefore, disambiguation will be required each time a lexical item that has both adjectival and nominal status is encountered in the absence of an overt head noun.

An aside: English has a rule that permits adjectives to be changed into nouns that indicate the class of people described by the adjective: poor → the poor; needy → the needy; underprivileged → the underprivileged. This process can, in theory, be applied to any adjective, but there are some semantic constraints. For example, it would be rather odd to use *the wet*, *the stiff*, or *the frustrated* (in any but humorous contexts), since these are generally not thought of as classes of people but as temporary conditions. These instances of nominalization are easily diagnosed because, apart from "the + adjective" instances, English does not permit head noun ellipsis after adjectives.

It is the work of lexicographers to determine whether an adjectivelike form has solidified as a nominal or, when used without a head noun, signals head noun ellipsis. In some cases, this is quite straightforward based on heuristics like frequency of use, permanency of the state, trait, or condition (e.g., *unfortunate* [*person*] is a nominal in Russian whereas *sad* [*person*] is not), the inability of a head noun to be overtly

expressed (not necessary but common), and membership of the item in a semantic class known to contain deadjectival nominals. Some such classes are illustrated for Russian in table 8.1. Rows 1–2 show nominals with a human referent that can be used in the singular or plural; 3–4, plural nominals with a human referent; 5–6, neuter singular nominals with a generalized referent; 7–8, nominals with fixed gender that agrees with the formerly overt but now canonically unrealized head noun (the word *komnata* 'room' in each case); 9–10, entities that can have an overt or missing head noun, with the missing-head-noun variant being a nominal as well (like *perm* in English, which is not currently an elliptical variant of *permanent wave*).

Distinguishing fixed nominals from freely created elliptical structures is not always easy. One way to test for independent nominal status is to use the word in a discourse-initial sentence like (25):

(25) В бар вошёл слепой.
V bar vošël slepoj.
into bar walked **blind-man**$_{\textbf{NOM.NOUN}}$
'A blind man walked into the bar.'

If the word can be used as a noun with no supporting context to provide an antecedent for ellipsis, chances are it has full-fledged nominal status. However, virtually any Russian adjective can occur bare in a text if its head noun is implied by the context. For example, if two guys are walking down the street and catch sight of a pretty girl, one can say to the other:

(26) Посмотри на эту чернявую ∅! Ничего, а?
Posmotri na ètu černjavuju ∅! Ničego, a?
look$_{\text{IMPER}}$ at that dark-haired$_{\text{ACC}}$ ∅$_{\textbf{ACC.ADJ}}$ not-bad eh
'Hey, look at that dark-haired girl! Not bad, eh?'

Чернявая/černjavaja 'dark-haired' is not a noun; it is an adjective that is used with an elided head noun, *девушка/devuška* 'girl', which is supplied from the real-world context.

TABLE 8.1 Examples of deadjectival nominals in Russian

No.	*Adjective*	*Transliteration*	*Translation*
1	раненый (-ая, -ые)	ranenyj (-aja, -ye)	wounded person
2	больной (-ая, -ые)	bol'noj (-aja, ye)	sick person
3	большие	bol'šie	adults, grown-ups
4	некоторые	nekotorye	some people
5	плохое	ploxoe	something bad, unpleasant
6	общее	obščee	something in common
7	приёмная	priëmnaja	foyer
8	столовая	stolovaja	dining room/hall, lunchroom
9	выходной (день)	vyxodnoj (den')	day off
10	борзая (собака)	borzaja (sobaka)	borzoi (dog)

Determining which *inanimate* adjectivelike words have independent nominal status and which do not is also not simple. For example, whereas the four-volume *Academy Dictionary* (Evgen'eva, 1981) considers *вчерашнее/včerašnee* 'what happened yesterday' to be a separate nominal sense of the corresponding adjectival entry, it does not consider *утреннее/utrenee* 'what happened this morning' or *недавнее/nedavnee* 'what happened not long ago' to have independent nominal status.

In closing, the status of bare adjectives is another case of potential ambiguity, which lends reason to researching and developing practical approaches to head noun ellipsis and lexical description together.

9

Verbal Ellipsis with One Licensor

Verbs are, in a sense, the backbone of any clause, since they determine what other elements can and must co-occur with them: generally a subject, often a direct or indirect object, and optionally various types of adjuncts. In addition, verbs impose restrictions on the semantic nature of their co-occurring elements. For example, the English verb *polish* requires a human agent or machine capable of polishing, a theme with a polishable surface, and, optionally, an instrument of polishing—perhaps a rag and some car, shoe, silver, furniture, or leather polish.

There are four well-studied configurations in which English permits verbal ellipsis: Gapping, Stripping, Sluicing, and VP Ellipsis.[1] The properties of each, largely stable cross-linguistically, have been discussed at length in the literature. I will consider the well-documented and widely agreed upon properties to be, by now, common knowledge and, in listing them, will not indicate who first discovered them, amended them, and so on.[2] Since an inventory of properties defines each of these types of verbal ellipsis, they can be thought of as superparameters, whose values cross-linguistically are 'yes' (the language permits that type of ellipsis) and 'no' (the language does not permit it). Cross-linguistic variation in the realization of these types of ellipsis can be accounted for by parameterizing those properties that can have different values. For example, in order for VP Ellipsis to obtain, the elided verb phrase must be licensed by a certain type of lexical licensor—this is an immutable property. However, the types of valid lexical licensors vary from language to language.

Describing elliptical phenomena in terms of parameters, values, and, when applicable, superparameters promises to permit research and development efforts for one language to be applied—with the necessary parametric adjustments—to other

languages. Knowledge sharing and resource reusability is particularly attractive in those domains that have been under close investigation in NLP, like VP Ellipsis and Gapping.

In this chapter, for each subtype of verbal ellipsis, the major properties are listed along with English examples. Russian and other languages are used as a means of showing some of the parameterizable aspects of these elliptical strategies.

1. Gapping

Gapping is an elliptical process that renders unexpressed the verb and, optionally, other elements of the verb phrase in the latter clause(s) of a coordinate or comparative structure:

(1) The soprano **sang** the high notes and the tenor ∅ the low notes.

(2) The old man **looked** at his dog like a lover ∅ at his beloved.

Gapping is licensed by parallelism between the antecedent and ellipsis clauses. This means that the verb (and its object or objects) can be omitted in the second part of the sentence because that same verb (and its object or objects) is overtly specified in the first part of the sentence and the two parts are parallel. The properties of Gapping described in the literature include the following.

An elliptical gap must be surrounded by the overt sentence elements being juxtaposed

According to this rule, an elliptical gap cannot be clause initial or clause final, as shown in all examples in this section.

Gapping can occur in coordinate (*1*) *and comparative* (*2*) *structures but not in subordinate structures* (*3*)

(3) *Mary **cooked dinner** on Tuesday because Peter ∅ on Wednesday.

The antecedent must be overt in the preceding clause

That is, the antecedent cannot merely be implied by the extralinguistic context. For example, one cannot peer at the neighboring table in a restaurant and say to one's dinner companion (4a); however, one can say (4b):

(4) a. *The man at the next table ∅ the fish and the woman ∅ the chicken salad.
b. The man at the next table **ordered** the fish and the woman ∅ the chicken salad.

Actually, one might be able to say something like (4a) in English (and one can definitely say it in Russian), but only in response to a question like *Can you tell what*

anyone else has ordered?—in which case the verb to be recovered in the Gapping structure is provided by the preceding utterance.

The antecedent must precede (*4b*)*, not follow* (*4c*)*, the gap*

(4) c. *The man at the next table ∅ the fish and the woman **ordered** the chicken salad.

This property, however common cross-linguistically, does not hold for every language. For example, in Hindi, so-called backward Gapping is permitted, in which the antecedent clause follows the Gapped clause (Sjoblom 1980). Backward Gapping has been attributed to Russian as well (Ross 1970), but the supporting examples are not universally accepted by native speakers.

The antecedent verb and the Gapped verb must match semantically, but they need not match morphologically

This means that both clauses must contain the same lexical entity, but it can have different values for gender, number, and so on. This is shown clearly on morphologically rich Russian:

(5) Мишка **сидел** впереди, а Танька ∅ сзади, обхватив Мишку поперёк живота (Токарева: 470).
Miška **sidel** vperedi, a Tan'ka ∅ szadi, obxvativ Mišku poperëk života (Tokareva: 470).
Mishka$_{NOM}$ **sat**$_{3SG.MASC}$ in-front and Tan'ka ∅$_{3SG.FEM}$ behind having-grabbed Mishka around waist
'Mishka sat in front and Tan'ka ∅ behind, wrapping her arms around Mishka's waist.'

In comparative Gapping structures, the antecedent verb and the Gapped verb need not match with respect to tense, aspect, or mood (*a property not in the canon but important for full semantic analysis*)

Such a mismatch commonly occurs when the antecedent clause presents a specific action and the Gapped clause presents a generalized one. For example, in (6) the verb in the antecedent clause is in the indicative mood, past tense—'looked'—whereas the elided verb would be either in the conditional mood or in the generalizing present tense:'would look/looks/looking':

(6) Подберёзовиков **поглядел** на помощницу, как редактор ∅ — на опечатку (Брагинский и Рязанов: 165).
Podberëzovikov **pogljadel** na pomoščnicu, kak redaktor ∅ [= gljadit/pogljadel by] — na opečatku (Braginskij i Rjazanov: 165).
Podberezovikov$_{NOM}$ **looked$_{3SG.MASC}$** at assistant like editor ∅ [looks$_{3SG}$ /look$_{3SG.MASC}$ CONDIT] at typo
'Podberezovikov looked at his assistant like an editor (looking) at a typo.'

Gapping is recursive

(7) Jane's birthday is in May, John's ∅ in June, and Rex's ∅ in July.

Remnants must be major constituents

In minimal pair (8a)–(8b), the prepositional phrase *with their dogs* is a major constituent. The words that comprise the prepositional phrase cannot be split up such that some are included in the gap and others are not. So, (8a) is fine because the whole prepositional phrase is outside of the gap (i.e., overt), while (8b) is bad because it chops up the prepositional phrase, Gapping part of it and leaving the other part overt:

(8) a. We **came** with our cats and they ∅ with their dogs.
b. *We **came with our** cats and they ∅ dogs.

There can be more than two elements contrasted in Gapping structures

(9) Peter **talked** to his boss on Tuesday and Betsy ∅ to her supervisor on Wednesday (Merchant 2001: 82 [empty category and boldface added]).

The elements juxtaposed in a Gapping structure must be semantically comparable

In other words, one cannot compare apples and oranges. In the following ungrammatical example, the second clause was intended to mean that "the queen of England rides her bike in the park," but since riding one's bike *with pleasure* and riding one's bike *in the park* are not logically comparable, the structure fails on semantic grounds:

(10) *I **ride my bike** with pleasure, and the queen of England ∅ in the park.

Although the compared/contrasted categories must be semantically comparable, they need not match with respect to part of speech. In (11), for example, the adverb *logically* is legally juxtaposed with the PP-adverbial *in practical terms*:

(11) You **are misguided** logically and your partner ∅, in practical terms.

Gapping can be employed cooperatively

(12) "Mitya **plays** the clarinet." "And Nina ∅ the piano."[3]

Several infinitivals in series can be Gapped

(13) John **was planning to start to study** Latin and Sasha ∅ French.

In fact, a Gap can include few or many of the verbal elements that occur in series, as shown in the variants of (14):[4]

(14) a. He **wants** to try to begin to write a novel and she ∅ to try to begin to write a play.
b. He **wants to try** to begin to write a novel and she ∅ to begin to write a play.
c. He **wants to try to begin** to write a novel and she ∅ to write a play.
d. He **wants to try to begin to write** a novel and she ∅ a play.

However, when the series of Gapped infinitives is long, some of the infinitives might not be reconstructed as part of the gap—a practical matter of memory rather than grammar.

Symmetric, not cause-effect, clause relations are implied in Gapping structures

As Levin and Prince (1986) note, Gapping does not preclude the fact that there could be some causal relationship between clauses, but there is nothing that forces that inference. So, whereas in (15a) both a symmetric and a cause-effect interpretation are possible, in (15b) the symmetric interpretation is the default, as in a context where the dogs react to a thunderclap:

(15) a. The Doberman **raced** into the kitchen and the poodle **raced** up the stairs.
b. The Doberman **raced** into the kitchen and the poodle ∅ up the stairs.

Although the basic Gapping strategy is the same cross-linguistically, there are parameterizable aspects. One, shown earlier, was the possibility or impossibility of backward Gapping. Three others, exemplified by Russian-English comparisons, are introduced here.

1.1 Gapping is stylistically restricted in English but not in Russian

In English, Gapping is most common in rather formal or planned speech, while in Russian it is used liberally in all speech registers. The following examples show stylistically neutral Gapping in Russian that is translated by non-Gapping structures in English:

(16) «Я **доложу о вас** прямо первому министру, а он ∅ королю» (Шварц 1: 87).
«Ja **doložu o vas** prjamo pervomu ministru, a on ∅ korolju» (Švarc 1: 87).
I_{NOM} will-inform about you directly first $minister_{DAT}$ and_{CONTR} he_{NOM} ∅ $king_{DAT}$
‘‘I’ll announce you directly to the first minister, and he’ll announce you to the king.’’

(17) «Волшебная палочка подобна дирижёрской. Дирижёрской — **повинуются** музыканты, а волшебной ∅ — всё живое на свете» (Шварц 3: 545).
«Volšebnaja paločka podobna dirižërskoj. Dirižërskoj — **povinujutsja** muzykanty, a volšebnoj ∅ — vsë živoe na svete» (Švarc 3: 545).

magic stick$_{NOM}$ similar-to conductor's conductor's$_{DAT}$ **obey$_{3PL}$** musicians$_{NOM}$
whereas magic$_{DAT}$ Ø everything$_{NOM}$ living$_{NOM}$ on earth
'"A magic wand is like a conductor's baton. The baton has power over musicians, while the magic wand has power over all living things."'

1.2 Morphologically realized case marking in Russian expands Gapping potential

Having morphological case marking on noun phrases is not a prerequisite for Gapping: after all, English permits Gapping despite its lack of case marking. However, there is one type of Gapping example that requires case marking and thus is found in Russian but not in English.[5]

Note: When reading the examples in this section, be sure to look at the English glosses (in single quotes) to see what the sentence is intended to mean. The grammaticality judgments are based on that interpretation.

Consider example (18), in which Gapping is fine in Russian but fails in English:

(18) Мама **попросила Мишу** спеть, а отец Ø — сыграть на гитаре.
Mama **poprosila Mišu** spet', a otec Ø — sygrat' na gitare.
*Mom **asked Misha** to sing and Dad Ø to play the guitar.
'Mom asked Misha to sing, and Dad asked Misha to play the guitar.'

A Gapping interpretation is possible in Russian because the NOM case marking on *отец/otec* 'Dad' unambiguously indicates that it is subject of the Gapped clause. In English, by contrast, there is no way to signal that *Dad* is the subject—it is, by default, interpreted as the direct object.[6]

There is, however, a grammatically valid interpretation of the English string of words *Mom asked Misha to sing and Dad to play the guitar*. Under this interpretation, the mother is making one request of Misha ("please sing") and another request of the father ("please play the guitar"). This makes both *Misha* and *Dad* objects of *ask*. Under this interpretation, the sentence contains no gaps—it is composed of a subject, verb, and coordinated complements: *Mom asked* [[*Misha to sing*] *and* [*Dad to play the guitar*]]. Such an interpretation is, of course, available in Russian as well, as long as *Dad* is case-marked ACC (*отца/otca*):

(19) Мама попросила Мишу спеть, а отца — сыграть на гитаре.
Mama poprosila Mišu spet', a otca — sygrat' na gitare.
Mom$_{NOM}$ asked Misha$_{ACC}$ to sing and Dad$_{ACC}$ to play the guitar
'Mom asked Misha to sing and (she asked) Dad to play the guitar.'

To reiterate, Russian has two interpretations of the string of words in (18)–(19), based on the case marking of *Dad* (*отец$_{NOM}$* vs. *отца$_{ACC}$*). English has only one interpretation, in which *Dad* is understood to be the object of the verb *ask*.

There are, however, instances in which Russian's case marking fails to distinguish between the subject and the object—for example, when the nouns in question are indeclinable foreign names. In such sentences, Russian, like English, strongly favors the non-Gapping interpretation as the default:

(20) a. ?Марго **заставила Бруно** починить машину, а Педро ∅ — починить грузовик.
?Margo **zastavila Bruno** počinit' mašinu, a Pedro ∅ — počinit' gruzovik.
*Margo **made Bruno** fix the car and Pedro ∅ fix the truck.
'Margo made Bruno fix the car and Pedro made Bruno fix the truck.'

(20) b. Марго заставила Бруно починить машину, а Педро — починить грузовик.
Margo zastavila Bruno počinit' mašinu, a Pedro — počinit' gruzovik.
Margo made Bruno fix the car and Pedro fix the truck.
'Margo made Bruno fix the car and (she made) Pedro fix the truck.'

Technically, the Gapping interpretation shown in (20a) is available in Russian, but only if strong contextual cues force it; for example, if Pedro was out of the country and therefore could not fix the truck but could demand over the telephone that Bruno do it. In English, by contrast, no degree of contextual support can produce a Gapping interpretation for (20a).

1.3. Conjunctionless Gapping

The most typical examples of Gapping contain an overt conjunction that links the clauses. However, Gapping structures do not require a conjunction. Conjunctionless Gapping is, in fact, quite common in Russian and is sometimes possible in English as well. (Language-wide, Russian permits conjunctions to be left out much more than English does.) In (21), both Russian and English permit Gapping without an overt conjunction, although in English the utterance is stylistically marked. In both languages, the Gapped structures must be pronounced with contrastive intonation and a significant pause must signal the gap:

(21) Моя щека **помнит** его щёку. Рука ∅ — его плечо (Токарева: 73).
Moja ščeka **pomnit** ego ščëku. Ruka ∅ — ego plečo (Tokareva: 73).
my cheek$_{NOM}$ **remembers** his cheek$_{ACC}$ hand$_{NOM}$ ∅ his shoulder$_{ACC}$
'My cheek remembers his cheek. My hand, his shoulder.'

Often Russian permits conjunctionless Gapping in contexts where English does not. For example, the Russian variants of (22)–(23) are perfectly natural, but direct English translations would sound awkward—except if one wanted to create stylistic effects, as in creative writing. Typical English translations require some sort of restructuring: having the second verb overt, inserting a conjunction, or fundamentally rewriting the sentence.

(22) Капитан Галкин никогда больше не появился. Мама **думала** трагически и возвышенно — потому что убит, бабушка ∅ — приземлённо — потому что бежал с часиками (Рачко: 46–47).
Kapitan Galkin nikogda bol'še ne pojavilsja. Mama **dumala** tragičeski i vozvyšenno — potomu čto ubit, babuška ∅ — prizemlënno — potomu čto bežal s časikami (Račko: 46–47).

Captain Galkin$_{NOM}$ never more NEG appeared Mom$_{NOM}$ **thought** tragically and loftily because was-killed$_{MASC.SG}$ Grandma$_{NOM}$ ∅ practically because ran-off$_{MASC.SG}$ with watch
'Captain Galkin never showed up again. Mom thought, tragically and loftily, that he had been killed; Grandma, being more practical, thought that he had run off with the watch.'

(23) «Очень трудно налаживать нормальные отношения, когда существуют вполне устоявшиеся взгляды на то, какой **должна быть** женщина, каким ∅ — мужчина» (Богуславская: 6–7).
«Očen' trudno nalaživat' normal'nye otnošenija, kogda suščestvujut vpolne ustojavšiesja vzgljady na to, kakoj **dolžna byt'** ženščina, kakim ∅ — mužčina» (Boguslavskaja: 6–7).
very difficult$_{IMPERS}$ create$_{INFIN}$ normal relations$_{ACC}$ when exist totally fixed ideas$_{NOM}$ about PARTICLE what-kind$_{INSTR}$ **must be** woman$_{NOM}$ what-kind$_{INSTR}$ ∅ man$_{NOM}$
'"It's very hard to normalize relations when there exist totally fixed ideas about how a woman should be and how a man should be."'

One final point regarding conjunctionless Gapping deserves mention. Some Gapping sentences in Russian that may appear to lack a conjunction actually do have one; it is simply not located between the clauses. For example, *же/же* is a clitic conjunction that occupies the second position in the Gapped clause, as shown in (24):

(24) К ней подходил высокий молодой человек, как я заключил, с целью пригласить её; он **был** от неё в двух шагах, я же ∅ — на противоположном конце залы (Толстой 2: 78).
K nej podxodil vysokij molodoj čelovek, kak ja zaključil, s cel'ju priglasit' eë; on **byl** ot neë v dvux šagax, ja že ∅ — na protivopoložnom konce zaly (Tolstoj 2: 78).
toward her$_{DAT}$ approached tall young man$_{NOM}$ as I$_{NOM}$ inferred with goal invite$_{INFIN}$ her$_{ACC}$ he$_{NOM}$ **was** from her$_{GEN}$ at two paces I$_{NOM}$ PARTICLE ∅ on other end hall$_{GEN}$
'A tall young man approached her, I gathered with the intention of asking her to dance; he was two steps away from her, while I was at the other end of the hall.'

Based on these English–Russian comparisons, Gapping is parameterizable according to at least the properties shown in table 9.1.

TABLE 9.1 Parameterizable aspects of Gapping

Parameters	*Values*
Backward Gapping	Possible, Impossible
Frequency of use	High, Low
Stylistic marking	Marked, None
Morphologically supported expanded Gapping potential?	Yes, No
Conjunctionless Gapping	Common, Rare, Impossible

2. Stripping

Stripping, like Gapping, is licensed by structural parallelism between the antecedent clause and the ellipsis clause. This elliptical process strips away all but one main constituent in the ellipsis clause under identity with the antecedent clause. Nonmain constituents, like adverbs or negation, can be overt in the Stripped clause as well. For example, in the second clause of (25) all categories are stripped away except for the main constituent *relatives* and the nonmain constituent *sometimes*:

(25) Neighbors often **come to visit her** and sometimes relatives ∅.

The following are among the fundamental properties of Stripping that appear to hold cross-linguistically.

The antecedent must precede, not follow, the elided category

(26) a. Mary **likes to call** Ivan but not ∅ Peter.
b. *Not ∅ Peter, but Mary **likes to call** Ivan.

Stripping cannot occur in a subordinate clause

(27) *This is the classroom where **we** usually **study** biology, and there is the classroom where sometimes ∅ chemistry.

Stripping can be realized cooperatively

(28) "Mary **likes to call** Ivan." "But not ∅ Peter?"

I have not discovered any parameterizable aspects of Stripping cross-linguistically. However, at a minimum, the parts of speech that can act as the main overt constituent in a Stripped clause and as nonmain facultative elements must be listed for each language.

3. Sluicing

Sluicing is a term introduced by Ross (1969) to describe sentences in which an interrogative clause is elided leaving only its *wh*-word (or phrase) overt. (See Merchant 2001 for extensive discussion of Sluicing cross-linguistically along with theoretical contributions.) In terms of licensing, this type of ellipsis is quite different from Gapping or Stripping: whereas Gapping and Stripping are licensed by structural parallelism between the antecedent and ellipsis clauses, the ellipsis of embedded questions is licensed by the preceding *wh*-word. So, in (29) the *wh*-word *who* licenses the ellipsis of the verb *to ask*; and in (30) the *wh*-word *why* licenses ellipsis of the verb phrase *dropped out of grad school.* The ellipsis licensors are underlined in the examples for emphasis.

(29) We need **to ask** someone, but we don't know <u>who</u> ∅.

(30) We know that **Peter dropped out of grad school**, but we don't know <u>why</u> ∅.

The key properties of Sluicing are as follows.

A wh*-element must license the ellipsis*

Other lexical items that can occupy the complementizer position cannot license Sluicing, like *that*, *whether*, *if*, and *for*.

The referent can precede the empty category, as in (*29*)–(*30*), *or follow it, as in* (*31*)

(31) Although I can't imagine why ∅, **someone stole my swimsuit from the locker room**.

The referent can be overt in the syntax or pragmatically understood

(32) [Being shown a dress by a saleswoman]
If you tell me how much ∅, I'll tell you if I like it.[7]

A phrase, not a head, is the target of ellipsis

For example, (33) is ungrammatical because only the head *bought* is included in the sluice rather than the verb phrase *bought the painting*:

(33) *Even though we aren't sure [who ∅ the painting], we know that someone **bought** the van Gogh (Lobeck 1995: 24).[8]

The case marking of the overt element in the Sluiced wh*-phrase must match the case marking of its referent* (*relevant only for case-marking languages, of course*)

(34) Л**иза сердится** на кого-то, но не знаю на кого ∅.
Liza serditsja na kogo-to, no ne znaju na kogo ∅.
Liza is-mad at someone$_{ACC}$ but NEG know$_{1SG}$ at whom$_{ACC}$ ∅
'Liza is mad at someone, but I don't know who.'

Parameterizable aspects of Sluicing include the following.

Preposition stranding is possible in sluices only in those languages in which it is possible under regular wh*-movement*

This is a rephrasing of Merchant's 2001 "Form-identity generalization II: Preposition-strandings" which states: "a language *L* will allow preposition-stranding under sluic-

ing iff *L* allows preposition-stranding under regular wh-movement" (p. 107). So preposition stranding is possible in English (*at someone* in the antecedent of [35] can be reduced to *who*—without 'at'—in the sluice) but not in Russian (*na kogo-to* 'at someone' in the antecedent cannot be reduced to *kogo* 'someone' in the sluice):

(35) *Л**иза сердится** на кого-то, но не знаю кого ∅.
*__Liza serditsja__ na kogo-to, no ne znaju kogo ∅.
Liza is-mad at someone but NEG know$_{1SG}$ who$_{ACC}$ ∅
'Liza got mad at someone, but I don't know who.'

Languages that permit the fronting of multiple wh-*words should permit Sluicing with multiple* wh-*words*

This generalization and Bulgarian example (36) are due to Merchant 2001 (109–110).

(36) Njakoj e vidjal njakogo, no ne znam [$_{CP}$ koj kogo ~~[$_{IP}$ e vidjal]~~].
someone AUX seen someone but not I know who whom AUX seen
'Someone saw someone, but I don't know who saw who.'

Even some languages that do not permit the fronting of multiple wh-*words permit multiple Sluicing*

This generalization and German example (37) are due to Merchant 2001 (110–111).

(37) Jemand hat was gesehen, aber ich weiß nicht, wer was.
someone has something seen but I know not who what
(lit.) 'Someone saw something, but I don't know who what.'

To sum up, Sluicing is parameterizable according to at least the parameters shown in table 9.2 (a more thorough incorporation of the work of Merchant 2001, which lies outside the scope of the current study, should add more parameters and values to this inventory).

TABLE 9.2 Parameterizable aspects of Sluicing

Parameters	*Values*
Does the language have morphological case marking (therefore requiring case matching in Sluicing)?	Yes, No
Is preposition stranding possible in the language in general?	Yes, No
It is preposition stranding possible in sluices?	Yes, No
Can multiple *wh*-words be fronted in general?	Yes, No
Can multiple *wh*-words be involved in a sluice?	Yes, No

4. Verb Phrase Ellipsis

VP Ellipsis is the ellipsis of a whole verb phrase, which includes the verb and its objects or adjuncts. VP Ellipsis is licensed by the immediately preceding auxiliary, underlined in the following examples:

(38) a. Jack doesn't **eat meat**, but Victor does ∅.
b. Terry **reads French novels**, but Barry doesn't ∅.
c. Mom isn't **going running today**, but I am ∅.
d. I'm forcing her to **do it** because I know she can ∅.
e. It's not that I can't **ask him for help**; it's that I don't want to ∅.
f. "I'm not **going**." "Are you sure? It'll be fun." "I'm sure." "OK, if you aren't ∅, then I'm not, either."
g. [Eyeing two slabs of chocolate cake] I will ∅ if you will ∅.

The following are the notable properties of VP Ellipsis.

If the antecedent is overt in the syntax, it need not immediately precede the ellipsis clause

Example (38f) supports this generalization.

If the referent is syntactically accessible, it need not be syntactically identical to the elided category

For example, in (39) the antecedent contains *himself*, whereas the reconstructed category would have to contain *him*:

(39) Bill$_i$ **defended himself$_i$** against the accusations because his lawyer$_j$ couldn't ∅ (Kehler 2002: 55; example formatting changed).

The antecedent may be understood from the real-world context

This is shown in (38g) earlier and (40):

(40) [The speaker, about to jump off the high diving board]
"Oh, I just can't ∅!"

The elided verb phrase may precede the antecedent

(41) It's your fault. I know that you didn't mean to ∅, but you **did it**.

The ellipsis clause can be subordinate to the antecedent clause

(42) We're not **going to that movie** because my husband doesn't want to ∅.

The language's inventory of auxiliaries is central

Since VP Ellipsis is licensed by auxiliaries, the contexts in which ellipsis may be employed depend upon the available auxiliaries. English and Russian, for example, have some, but not all, auxiliaries in common. The ones they have in common include *мочь/moč'* 'can', *уметь/umet'* 'know how (to)', *собираться/sobirat'sja* 'plan to', *начинать* (*начать*) */ načinat'* (*načat'*) 'start', *перестать* 'stop', *прекращать* (*прекратить*) */ prekraščat'* (*prekratit'*) 'stop', *быть/byt'* (in future forms) 'will', *хотеть/xotet'* 'want to'.

(43) «Бывает, конечно, что хочется **посидеть дома**, но по экономическим соображениям не можешь ∅» (Богуславская: 7).
«Byvaet, konečno, čto xočetsja **posidet' doma**, no po èkonomičeskim soobraženijam ne možeš' ∅» (Boguslavskaja: 7).
happens of-course that want$_{3SG}$ **sit$_{INFIN}$ home** but for economic coniderations NEG can$_{2SG}$ ∅$_{INFIN}$
'"It sometimes happens, of course, that you want to stay home [i.e., be a stay-at-home mom] but for economic reasons you can't."'

(44) «Я думала, что после этого он всё будет **заходить** чаще и чаще, — не тут-то было. Он почти совсем перестал ∅» (Достоевский 1: 142).
«Ja dumala, čto posle ètogo on vsë budet **zaxodit'** čašče i čašče, — ne tut-to bylo. On počti sovsem perestal ∅» (Dostoevskij 1: 142).
I$_{NOM}$ thought that after this he PARTICLE will **stop-by$_{INFIN}$** more-often and more-often not at all (idiom) he$_{NOM}$ practically stopped ∅$_{INFIN}$
'"I thought that after that he'd stop by more and more often, but not even close. He practically stopped altogether."'

Although Russian and English have the aforementioned auxiliaries in common, there are many auxiliaries that are found in only one or the other language. For example, Russian has a number of impersonal predicate words that function like auxiliaries in their ability to license VP Ellipsis: for example, *можно/možno* '(it-is)-possible, (one)-may'; *нельзя/nel'zja* '(it-is)-impossible, (one)-must-not'; *надо/nado* '(it-is)-necessary, (one)-must'. English lacks impersonal predicate words and must use other auxiliaries instead, as shown in (45)–(47):

(45) «А я думала, чего доброго, не **попрошавшись уедете**». — «Ну что вы, как можно ∅!» (Вампилов 5: 379–380).
«A ja dumala, čego dobrogo, ne **poproščavšis' uedete**». — «Nu čto vy, kak možno ∅!» (Vampilov 5: 379–380).
but I$_{NOM}$ thought for-all-I-knew NEG **having-said-goodbye will-leave$_{3SG}$** what-are-you-talking-about how is-possible$_{IMPERS}$ ∅$_{INFIN}$
'"I thought you'd left without saying good-bye." "What are you talking about, how could I?!"'

(46) «Они **развели конспирацию**, вот я и решил, что нам тоже надо ∅ . . . » (Хмелевская 4: 40).

«Oni **razveli konspiraciju**, vot ja i rešil, čto nam tože nado ∅ . . . » (Chmielewska 4: 40).
they$_{NOM}$ **established conspiracy$_{ACC}$** so I$_{NOM}$ PARTICLE decided that us$_{DAT}$ also necessary$_{IMPERS}$ **∅$_{INFIN}$**
‘“They’ve gotten a conspiracy going, so I figure we should, too . . .”’

(47) Но он понятия не имел, о чём можно **с ней говорить** и о чём нельзя ∅ . . . (Золотые: 76–77).
No on ponjatija ne imel, o čëm možno **s nej govorit’** i o čëm nel’zja ∅ . . . (Zolotye: 76–77).
but he$_{NOM}$ idea$_{GEN}$ NEG had about what possible$_{IMPERS}$ **with her talk$_{INFIN}$** and about what impossible$_{IMPERS}$ **∅$_{INFIN}$**
‘But he had no idea what he could talk to her about and what he couldn’t . . .’

English, for its part, has a number of auxiliaries that Russian lacks: *does/doesn’t, is/isn’t, has/hasn’t*, and so on. For example, in (48) the English auxiliary *didn’t* licenses ellipsis of the verb phrase *know*. Russian, lacking this option, employs a *нет/net* ‘not’ construction that arguably does not represent syntactic ellipsis. The *нет/net* construction differs from typical elliptical constructions in that the verb cannot be reinserted: *. . . *а бабушка нет знала/a babuška net znala/but Grandma not knew* is ungrammatical. Therefore, although this represents ellipsis, it is semantic ellipsis, not syntactic ellipsis.

(48) Я знала шестым чувством, что такие люди часиков не крадут, и мама знала, а бабушка — нет (Рачко: 46).
Ja znala šestym čuvstvom, čto takie ljudi časikov ne kradut, i mama znala, a babuška — net (Račko: 46).
I$_{NOM}$ knew sixth sense$_{INSTR}$ that such people$_{NOM}$ watches$_{ACC}$ NEG steal and Mom$_{NOM}$ knew but Grandma$_{NOM}$ no
‘I knew by some sixth sense that people like him don’t steal watches, and Mom knew, but Grandma didn’t.’

(49) «Меня всегда интересовало, почему одним удаётся осуществить свои мечты, повысить своё благосостояние, а другим — нет» (Богуславская: 226).
«Menja vsegda interesovalo, počemu odnim udaëtsja osuščestvit’ svoi mečty, povysit’ svoë blagosostojanie, a drugim — net» (Boguslavskaja: 226).
me$_{ACC}$ always interested why some(-people)$_{DAT}$ succeeds$_{3SG}$ realize$_{INFIN}$ self’s dreams$_{ACC}$ raise$_{INFIN}$ self’s standard-of-living$_{ACC}$ and other(-people)$_{DAT}$ no
‘“I’ve always wondered why it is that some people can fulfill their dreams and raise their standard of living while others can’t.”’

Other parameterizable aspects of VP Ellipsis (that is, apart from the language-specific inventory of ellipsis-licensing auxiliaries) come from the literature devoted to the more difficult aspects of formally delineating meaning-recoverability strategies for elided verb phrases in English. Some examples follow.

A language may or may not permit antecedent-contained VP ellipsis

Antecedent-contained VP ellipsis is the ellipsis of a verb phrase within the verb phrase that contains its antecedent. English permits this type of ellipsis, as shown by examples (50a)–(50b), drawn from Lappin 1999, p. 1 (a source that provides a Head-driven Phrase Structure Grammar account of antecedent-contained ellipsis and a bibliography of relevant work):

(50) a. John read every book which Mary did.
b. Mary stood beside every painting which John did.

Russian, by contrast, does not permit this type of ellipsis, to the point of making it impossible to even posit ungrammatical examples for illustration.

A language may or may not permit multiple VP ellipsis

Klein and Stainton-Ellis 1989 delineates three types of multiple VP ellipsis—Nesting, Crossing, and Mixed—all of which English permits but Russian does not, as shown by the contrast between the (a) and (b) variants of (51)–(53), respectively.[9] The Russian variants show either repetition of the given verbs (51b), a fixed construction (52b), or a combination of same-verb repetition and use of other referential verbs (lit.: *I didn't* ***do that***) (53b):

(51) a. I promised myself I [wouldn't go to the park]$_1$ until I [had sorted my papers]$_2$. I didn't $\varnothing_2$, so I didn't $\varnothing_1$.
b. Ja rešila, čto ne poidu v park, poka ne razberu bumagi. Ne razobrala, značit, ne pošla.
I_{NOM} decided that NEG will-go to park until PARTICLE will-sort papers NEG sorted therefore NEG went

(52) a. If you [work hard]$_1$, you'll [make progress]$_2$. If you don't $\varnothing_1$, you won't$\varnothing_2$.
b. Esli budeš' rabotat' kak sleduet, dob'eš'sja uspexa. Esli net, to net.
if $will_{2SG}$ $work_{INFIN}$ as is-necessary will-achieve progress if not then not

(53) a. I [was thin]$_1$ then and tried on some designer jeans that [looked good on me]$_2$ and I [should have bought them]$_3$. But I didn't $\varnothing_3$ and now I'm not $\varnothing_1$, and they wouldn't $\varnothing_2$.
b. Ja togda byla xudaja i pomerila dizajnerskie džinsy, v kotoryx ja vygljadela xorošo; nado bylo ix kupit'. No ja ètogo ne sdelala, a teper' ja ne xudaja, i oni na menja ne nalezut.
I_{NOM} then was thin and tried-on designer jeans in which I_{NOM} looked good $necessary_{IMPERS}$ was $them_{ACC}$ buy_{INFIN} but I_{NOM} that NEG did and now I_{NOM} NEG thin and $they_{NOM}$ onto me NEG will-fit

Sentences that show the well-known strict versus sloppy anaphora ambiguity may or may not represent VP ellipsis as such

This is shown by the English-Russian contrast in (54), in which Tom loving his own sister represents the strict interpretation whereas Tom loving Mike's sister represents the sloppy one:

(54) a. Mike loves his sister and Tom does, too.
b. Majk ljubit svoju sestru i Tom tože.
Mike$_{NOM}$ loves self's sister$_{ACC}$ and Tom$_{NOM}$ also

The Russian variant, like the English one, shows syntactic ellipsis and permits both interpretations, but the overt categories in the elliptical conjunct are quite different—there is no auxiliary to license the VP ellipsis, and the overt categories, in fact, are more typical of a Stripping structure. This surfacy contrast is particularly important because the strict-versus-sloppy problem has received much attention in the computational and theoretical literature.[10] Once approaches to the semantic-resolution resolution issues are developed, they should be applied not only to VP Ellipsis but also to other syntactic realizations of verbal ellipsis, like the Russian one shown in (54b).

The parameterizable aspects of VP Ellipsis are summarized in table 9.3.

One of the key debates, both theoretically and computationally, is whether the elided category should be reconstructed syntactically or semantically. Kehler (2002), for example, presents analysis that supports the claim that contexts that show a Resemblance relation are subject to a syntactic account, whereas contexts that showing a Cause-Effect relation must be analyzed semantically. However, since heuristics for determining which discourse relation applies to which sentences remains an outstanding—and extremely difficult—research issue, this theoretically interesting insight would be difficult to apply to practical systems.

In the Ontological Semantic approach to language processing, semantic analysis will be supported by syntactic and other heuristics. For instance, examples like (38a)–(38c), repeated here for convenience, are quite constructionlike and can most simply (although, granted, not with 100% accuracy) be handled by algorithms like the one shown in (55):

TABLE 9.3 Parameterizable aspects of VP Ellipsis

Parameters	*Values*
What types of auxiliarylike elements can license VP Ellipsis?	Inflected verb forms Impersonal predicates
Does the language have a negation strategy similar to the Russian *net* strategy that replaces VP Ellipsis in some instances?	Yes, No
Is antecedent-contained VP ellipsis permitted?	Yes, No
Is multiple VP Ellipsis permitted?	Yes, No
Are structures that show strict-sloppy readings of anaphora actually VP Ellipsis or some other strategy (like Russian *tože* 'too' in [54])?	VP Ellipsis Other

(38) a. Jack doesn't **eat meat**, but Victor does **∅.**
b. Terry **reads French novels**, but Barry doesn't **∅.**
c. Mom isn't **going running today**, but I am **∅.**

(55) If an elided verb phrase is detected
And if the elided verb phrase is located in the second conjunct of a contrastive coordinate structure
And if the subjects of each clause are semantically comparable [determined ontologically]
Then corefer the elided verb phrase with the verb phrase in the preceding conjunct.
Else, go to algorithm *x*.

Similarly, as Kehler (2002: 58) points out, *and . . . too* constructions almost always require sloppy identity between the elided verb phrase and its syntactic antecedent, permitting the construction of another useful heuristic.[11] For example, algorithm (57) might be used to analyze sentences like (56b):

(56) a. *John$_i$ defended himself$_i$ and Bob$_j$ did too. [defend himself$_i$] (Kehler 2002: 58).
b. John$_i$ defended himself$_i$ and Bob$_j$ did too. [defend himself$_j$]

(57) If an elided verb phrase is detected
And if the ellipsis clause is flanked by *and . . . too*
Then corefer the elided verb phrase with the verb phrase in the preceding conjunct or clause, using sloppy interpretation of anaphors.
Else, go to algorithm *x*.

Algorithm *x* will be the control that can access algorithms that represent any number of heuristics-based algorithms that should cover many of the typical cases and also launch Ontological Semantic resolution procedures if the former should fail.

Many lines of work must be carried out in order to convert the massive literature on verbal ellipsis (largely English-based) into a cross-linguistically applicable system of parameters and values. For example, English-specific lexical effects must be targeted and tested to see if there exist similar correspondences in other languages. One such case involves English *to*, which—although an ellipsis licensor—imposes restrictions on licit configurations. For example, *to* cannot license ellipsis when the infinitive would function as a subject ([58a] vs. [58b]) or when the infinitive is embedded in a noun phrase (59):

(58) a. You shouldn't play with rifles because it's dangerous to **∅**
b. *You shouldn't play with rifles because to **∅** is dangerous. (Lobeck 1995: 165)

(59) *? I reviewed Joe's attempt to find Holly while you reviewed José's attempt to (Johnson 2001, ex. 24b).

Since few languages have an infinitival particle like *to*, the many elliptical issues it raises might be lexically idiosyncratic and not central to a generalized theory of ellipsis. However, they might suggest more generalized parameters and values that affect ellipsis licensing for other licensors in other languages, like "the licensing category functions as a subject" or "the licensing category is within a noun phrase." In other words, the minimal pairs, affecting factors, and theoretical insights gained from work on English can be the ground upon which an extended set of parameters and values can be developed. Such parameterization will permit solutions to the difficult problems of interpretation and processing to be applied cross-linguistically.

10

Verbal Ellipsis with a Combination of Licensors

All of the well-studied types of verbal ellipsis discussed in chapter 9 are licensed by one of two basic strategies: interclause parallelism (Gapping and Stripping) or a lexical licensor (Sluicing and VP Ellipsis). There is, however, another strategy for licensing verbal ellipsis that has not made it to the agenda of mainstream theories or practical systems: licensing by a combination of lexical categories, as illustrated by Russian examples (1) and (2).[1]

(1) «Прощай, король». — «**Куда ты ∅**?» — «Пойду к соседнему королю» (Шварц 1: 69).
«Proščaj, korol'». — «**Kuda ty ∅**?» — «Pojdu k sosednemu korolju» (Švarc 1: 69).
good-bye King **where$_{\text{DIRECTIONAL}}$ you$_{\text{NOM}}$ ∅** will-go$_{\text{1SG}}$ to neighboring king$_{\text{DAT}}$
'"Good-bye, King." "Where are you going?" "I'm going to see the neighboring king."'

(2) [Discussing what seems like an unfair decision]
«Ведь **он не из мести ∅**, вы понимаете, что не из мести?» (Вампилов 1: 26).
«Ved' **on ne iz mesti ∅**, vy ponimaete, čto ne iz mesti?» (Vampilov 1: 26).
after-all **he$_{\text{NOM}}$ NEG from revenge$_{\text{GEN}}$ ∅** you$_{\text{NOM}}$ understand that NEG from revenge$_{\text{GEN}}$
'"After all, he's not doing this for revenge—you *do* understand that he's not out for revenge?"'

In (1) the combination of directional *where* and NOM-case *you* implies motion-related action, while in (2) the combination of NOM-case *he* and the adverbial *not for revenge*

implies behaving in some manner that must be understood from the context. So it is the combined semantics of the overt categories that both licenses the ellipsis and, sometimes with the help of the context, ensures recoverability of verbal meaning. I will refer to this as Multilicensor Verbal Ellipsis, or Multi-VE for short.

One salient aspect of Multi-VE is that often, but not always, a whole semantic class of verbs is implied rather than one specific meaning associated with one specific verb. The notion "semantic class" requires further specification. For example, when a verb of motion is elided, the motion might be on foot or in a vehicle, fast or slow; when a verb of speaking is elided, the speech might be storytelling, asking, lecturing, or blathering on; and when a verb of hitting is elided, the hitting might be punching, smacking, or walloping with a frying pan.

Multi-VE, which is highly productive in Russian and is also used, although more limitedly, in Polish and Czech, is almost exclusively found in colloquial speech. Because of this register restriction and because this is not among the simpler types of ellipsis to process, Multi-VE is probably not a priority for most current NLP systems. However, it is important to the current study for two reasons. First, it must be included in a full description of ellipsis in those languages that use it, whether or not one chooses to include rules for it in a given NLP system. Second, this is a realm where Ontological Semantic text processing can flex its muscles, so it serves as a good example of the goal toward which semantics-rich systems are moving.

1. Orienting the description toward processing

Unlike most of the types of ellipsis described so far, for which one could posit a relatively neutral description that could be applied to many fields, this subclass is best described with a class of applications in mind (like NLP), since the form of the description could vary significantly, being more oriented toward semantics, syntax, patterns with variables, and so on.[2] Each of the Subsections 1.1–1.4 discusses a type of preparatory descriptive work that would be required for taking the approach to processing these sentences that is put forward in section 2.

1.1. Compile a list of syntactic structures

Analysis of extensive examples of Multi-VE shows that, in most clauses that employ it, the overt syntactic categories fall into a number of patterns. The most common of these patterns—whose elements can occur in any order in Russian—are shown in table 10.1 (when other patterns are used, they are limited lexically and should be covered by special rules). When turned into syntactic rules, this list will constitute an inventory of "legal" sentence structures that, when encountered, will trigger the search for the elided verb based on the semantics of the overt categories.

The licensors in table 10.1 are intentionally not called minimal licensors, which is a matter of more theoretical than practical concern. For example, whereas (1) and (2) each have two licensors, in examples like (3) it is more difficult to draw the line between licensors and facultative categories:

TABLE 10.1 Syntactic structures in which Multi-VE is commonly used

Pattern	*Agent or experiencer*	*Tensed auxiliary or impersonal predicate word*	*PP or Adv*	*Object*
1	$\text{Agent}_{\text{NOM}}$	Tensed aux.	PP or Adv	—
2	$\text{Agent}_{\text{NOM}}$	—	PP or Adv	—
3	—	Tensed aux.	PP or Adv	—
4	—	—	PP or Adv	Object
5	$\text{Agent}_{\text{NOM}}$	—	—	Object
6	$\text{Experiencer}_{\text{DAT}}$	Impers. pred. word	PP or Adv	—
7	$\text{Experiencer}_{\text{DAT}}$	—	PP or Adv	—
8	—	Impers. pred. word	PP or Adv	—
9	—	Impers. pred. word	—	Object
10	—	*Davaj*(*te*) 'Let's'	PP or Adv	—
11	—	*Davaj*(*te*) 'Let's'	—	Object

(3) a. «Беги за билетами, если ещё **хочешь** ∅ со мной **в кино**» (Вампилов 2: 124).
«Begi za biletami, esli eščë **xočeš'** ∅ so mnoj **v kino**» (Vampilov 2: 124).
$\text{run}_{\text{IMPER}}$ for tickets if still $\textbf{want}_{\textbf{2SG}}$ ∅ with me_{INSTR} $\textbf{to movies}_{\textbf{ACC}}$
'"Run and get tickets if you still want to go to the movies with me."'

b. «Беги за билетами, если **ты** ещё **хочешь** ∅ со мной **в кино**».
«Begi za biletami, esli **ty** eščë **xočeš'** ∅ so mnoj **v kino**».
$\text{run}_{\text{IMPER}}$ for tickets if $\textbf{you}_{\textbf{NOM}}$ still $\textbf{want}_{\textbf{2SG}}$ ∅ with me $\textbf{to movies}_{\textbf{ACC}}$
'"Run and get tickets if you still want to go to the movies with me."'

In (3a), the ellipsis is licensed and recovered based on the combination of $\textit{want}_{\textit{2.SG}}$ and *to the movies*; *with me* is uninvolved in licensing. In (3b) there is one extra category that may or may not play a role in licensing the ellipsis: the subject, $\textit{you}_{\textit{NOM}}$. Obviously, *you* cannot be considered a *minimal* licensor in (3b) because (3a), which lacks it, is grammatical. However, it is possible that *you* somehow facilitates the ellipsis without being minimally required to license it. Thus, categories within elliptical sentences can have a number of statuses: they can clearly be minimal licensors, perhaps be minimal licensors, not be minimal licensors but potentially facilitate the ellipsis, or not be involved in licensing/facilitating the ellipsis at all. For practical purposes, any category that potentially facilitates the ellipsis and is typically a part of the given elliptical pattern will be listed among the licensors. (There is no practical benefit to reducing the number of patterns.)

Since Russian permits ellipsis in general so widely, it is common for a given category to imply the existence of another, which can thereby be elided. This was discussed in chapter 1 with respect to direct and reverse valency. Some examples:

- Pattern 3 lacks but implies a subject, since the inflectional ending on the tensed auxiliary licenses its ellipsis and provides clues to the nature (person, number, and sometimes gender) of its referent.

- Pattern 7 lacks but implies an impersonal predicate word that conveys modality, since the DAT case marking on the experiencer must be assigned by such a category (the modality must be contextually determined).
- Patterns 8 and 9 lack but imply an experiencer, since the modality of impersonal predicate words must apply to some specific or generalized persons/animals.

1.2. Prepare the system to diagnose each type of category

In order to recognize the categories used in describing syntactic patterns, the system must have access to diagnostics for or inventories of those categories. To cover the categories listed in table 10.1, the following types of information are required, many of which are basic requirements for most NLP systems anyway:

- analysis of the function of case marking (e.g., NOM is for subjects; ACC is for direct objects);
- knowledge that agents and experiencers must be sentient (humans or animals);
- access to part-of-speech information, like preposition, adverb, noun;
- an inventory of auxiliaries that can inflect; for Russian this includes *moč'* 'can', *xotet'* 'want', *uspet'* 'have time', *dumat'* be considering, and others; and
- an inventory of impersonal predicate words (most in Russian are fixed in form): *možno* 'it is possible', *nužno* 'it is necessary', *nado* 'it is necessary', *ne nado* '(one) shouldn't/does not have to', *vozmožno* 'it is possible', *nevozmožno* 'it is impossible', *nel'zja* 'it is impossible/(one) mustn't', *xočetsja/xotelos'* '(one) feels/felt like', *zaxočetsja/zaxotelos'* '(one) will feel like/(one) felt like'.

1.3. Compile a list of common collocations

Although most instances of Multi-VE in Russian can be analyzed productively (i.e., these are not idiomatic constructions), for practical purposes it would be helpful to compile a list of the most common collocations and explicitly encode their semantic represention. This promises the same benefits for NLP as having a large phrasal lexicon. A short inventory of such collocations in Russian is provided here, with examples following:

AGENT$_{NOM}$ *nečajanno* 'accidentally' → AGENT did it by mistake
AGENT$_{NOM}$ *sam/sama/sami* 'self' → AGENT will do it 'oneself' (select correct reflexive pronoun based on the referent for AGENT)
AGENT$_{NOM}$ *ne* 'not' *mog/mogla/mogli* 'could' *inače* 'otherwise' → AGENT couldn't do otherwise
AGENT$_{NOM}$ *ničego* 'nothing' *ne* 'not' *možet* 'can' → AGENT can't do anything
Davaj(*te*) 'Let's' *po porjadku* 'in order' → Let's take this step-by-step
Davaj(*te*) 'Let's' *po novoj* 'in new-way' → Let's start again
Tak 'in this way' *nel'zja* 'it is impossible/(one) must not' → This isn't right, (One) shouldn't do this
Čego/Čto 'what' *èto* 'particle' *ty/vy* 'you$_{NOM}$' → What did you do that for?

(4) [Vlasakiev's neighbor stained Vlasakiev's doormat while painting his own front door]
«Может, **он ∅ нечаянно**?» — предложила Лидуся. «Может, нечаянно, а может, и нарочно!» — с жаром откликнулся Власакиев (Шинов: 55).
«Možet, on ∅ nečajanno?» — predložila Lidusja. «Možet, nečajanno, a možet, i naročno!» (Šinov: 55).
maybe **he$_{NOM}$ ∅ accidentally** suggested Lidusja$_{NOM}$ maybe accidentally but maybe PARTICLE intentionally with ardor responded Vlasakiev$_{NOM}$
'"Maybe he did it by mistake?" suggested Lidusja. "Maybe by mistake, but maybe on purpose!" retorted Vlasakiev heatedly.'

(5) «Уже уходишь? — грустно спросил Вадим. — Хочешь, я тебя провожу?» — «Нет, нет, **я ∅ сама**» (Войнович 1: 56).
«Uže uxodiš'? — grustno sprosil Vadim. — Xočeš', ja tebja provožu?» — «Net, net, **ja ∅ sama**» (Vojnovič 1: 56).
already leave$_{2SG}$ sadly asked Vadim want$_{2SG}$ I$_{NOM}$ you$_{ACC}$ accompany$_{1SG}$ no no **I$_{NOM}$ ∅ myself$_{NOM}$**
'"Leaving so soon?" asked Vadim sadly. "Do you want me to walk you home?" "No, I'll go myself."'

(6) «Я всё поняла. Ты сделал это не ради удовольствия, поняла. **Ты не мог иначе ∅**, поняла . . . » (Вампилов 1: 66).
«Ja vsë ponjala. Ty sdelal èto ne radi udovol'stvija, ponjala. **Ty ne mog inače ∅**, ponjala . . . » (Vampilov 1: 66).
I$_{NOM}$ everything understood$_{1SG}$ you$_{NOM}$ did that NEG for pleasure understood$_{1SG}$ **you$_{NOM}$ NEG could otherwise ∅** understood$_{1SG}$
'"I understand everything. You didn't do this for pleasure, I understand. You couldn't have done otherwise, I understand . . ."'

(7) Странный человек этот Вадим. **Он** ни к чему не приспособлен, **ничего не может ∅** (Войнович 1: 55).
Strannyj čelovek ètot Vadim. **On** ni k čemu ne prisposoblen, **ničego ne možet ∅** (Vampilov 1: 55).
strange person this Vadim **he$_{NOM}$** NEG to anything NEG fit **nothing NEG can$_{3SG}$ ∅**
'This Vadim was a strange guy. He wasn't fit for anything, he couldn't do anything.'

(8) [The speaker is trying to understand the situation]
«Так. **Давайте ∅ по порядку**. Выходит, ты — его отец, а он — твой сын. Так, что ли?» (Вампилов 2: 95).
«Tak. **Davajte ∅ po porjadku.** Vyxodit, ty – ego otec, a on – tvoj syn. Tak, čto li?» (Vampilov 2: 95).
so **let's ∅ in order** turns-out you$_{NOM}$ his father$_{NOM}$ and he$_{NOM}$ your son$_{NOM}$ that-way COMPOUND PARTICLE
'"OK. Let's take this step-by-step. You're his father and he's your son. Right?"'

(9) «Снова тост? Нет, **так нельзя ∅**, только выпили и снова. Дайте закусить» (Вампилов 1: 8).

«Snova tost? Net, **tak nel'zja ∅**, tol'ko vypili i snova. Dajte zakusit'» (Vampilov 1: 8).
again toast no **this-way impossible$_{\text{IMPERS}}$ ∅** just drank$_{\text{1PL}}$ and again permit$_{\text{IMPER}}$ eat$_{\text{INFIN}}$
'"Another toast? No, this isn't the way to do things: we just had a drink and now another. Let us have a bite to eat first."'

(10) «А я тебе подарок приготовил. — Мишка протянул подарок. — Померь». . . — «**Чего это ты ∅**?» (Токарева: 503).
«A ja tebe podarok prigotovil. — Miška protjanul podarok. — Pomer'» . . . — «**Čego èto ty ∅**?» (Tokareva: 503).
and I$_{\text{NOM}}$ you$_{\text{DAT}}$ gift$_{\text{ACC}}$ prepared$_{\text{1SG}}$ Mishka held-out gift$_{\text{ACC}}$ try-on$_{\text{IMPER}}$ **what-for this$_{\text{ACC}}$ you$_{\text{NOM}}$ ∅**
'"I've got a gift for you." Mishka held out the gift. "Try it on." . . . "What'd you do that for?"'

(11) «Слушай, а как тебя зовут? Извини, там, в кафе, я толком не расслышал». — «Я тоже не расслышал». — **Давай ∅ по новой**, что ли . . . » (*Трясут друг друга за руку*.) «Бусыгин. Владимир». — «Севостьянов. Семён» (Вампилов 2: 76).
«Slušaj, a kak tebja zovut? Izvini, tam, v kafe, ja tolkom ne rasslyšal». — «Ja tože ne rasslyšal». — «**Davaj ∅ po novoj**, čto li . . . » (*Trjasut drug druga za ruku*.) «Busygin. Vladimir». — «Sevost'janov. Semën» (Vampilov 2: 76).
listen$_{\text{IMPER}}$ and how you$_{\text{ACC}}$ are-called$_{\text{3PL}}$ sorry there in café I$_{\text{NOM}}$ clearly NEG hear I$_{\text{NOM}}$ too NEG heard **let's ∅ PREPOSITION new$_{\text{ADJ. DAT}}$** COMPOUND PARTICLE
'"What was your name again? Sorry, but there, in the café, I didn't really catch it." "I didn't catch yours, either." "Let's start again . . ." (*They shake hands*.) "Busygin. Vladimir." "Sevost'yanov. Semyon."'

1.4. Describe special subtypes

Of the semantic classes of verbs that can be elided using this licensing strategy in Russian, verbs that express motion, speaking, and hitting (illustrated by [12]–[14]) are privileged in Russian, at least in terms of frequency of use:

(12) Я ∅ в кино.
Ja ∅ v kino.
I$_{\text{NOM}}$ ∅ to movies$_{\text{ACC}}$
'I'm going to [or: I'm off to, I'm heading to] the movies.'

(13) О чём он ∅?
O čëm on ∅?
about what$_{\text{PREP}}$ he$_{\text{NOM}}$ ∅
'What is he talking [or: asking, yelling, and so forth] about?'

(14) Я ему ∅ не сильно (Земская 1973: 306).
Ja emu ∅ ne sil'no (Zemskaja 1973: 306).

I_{NOM} him_{DAT} ∅ NEG hard
'I didn't hit [or: punch, smack, and so forth] him hard.'

That is, when a human subject combines with a destination like *to the movies*, the most typical action is motion related, as in (12); when a human subject combines with a phrase like *about what*, the most typical action is speech related, as in (13); when a human subject combines with a human DAT object and an adverb like *hard*, the most typical action is hitting, as in (14).

Zolotova (1982: 199) suggests that elided verbs of motion, speaking, and hitting have yet another special property—the nuances of speed and immediacy and the feeling that the action is occurring in the present tense, as in (15):

(15) Стали мы болтать о том, о сём: вдруг смотрю, **Казбич** вздрогнул, переменился в лице — и ∅ **к окну** . . . «Что с тобой?» — спросил я (Лермонтов, from Zolotova 1982: 199).
Stali my boltat' o tom, o sëm: vdrug smotrju, **Kazbič** vzdrognul, peremenilsja v lice — i ∅ **k oknu** . . . «Čto s toboj?» — sprosil ja (Lermontov).
began we_{NOM} $chat_{INFIN}$ about this about that suddenly $look_{1SG}$ $Kazbich_{NOM}$ shuddered changed in face and ∅ **to window** what with you_{INSTR} asked I_{NOM}
'We began chatting about this and that. Suddenly I look: Kazbich shuddered, the expression on his face changed, and he raced to the window . . . "What's the matter?" I asked.'

Each of the following subsections is devoted to a different semantic class of verbs. Possible combinations of licensors are listed, each being described semantically and syntactically, and examples are provided. The reason for correlating the syntactic realizations and semantics of the overt categories is to delimit the number of candidate analyses.

Elided verbs of motion

Semantically speaking, eliding verbs of motion requires overtly specifying the person or thing moving (which is realized as NP_{NOM} unless otherwise specified) and the destination, source, goal, purpose, or (in limited instances) time frame of the motion. I will refer to the person/thing moving as the MOVER, avoiding the complexities associated with whether people/things that move have the thematic role of 'theme', 'agent', and so on. The inventory of means by which these latter can be realized—which will be referred to hereafter as MOTION-GOAL (with "goal" covering a broader than typical sense for brevity of the notation)—is shown in table 10.2.

Motion 1: MOVER + MOTION-GOAL

(16) **Ефим** встаёт и . . . шлёпает в коридор, подбирает газету и с газетой ∅ — **в уборную** (Войнович 2: 19–20).
Efim vstaët i . . . šlëpaet v koridor, podbiraet gazetu i s gazetoj ∅ — **v ubornuju** (Vojnovič 2: 19–20).

TABLE 10.2 Meanings and realizations of MOTION-GOAL licensors

Syntactic category	*Semantic class*
PP = *v/na* 'to' + NP_{ACC}	Place (destination)
PP = *iz/s* 'from' + NP_{GEN}	Place (source)
PP = *k* 'to' + NP_{DAT}	Person (destination)
PP = *ot* 'from' + NP_{GEN}	Person (source)
PP = *za* 'for' + NP_{INSTR}	Thing sought
PP or Adverb	Motion related or time related
Infinitive	Any action, but particularly those connected with a certain location (information that would be stored in ontological scripts)

Yefim$_{NOM}$ gets-up and tramps into hall picks-up newspaper and with newspaper **Ø into bathroom**$_{ACC}$
'Yefim gets up and . . . tramps into the hall; he picks up the newspaper and heads into the bathroom with it.'

(17) «**Мы Ø** сейчас **с поезда**» (Вампилов 2: 88).
«**My Ø** sejčas **s poezda**» (Vampilov 2: 88).
we$_{NOM}$ **Ø** now **from train**$_{GEN}$
'"We're on our way from the train."'

(18) «Проходите, присаживайтесь . . . » — «Я собственно . . . **Я Ø к Владимиру Алексеевичу**» (Вампилов 1: 30).
«Proxodite, prisaživajtes' . . . – Ja sobstvenno . . . **Ja Ø k Vladimiru Alekseeviču**» (Vampilov 1: 30).
come-in$_{IMPER}$ sit-down$_{IMPER}$ I_{NOM} actually **I**$_{NOM}$ **Ø to Vladimir**$_{DAT}$ **Alekseyevich**$_{DAT}$
'"Come in, sit-down. . . ." "Actually I . . . I've come to see Vladimir Alekseyevich."'

(19) «А я тебя не заметил, — сказал Ефим виновато. — **Ты** тоже **Ø за шапкой**?» (Войнович 2: 48).
«A ja tebja ne zametil, — skazal Efim vinovato. — **Ty** tože **Ø** za šapkoj?» (Vojnovič 2: 48).
but I_{NOM} you$_{ACC}$ NEG noticed$_{MASC.SG}$ **you**$_{NOM}$ also **Ø for hat**$_{INSTR}$
'"Oh, I didn't see you," said Yefim in a guilty tone. "Have you come for a hat, too?"'

(20) «**Куда Ø он, туда Ø и я.** Мы с ним неразлучные» (Вампилов 2: 105).
«**Kuda Ø on, tuda Ø i ja**. My s nim nerazlučnye» (Vampilov 2: 105).
where$_{DIRECTIONAL}$**Ø he**$_{NOM}$ **there**$_{DIRECTIONAL}$ **Ø also I**$_{NOM}$ we$_{NOM}$ with him$_{INSTR}$ inseparable
'Wherever he goes, I go. We're inseparable.'

(21) «Кофе будешь?» — «Нет, нет, **я Ø на минутку**» (Войнович 2: 79).
«Kofe budeš'?» — «Net, net, **ja Ø na minutku**» (Vojnovič 2: 79).
coffee$_{ACC}$ will$_{2SG}$ no no **I**$_{NOM}$ **Ø for minute**
'"Will you have some coffee?" "No, I've just stopped by for a minute."'

(22) «**Я – обедать**. Вернусь через час».
«**Ja – obedat'**. Vernus' čerez čas».
I$_{NOM}$ Ø eat$_{INFIN}$ will-return$_{1SG}$ in hour
'"I'm going to lunch. I'll be back in an hour."'

Motion 2: NP$_{NOM}$ + Tensed auxiliary + MOTION-GOAL

(23) «На твоём месте я бы сначала доучился**. В тайгу ты** всегда **успеешь Ø**» (Вампилов 2: 106).
«Na tvoëm meste ja by snačala doučilsja**. V tajgu ty** vsegda **uspeeš' Ø**» (Vampilov 2: 106).
in your place I$_{NOM}$ CONDIT first finished-studying **to taiga$_{ACC}$ you$_{NOM}$** always **will-have-time Ø**
'"If I were you I'd finish school first. You have plenty of time to go to the taiga."'

Motion 3: Tensed auxiliary + MOTION-GOAL

Here the tensed auxiliary provides morphological information for the recovery of the person or thing undergoing the motion:

(24) «Надолго к нам?» — «Нет, завтра же **думаю Ø в Москву**» (Чехов 2: 434–435).
«Nadolgo k nam?» — «Net, zavtra že **dumaju Ø v Moskvu**» (Čexov 2: 434–435).
for-long to us$_{DAT}$ no tomorrow PARTICLE **think$_{1SG}$ Ø to Moscow$_{ACC}$**
'"Are you here for long?" "No, I think I'll head off to Moscow tomorrow."'

Motion 4: NP$_{DAT}$ (= logical subject) + Impersonal predicate word + MOTION-GOAL (*Logical subject* is a term commonly used to refer to DAT-case entities that have a subject-like function in the sentence.)

(25) «Хорошо ещё, что сегодня **ему не надо Ø в школу**» (Вампилов 2: 106).
«Xorošo eščë, čto segodnja **emu ne nado Ø v školu**» (Vampilov 2: 106).
good still that today **him$_{DAT}$ NEG necessary$_{IMPERS}$ Ø to school$_{ACC}$**
'"And it's good that he doesn't have to go to school today."'

(26) «Простите, но **мне хочется Ø на свежий воздух**. У меня разболелась голова» (Пелевин: 160).
«Prostite, no mne **xočetsja Ø na svežij vozdux**. U menja razbolelas' golova» (Pelevin: 160).
excuse$_{IMPER}$ me but **me$_{DAT}$ wants$_{3SG}$ Ø to fresh$_{ACC}$ air$_{ACC}$** at me got-achy head$_{NOM}$
'"Excuse me, but I need some air. I've gotten a headache."'

Motion 5: Impersonal predicate word + MOTION-GOAL

Here the logical subject, whose existence is implied by the impersonal predicate word, is elided and must itself be recovered (cf. chapter 13, section 3.3):

(27) «Кофе выпит, **можно Ø на покой**» (Чехов 1: 566).
«Kofe vypit, **možno Ø na pokoj**» (Čexov 1: 566).

coffee$_{ACC}$ drunk$_{PAST\text{-}PARTICIPLE}$ **possible$_{IMPERS}$ Ø to peace$_{ACC}$**
‘“The coffee’s drunk, now we can retire.”’

Motion 6: NP$_{DAT}$ (= logical subject) + MOTION-GOAL

In these structures, an impersonal predicate word has been elided, evidenced by two things: first, the sentences imply the type of modality that is typical of impersonal predicate words (*must, should,* etc.); and second, the DAT case marking on the understood subject must be assigned somehow:

(28) Потом он отправил Кукушу спать (**ей** утром опять Ø **на работу**), а сам перетаскал на кухню и там долго мыл посуду. . . (Войнович 2: 57).
Potom on otpravil Kukušu spat’ (**ej** utrom opjat’ Ø **na rabotu**), a sam peretaskal na kuxnju i tam dolgo myl posudu . . . (Vojnovič 2: 57).
then he$_{NOM}$ sent Kukusha$_{ACC}$ sleep$_{INFIN}$ **her$_{DAT}$** morning$_{INSTR}$ again **Ø to work$_{ACC}$**
and self$_{NOM}$ trudged to kitchen and there for-a-long-time washed dishes$_{ACC}$
‘Then he sent Kukusha to bed (she had to go to work again in the morning), and he trudged off to the kitchen where he spent a long time washing dishes . . .’

(29) «А я знаю дорогу в деревню». — «А зачем **нам в деревню**?» — «За молоком» . . . (Золотые: 62).
«A ja znaju dorogu v derevnju». — «A začem **nam v derevnju**?» — «Za molokom» . . . (Zolotye: 62).
but I$_{NOM}$ know road$_{ACC}$ to village but why **us$_{DAT}$ to village$_{ACC}$** for milk$_{INSTR}$
‘“I know how to get to the village.” “Why should we go to the village?” “For milk” . . .’

Motion 7: MOTION-GOAL + other category

This class is a repository for patterns in which constraints on the nature of the “other category” are as yet unclear, as in (30):

(30) «Какого же мы дурака сваляли, что Хабра не прихватили!» — «Ясно, дураки. Теперь **без него никуда** Ø!» (Хмелевская 1: 201–202).
«Kakogo že my duraka svaljali, čto Xabra ne prixvatili!» — «Jasno, duraki. Teper’ **bez nego nikuda** Ø!» (Chmielewska 1: 201–202).
[MULTIWORD IDIOM: what idiots we are] that Khabr$_{ACC}$ NEG took-along obviously idiots now **without him$_{GEN}$ nowhere$_{DIRECTIONAL}$ Ø**
‘“How stupid of us not to take Khabr [a dog] along!” “Yeah, that was really dumb. From now on we don’t go anywhere without him!”’

Elided verbs of speaking

Eliding verbs of speech requires overtly specifying the SPEAKER, realized as NP$_{NOM}$, and one or more of the following: the SPEECH-CONTENT, whose realizations are shown in table 10.3; the ADDRESSEE, realized as NP$_{DAT}$; and an adverb that expresses the manner in which or reason that something is said (ADVERB$_{MANNER/REASON}$).

TABLE 10.3 Meanings and realizations of SPEECH-CONTENT licensors

Syntactic category	*Semantic class*
PP = *o* 'about' + NP_{PREP}	The topic being spoken about
PP = *pro* 'about' + NP_{ACC}	The topic being spoken about
Clause as direct quote	What is said

Speech 1: SPEAKER + SPEECH-CONTENT

(31) «Интересно, что этот кретин выкинул?» — говорил Павлик. «**Ты о ком Ø**?» — «Да о её сыночке» (Хмелевская 1: 239).
«Interesno, čto ètot kretin vykinul?» — govoril Pavlik. «**Ty o kom Ø**?» — «Da o eë synočke» (Chmielewska 1: 239).
interesting what that $creep_{NOM}$ did said $Pavlik_{NOM}$ **you_{NOM} about whom Ø** PARTICLE about her son
'"I wonder what this creep has got up to," said Pavlik. "Who are you talking about?" "Her son."'

(32) «Вот я всё хочу тебя спросить . . . Зачем ты это делаешь?» — «**Вы Ø про палисадник**?.. Зачем я его чиню?» (Вампилов 4: 333).
«Vot ja vsë xoču tebja sprosit' . . . Začem ty èto delaeš'?» — «**Vy Ø pro palisadnik**?.. Začem ja ego činju?» (Vampilov 4: 333).
so I_{NOM} keep wanting you_{ACC} ask_{INFIN} why you_{NOM} $that_{ACC}$ do_{2SG} **you_{NOM} Ø about $fence_{ACC}$** why I_{NOM} it_{ACC} fix_{1SG}
'"I keep meaning to ask you . . . Why are you doing that?" "You mean the fence? Why am I fixing it?"'

Speech 2: SPEAKER + ADDRESSEE

(33) «Молодой человек!» — позвала почтальонша Зинаида. Лётчик не обернулся. «Мужчина!» — поправила себя Зинаида. «**Вы Ø мне**?» — «Вам, а кому же ещё . . . » — сказала Фрося (Токарева: 475).
«Molodoj čelovek!» — pozvala počtal'onša Zinaida. Lëtčik ne obernulsja. «Mužčina!» — popravila sebja Zinaida. «**Vy Ø mne**?» — «Vam, a komu že eščë . . . » — skazala Frosja (Tokareva: 475).
young man_{NOM} called $postal\text{-}carrier_{NOM}$ $Zinaida_{NOM}$ $pilot_{NOM}$ NEG turned-around $mister_{NOM}$ corrected herself $Zinaida_{NOM}$ **you_{NOM} Ø me_{DAT}** you_{DAT} and $whom_{DAT}$ PARTICLE else said $Frosya_{NOM}$
'"Young man!" called the postal carrier, Zinaida. The pilot didn't turn around. "Mister!" Zinaida corrected herself. "**Are you talking to me**?" "Of course, who else would she be talking to . . ." said Frosya.'

Speech 3: SPEAKER + ADDRESSEE + SPEECH-CONTENT

(34) . . . Жилец в это время домой приходил. «Здравствуйте!» — говорит. **Я ∅ ему: «Здравствуйте!»** (Достоевский 1: 141).
. . . Žilec v èto vremja domoj prixodil. «Zdravstvujte!» — govorit. **Ja ∅ emu: «Zdravstvujte!»** (Dostoevskij 1: 141).
tenant$_{NOM}$ at that time home$_{DIRECTIONAL}$ came hello says$_{3SG}$ **I$_{NOM}$ ∅ him$_{DAT}$ hello**
'. . . The tenant was just arriving home. "Hello!" he says. "Hello!" I say back.'

(35) «Значит, так: **ты ∅ ему „здрасте"**, **он ∅ тебе „здрасте"**, — напомнила Вероника (Токарева: 476).
«Značit, tak: **ty ∅ emu „zdraste", on ∅ tebe „zdraste",** — napomnila Veronika (Tokareva: 476).
so this-way **you$_{NOM}$ ∅ him$_{DAT}$ hi he$_{NOM}$ ∅ you$_{DAT}$ hi** reminded Veronika$_{NOM}$
'"So, here's what you do: you go to him 'hi' and he goes to you 'hi'" Veronika reminded her.'[3]

Speech 4: SPEAKER + ADVERB$_{MANNER/REASON}$

(36) «Ты его не знаешь. Шутит **он** или **∅ серьёзно** — сразу у него не поймёшь . . . » (Вампилов 4: 358).
«Ty ego ne znaeš'. Šutit **on** ili **∅ ser'ëzno** — srazu u nego ne pojmëš' . . . » (Vampilov 4: 358).
you$_{NOM}$ him$_{ACC}$ know$_{2SG}$ jokes$_{3SG}$ **he$_{NOM}$** or **∅ seriously** immediately at him NEG understand$_{2SG}$
'"You don't know him. Whether he's joking or being serious—you just can't tell right away with him . . ."'

(37) «Мне говорили, будто она повела какую-то особенную жизнь. В чём дело?» — «Это, доктор, длинная история». — «А **вы ∅ покороче**» (Чехов 2: 432).
«Mne govorili, budto ona povela kakuju-to osobennuju žizn'. V čëm delo?» — «Èto, doktor, dlinnaja istorija». — «A **vy ∅ pokoroče**» (Čexov 2: 432).
me$_{DAT}$ told$_{3PL}$ as-if she$_{NOM}$ led some-kind-of strange life$_{ACC}$ in what matter that doctor long story but **you$_{NOM}$ ∅ more-briefly**
'"I was told that she lead some kind of peculiar lifestyle. What's that all about?" "That, Doctor, is a long story." "Well, **give me the short version**."'

Elided verbs of hitting

All licensing strategies for eliding verbs of hitting require three licensors. In all cases, the HITTER (realized as NP$_{NOM}$) and the person or animal hit (called PATIENT and realized as NP$_{ACC}$) must be overt. The third licensor, however, can take various forms. It can be:

- an adverbial that refers to the manner or reason of hitting (called ADVERB$_{MANNER/REASON}$ and realized by a PP or adverb);

- an adjunct that refers to the INSTRUMENT of hitting (realized as NP_{INSTR});
- an adjunct that refers to the place hit (called SITE and realized by various types of prepositional phrases: *po* + NP_{DAT}, *v* + NP_{ACC}); or
- an idiosyncratic combination of restrictions on the nature of the HITTER—first-person singular pronoun only—plus emphatic intonation on that pronoun.[4]

Hitting 1: HITTER + PATIENT + SITE

(38) **Я** от неё отворачиваюсь / **она ∅ меня** опять **по лицу** (Земская 1973: 306).
Ja ot neë otvoračivajus' / **ona ∅ menja** opjat' **po licu** (Zemskaja 1973: 306).
I_{NOM} from her_{GEN} turn **she_{NOM} ∅ me_{ACC}** again **in $face_{DAT}$**
'I turn away from her / she smacks me in the face again.'

Hitting 2: HITTER + PATIENT + $ADVERB_{MANNER/REASON}$

(39) А **здорово они его ∅** (Земская 1973: 306).
A **zdorovo oni ego ∅** (Zemskaja 1973: 306).
but **intensely $they_{NOM}$ him_{ACC} ∅**
'And they really let him have it.'

(40) Откуда я знаю **за что они его ∅**? (Земская 1973: 306).
Otkuda ja znaju **za čto oni ego ∅**? (Zemskaja 1973: 306).
from-where I_{NOM} $know_{1SG}$ **for $what_{ACC}$ $they_{NOM}$ him_{ACC} ∅**
'How am I supposed to know why they beat him up?'

Hitting 3: HITTER + PATIENT + SITE + INSTRUMENT

(41) **Они** заманили **её** в подъезд и **чем-то там тяжёлым ∅ по голове** / но она всё-таки жива осталась (Земская 1973: 306).
Oni zamanili **ë** v pod"ezd i **čem-to tam tjažëlym ∅ po golove** / no ona vsë-taki živa ostalas' (Zemskaja 1973: 306).
$they_{NOM}$ lured her_{ACC} into doorway and **$something_{INSTR}$ $heavy_{INSTR}$ ∅ in $head_{DAT}$**
but she_{NOM} nevertheless alive remained
'They lured her into the doorway and hit her on the head with something heavy, but she survived.'

(42) **Я его ∅ кулаком в бок** — пусть не лезет без очереди.
Ja ego ∅ kulakom v bok — pust' ne lezet bez očeredi.
I_{NOM} him_{ACC} ∅ $fist_{INSTR}$ in side let-it-be NEG creep-in_{3SG} without line
'I punched him in the side—that'll keep him from cutting in line.'

Hitting 4: $HITTER_{1.SG}$ + PATIENT + emphatic intonation on HITTER

(43) «Вдруг всё провалится, — подумал он, холодея, — да **я его ∅**!» (Тургенев quoted from Fominyx 1965: 113).
«Vdrug vsë provalitsja, — podumal on, xolodeja, — da **ja ego ∅**!» (Turgenev).

suddenly everything$_{NOM}$ goes-to-pieces thought he$_{NOM}$ getting-cold PARTICLE **I$_{NOM}$ him$_{ACC}$ ∅**
'"What if everything goes to pieces," he thought, feeling a chill up his spine, "then I'll really give it to him!"'

Although the semantic fields of motion, speaking, and hitting are the most widely used types that can be captured by patterns like these, there are others as well. For example, when the overt elements include (1) a finite verb form or impersonal predicate word that means "it is impossible," (2) a prepositional phrase headed by the preposition *bez* 'without', and (3) an indication of realm (PP) or people affected (NOM subject of finite verb form or DAT experiencer of impersonal predicate), the verbal meaning has the broad sense of "live, exist, go on, survive." The patterns are shown in (44)–(47):

(44) «Нам не нужны полицейские, раз у нас живёт Пеппи», — крикнул кто-то из толпы . . . «Нет, **без полицейских** всё же **нельзя ∅ в городе», —** возразила Пеппи (Линдгрен: 120).
«Nam ne nužny policejskie, raz u nas živët Peppi», — kriknul kto-to iz tolpy . . . «Net, **bez policejskix** vsë že **nel'zja v gorode**», — vozrazila Peppi (Lindgren: 120).
{ . . . } no **without policemen$_{GEN}$** after all (idiom) **impossible$_{IMPERS}$ ∅ in city** retorted Pippi$_{NOM}$
'"We don't need policemen now that we've got Pippi," shouted someone from the crowd. . . . "Not true, a city can't be without policemen," Pippi disagreed.'

(45) «Было у меня два ассистента, — сказал он, — вашего примерно возраста. Такие, знаете, ассенизаторы реальности. Сейчас **без этого в бизнесе нельзя ∅**» (Пелевин: 124).
«Bylo u menja dva assistenta, — skazal on, — vašego primerno vozrasta. Takie, znaete, assenizatory real'nosti. Sejčas **bez ètogo v biznese nel'zja ∅**» (Pelevin: 124).
was at me two assistants$_{NOM}$ said he$_{NOM}$ your$_{GEN}$ approximately age$_{GEN}$ such$_{NOM.PL}$ know$_{2PL}$ sewage-disposers$_{NOM}$ reality$_{GEN}$ now **without that$_{GEN}$ in business impossible$_{IMPERS}$ ∅**
'"I had two assistants," he said, "about your age. You know, sewage disposers of reality. These days you can't do without them in business."'

(46) «Только вот **без дела не могу ∅**, мамочка. Мне каждую минуту надо что-нибудь делать» (Чехов 1: 587).
«Tol'ko vot **bez dela ne mogu ∅**, mamočka. Mne každuju minutu nado čto-nibud' delat'» (Čexov 1: 587).
only PARTICLE **without occupation$_{GEN}$ NEG can$_{1SG}$ ∅** Mom me$_{DAT}$ every minute necessary$_{IMPERS}$ something$_{ACC}$ do$_{INFIN}$
'"But Mom, I just can't stand having nothing to do. I need to be doing something every minute."'

(47) «Сьюзен полагает, что власть в мужском мире — это основа. «**Без власти женщине невозможно ∅**. Не получив её, она никуда не пройдёт» (Богуславская: 173).
«S'juzen polagaet, čto vlast' v mužskom mire — èto osnova. **«Bez vlasti ženščine nevozmožno ∅**. Ne polučiv eë, ona nikuda ne projdët» (Boguslavskaja: 173).
Susan$_{NOM}$ figures that power$_{NOM}$ in man's world it-is basis **without power$_{GEN}$ woman$_{DAT}$ impossible$_{IMPERS}$ ∅** NEG receiving it$_{ACC}$ she$_{NOM}$ never NEG will-move-forward
'Susan thinks that power is key in a man's world. "Without power, a woman can't survive. Without it, she can't get anywhere."'

Once again, the benefit of describing patterns like these is that one can specify the relationship between the syntactic realization of a licensor and its semantics, whose correlation increases the likelihood of a correct analysis.

2. Processing algorithm

One could write many types of algorithms for processing Multi-VE sentences. The following one reflects a progression from more straightforward to more complex issues. Although the algorithm is written such that processing stops after a given analysis succeeds, this need not be the case—processing could continue, with the best of alternative analyses being selected later based on a comparison of confidence levels for each analysis.

1. *Is the clause* (*minus possible adjuncts*) *among the fixed collocations? If yes, use this analysis. If not, go to 2*. Sentences (4)–(11) would be analyzed this way.

2. *Does the combination of subcategorization and selectional restrictions limit the verb choice to one? If yes, use this analysis. If not, go to 3*. This analysis would succeed in examples like (48): the only verb that takes a complement with lexical INSTR case marking and regularly collocates with objects like languages and music is *zanimat'sja* 'to study, be engaged in'.

(48) **Бедные дети ∅ / и музыкой / и языком / и черте** [sic[5]] **чем** (Земская 1973: 305).
Bednye deti ∅ / i muzykoj / i jazykom / i čerte [sic] **čem** (Zemskaja 1973: 305).
poor kids$_{NOM}$ ∅ both music$_{INSTR}$ and language$_{INSTR}$ and God-knows what$_{INSTR}$
'Those poor kids: they do music and foreign language and God knows else.'[6]

3. *Would the preceding "major" event/action* (*in Ontological Semantic terms, the head of the most recent chunk of text-meaning representation*) *in the context fit the subcategorization and selectional restrictions of the overt elements in this clause? If yes, use that analysis. If not, go to 4*. This method accounts for quite a few examples, but heuristics must be developed to

distinguish major from nonmajor events. For example, all speech events (*he said, she exclaimed*) would be nonmajor and therefore should be skipped in backtracking the text for a potential antecedent. In addition, interpretation of the lexical items in the preceding context might be necessary: for example, in (55) *spilled blood* must be recognized as *killed*—in fact, it should instantiate an instance of the concept KILL. Then this concept, with its associated patient slot, will be an appropriate antecedent for the elided verb. The overt antecedents are underlined in this set of examples to distinguish them from the boldface licensors of the ellipsis.

(49) «Чё? Снова за ремонт? Ну, Валюша, подписалась ты с этим палисадником! . . . Ладно. Дай я его налажу». — «Не надо». — «Да **я его ∅ мигом**» (Вампилов 4: 366).
«Čë? Snova za remont? Nu, Valjuša, podpisalas' ty s ètim palisadnikom!.. Ladno. Daj ja ego nalažu». — «Ne nado». — «Da **ja ego ∅ migom**» (Vampilov 4: 366).
{ . . . } I_{NOM} it_{ACC} fix NEG $necessary_{IMPERS}$ PARTICLE **I_{NOM} it_{ACC} in-a-flash**
'"At it again? Valyusha, you're taking this garden way too seriously! . . . Oh, all right. Let me fix it for you." "No need." "I'll do it in a flash."'

(50) [The bus driver is holding one hand up and driving with the other while being hijacked by a couple of kids]
«Почему вы смеётесь?» — удивилась Лидуся. «**Это я ∅ над собой**! Я, наверное, ужасно смешно выгляжу с поднятой рукой!» (Шинов: 32).[7]
«Počemu vy smeëtes'?» — udivilas' Lidusja. «**Èto ja ∅ nad soboj**! Ja, navernoe, užasno smešno vygljažu s podnjatoj rukoj!» (Šinov: 32).
why you_{NOM} laugh was-surprised Lidusja PARTICLE **I_{NOM} ∅ at $self_{INSTR}$** I_{NOM} probably terribly funny look with raised hand
'"Why are you laughing?" Lidusja asked, surprised. "I'm laughing at myself! I must look awfully silly with my hand held up!"'

(51) «На площади выкопал орхидею. Цветок такой . . . Да разве это хулиганство?» — «Это неслыханная наглость. На площади, под носом у милиции. **Вы** что, **в другом месте не могли ∅**?» (Вампилов 1: 23).
«Na ploščadi vykopal orxideju. Cvetok takoj . . . Da razve èto xuliganstvo?» — «Èto neslyxannaja naglost'. Na ploščadi, pod nosom u milicii. **Vy** čto, **v drugom meste ne mogli ∅**?» (Vampilov 1: 23).
on square $dug\text{-}up_{1SG}$ $orchid_{ACC}$ flower of-a-kind PARTICLE really that(-is) hooliganism that(-is) incredible impertinence on square under nose at militia **you_{NOM}** what **in another place NEG $could_{3PL}$ ∅**
'"I dug up an orchid in the square. You know, the flower . . . Could that really be considered hooliganism?" "It's incredible impertinence. In the square, right under the militia's nose. Couldn't you have done it somewhere else?"'

(52) «У меня соринка в глаз попала. А ну посмотри! — Валя полезла руками Мишке в глаз, чтобы оттянуть веко. — Да **ты ∅ глазами, глазами**!» (Токарева: 482).

«U menja sorinka v glaz popala. A nu posmotri! — Valja polezla rukami Miška v glaz, čtoby ottjanut' veko. — Da **ty ∅ glazami, glazami**!» (Tokareva: 482).
at me speck-of-dust$_{NOM}$ into eye$_{ACC}$ fell COMPOUND PARTICLE look$_{IMPER}$ Valja$_{NOM}$ went hands$_{INSTR}$ Mishka$_{DAT}$ into eye$_{ACC}$ in-order-to pull-back eyelid$_{ACC}$ PARTICLE **you$_{NOM}$ ∅ eyes$_{INSTR}$ eyes$_{INSTR}$**
'"I got a speck of dust in my eye. Take a look!" Valja poked her hands toward Mishka's eye to pull back the eyelid. "Eh, look with your *eyes*, with your *eyes*!"'[9]

(53) «Послушайте, говорит, хотите со мною в театр поехать?» — В театр? как же бабушка-то?» — «Да **вы**, говорит, **∅ тихонько от бабушки** . . . » (Достоевский 1: 141).
«Poslušajte, govorit, xotite so mnoju v teatr poexat?» — «V teatr? kak že babuška-to?» — «Da **vy**, govorit, **∅ tixon'ko ot babuški** . . . » (Dostoevskij 1: 141).
listen$_{IMPER}$ says$_{3SG}$ want$_{2PL}$ with me$_{INSTR}$ to theater go$_{INFIN}$ to theater what about grandma PARTICLE **you$_{NOM}$** says$_{3SG}$ **∅ secretly from grandma$_{GEN}$**
'"Listen," he says, "do you want to go to the theater with me?" "The theater? But what about Grandma?" "Don't tell her," he says.'

(54) Она была, правда, грустна и рассеянна и всё отставала, но он не понимал ничего, а думал, что **то она ∅ от усталости** (Золотые: 113).
Ona byla, pravda, grustna i rassejanna i vsë otstavala, no on ne ponimal ničego, a dumal, čto **èto ona ∅ ot ustalosti** (Zolotye: 113).
she$_{NOM}$ was it-is-true sad and distracted and continually fell-behind **but he$_{NOM}$** NEG understood and thought that **that she$_{NOM}$ ∅ from tiredness**
'It's true, she was sad and distracted and kept falling behind, but he didn't understand what was going on, he thought she was just tired.'

(55) [Everyone is looking at a dead bird on the ground]
«Пролили кровь, раскаиваемся». — «Что за зверь?» — «Сорока». — «**За что вы её ∅**, бедняжку?» (Вампилов 1: 50).
«Prolili krov', raskaivaemsja». — «Čto za zver'?» — «Soroka». — «**Za čto vy eë ∅,** bednjažku?» (Vampilov 1: 50).
spilled$_{1PL}$ blood$_{ACC}$ repent$_{1PL}$ [idiom: what kind of creature is it] magpie **for what you$_{NOM}$ it$_{ACC.FEM}$ ∅** poor-thing
'"We spilled blood, we repent." "What kind of creature is it?" "A magpie." "Why did you kill it, the poor thing?"'

4. *Do the overt categories in the clause fit the syntax and semantics of the closely described patterns* (*here—for motion, speaking, or hitting*)*? If yes, use this analysis. If no, go to 5.* Examples (12)–(47) would be analyzed thus.

5. *If there is an object, does it strongly suggest—based on semantics and case marking—some action? If yes, use this analysis. If no, go to 6.* Such analysis would require use of an ontology and its associated lexicon, which would be searched for instances in which the semantic class of object was used as the default filler for the patient role. Then the lexicon could be checked for which verbs of that semantic class permit an object with the given case marking. If only one verb, then it is selected. If more than one verb, whichever

ontological concept includes all valid verbs (perhaps a parent of several children concepts) will be selected.

(56) . . . [Мы] встали и отправились играть. «Ну во что? — спросила Любочка, щурясь от солнца и припрыгивая по траве. — **Давайте ∅ в Робинзона**» (Толстой 2: 30).
. . . [My] vstali i otpravilis' igrat'. «Nu vo čto? — sprosila Ljubočka, ščurjas' ot solnca i priprygivaja po trave. — **Davajte ∅ v Robinzona**» (Tolstoj 2: 30).
[we]$_{NOM}$ got-up and set-off play$_{INFIN}$ so at what asked Lyubochka$_{NOM}$ squinting from sun and jumping-around on grass **let's ∅ at Robinson$_{ACC}$**
'. . . We got up and set off to play. "What are we going to play?" asked Lyubochka, squinting from the sun and jumping up and down on the grass. "Let's play Robinson and Crusoe."'

(57) Это **вы ∅ сарделькy**? (Земская 1973: 295).
Èto **vy ∅ sardel'ku** (Zemskaja 1973: 295).
PARTICLE **you$_{NOM}$ ∅ hot-dog$_{ACC}$**
'Are you eating a hot dog?'

(58) Это **вы ∅ Ремарка**? (Земская 1973: 295).
Èto **vy ∅ Remarka**? (Zemskaja 1973: 295).
PARTICLE **you$_{NOM}$ ∅ Remarque$_{ACC}$**
'Are you reading (something written by) Remarque?'

6. *If the class of actions that could have the given object as a patient is too large, search for sentence clues to suggest one or another interpretation. If found, use that interpretation; if not, go to 7*. The notion of "default action associated with an object" is, of course, a simplification—after all, a flyswatter can be used to reach something that has fallen behind the stove, and a bicycle can be polished rather than ridden. In order to disambiguate in verbless sentences, there must be lexical or contextual clues. In (59) the instrument *shampoo* suggests that the theme, *hair*, is being washed, and in (60b)—in contrast to (60a)—the adjunct *in the store* suggests that the meat is being bought, not eaten.[9]

(59) **Голову я ∅ шампунем** обычно (Земская 1973: 306–307).
Golovu ja ∅ šampunem obyčno (Zemskaja 1973: 306–307).
head$_{ACC}$ I$_{NOM}$ ∅ shampoo$_{INSTR}$ usually
'I usually wash my hair with shampoo.'

(60) a. Это **вы ∅ фарш**? (Земская 1973: 296).
Èto **vy ∅ farš**? (Zemskaja 1973: 296).
PARTICLE **you$_{NOM}$ ∅ meat-stuffing$_{ACC}$**
'Is that meat stuffing you're making [or: you bought, etc.]?'
b. Это **вы ∅ фарш такой в магазине**? (Земская 1973: 296).
Èto **vy ∅ farš takoj v magazine**? (Zemskaja 1973: 296).
PARTICLE **you$_{NOM}$ meat-stuffing$_{ACC}$ that-kind$_{ACC}$ in store$_{PREP}$**
'Did you really get meat stuffing like this at the store? [implying: it's SO good!]'

Using contextual clues requires reference to ontological concepts that represent simple events or complex events (scripts). When processing (60b), the simple events BUY, SELL, and SHOP should be the only candidates because they should be the only concepts with a default LOCATION: STORE. SHOP will be excluded as a candidate because none of the Russian verbs linked to it take an ACC direct object. Disambiguating between BUY and SELL might be accomplished using heuristics, for example, giving preference to the BUY interpretation if the subject is specific and leaving both options open for later disambiguation if the subject is generalized (e.g., *Do **they** sell such things in stores?* in which case the Russian variant would use the Indefinite Personal Construction, decribed in chapter 13). This and other heuristics can only be evaluated during system testing.

Processing (59) could exploit a WASH-HAIR script, if available, or the simple event WASH if not. A WASH-HAIR script would have default INSTRUMENT: SHAMPOO and the default THEME: HAIR.[10] The simple event WASH, by contrast, would have the default INSTRUMENT: SOAP and no default THEME. Since shampoo is a type of soap, for a generalized ontology it would arguably be best to link the lexical item *shampoo* to the concept SOAP with *lexical* restriction of the THEME to "head/hair" (i.e., there is no need for a concept SHAMPOO). This combination of lexical and ontolgical specification should lead to a correct analysis of the missing verb.

7. *Is there any script into which the overt elements and any elements in the preceding sentences readily fit? If so, instantiate the script and analyze the verb accordingly. If not, go to 8.* In this instance, unlike the last one, there are no natural events associated with the object, so searching for an applicable script is necessary.

Bar or being drunk script

(61) [Sayapin and Kuzakov are preparing to drag Zilov out of a café where he got dead drunk]
«Дождь пошёл». — А вот **мы его ∅ по дождичку**» (Вампилов 3: 230).
«Dožd' pošël» — «A vot my ego **∅ po doždičku**» (Vampilov 3: 230).
rain_{NOM} started COMPOUND PARTICLE $\textbf{we}_{\textbf{NOM}}$ $\textbf{him}_{\textbf{ACC}}$ **∅ through rain**
'"It's started to rain." "Then we'll haul him off in the rain."'[11]

Studying script

(62) **И мы за полгода ∅ весь курс** (Земская 1973: 305).
I my za polgoda **∅ ves' kurs** (Zemskaja 1973: 305).
and $\textbf{we}_{\textbf{NOM}}$ **in half year ∅** $\textbf{whole course}_{\textbf{ACC}}$
'We did the whole course in half a year.'

Asking and answering questions script

(63) «Я понимаю, вопрос деликатный, но **я ∅ не официально**, не с агентурной — ха-ха — точки зрения, а как . . . отец и даже как дед . . . » (Войнович 2: 69).
«Ja ponimaju, vopros delikatnyj, no **ja ∅ ne oficial'no**, ne s agenturnoj — xa-xa — točki zrenija, a kak . . . otec i daže kak ded . . . » (Vojnovič 2: 69).

I_{NOM} understand question$_{NOM}$ delicate but **I_{NOM} Ø NEG officially** NEG from secret-service ha-ha point of-view but like father and even like grandfather
'"I understand that this is a delicate question, but I'm not asking officially, not from a secret-service—ha-ha—point of view, but like . . . a father or even a grandfather . . ."'

Marrying script (*with country-specific, religion-specific, and so forth, variations*)

(64) «Может, ты в этом Тбилиси уже штампик в паспорт проставил?» Изумлённый таким оборотом дела, Деточкин полез в пиджак и предъявил Любе свой неженатый паспорт. «Это ничего не значит, — вздохнула Люба, **— можно Ø и без печати**» (Брагинский и Рязанов: 67).
«Možet, ty v ètom Tbilisi uže štampik v pasport prostavil?» Izumlënnyj takim oborotom dela, Detočkin polez v pidžak i pred"javil Ljube svoj neženatyj pasport. «Èto ničego ne značit, — vzdoxnula Ljuba, **— možno Ø i bez pečati**» (Braginskij i Rjazanov: 67).
{ . . . } that$_{NOM}$ nothing NEG means sighed Ljuba$_{NOM}$ **possible$_{IMPERS}$ Ø even without stamp$_{GEN}$**
'"Maybe you already got your passport stamped 'married' in Tbilisi?" Shocked by such a suggestion, Detochkin stuck his hand in his jacket pocket and showed Lyuba his bachelor's passport. "That doesn't mean anything," Lyuba sighed. "You can do it without a stamp, too."'

8. *Represent the verb as generalized 'do'* (*ontologically, EVENT*) *and see if later processing provides further information.* This option is actually no different from processing a very general verb like *do* when it is encountered overtly in text.

(65) Он целует меня . . . Отвожу лицо. «Не надо». — «Почему?» — «Не надо». **Это я Ø из самолюбия** (Токарева: 73).
On celuet menja . . . Otvožu lico. «Ne nado». — «Počemu?» — «Ne nado». **Èto ja Ø iz samoljubija** (Tokareva: 73).
he$_{NOM}$ kisses me$_{ACC}$ turn-away$_{1SG}$ face$_{ACC}$ NEG necessary$_{IMPERS}$ why NEG necessary$_{IMPERS}$ PARTICLE **I_{NOM} Ø from pride**
'He kisses me . . . I turn away. "Don't." "Why not?" "Don't." It's my pride that's making me act this way.'

3. The parameters and values for cross-linguistic description

The inventory of parameters and values required to describe Multi-VE includes those shown in table 10.4.

At least two other languages employ Multi-VE, although in a much more limited range of contexts: Polish and Czech. The tight usage restrictions in these languages might be, at least in part, explained by the fact that both Polish and Czech use subject *pro*-drop more consistently than Russian, making overt subjects relatively less

TABLE 10.4 Parameters and values for describing Multi-VE (all but the first apply separately to each meaning: motion, speaking, etc.)

Parameters	*Values*
Meaning (semantic class) of elided verb	Motion
	Speaking
	Hitting
	Living/surviving
	Doing/acting (in a context-specified manner)
	Other
Meaning (semantic class) of licensors	Human/animal agent
	Human/animal experiencer
	Adverbial expressing . . .
	Object indicating . . .
	Etc.
Syntactic realization of licensors	NP argument
	NP adverbial
	PP (possibly with certain preposition and/or object case marking)
	Adverb
	Auxiliary verb
	Impersonal predicate
Number of licensors	2, 3, 4
Prosodic requirements	None
	Emphatic intonation
	Other intonation
Productivity of pattern	Highly productive
	Of limited productivity
	Unproductive (fixed expressions only)
Stylistic nuances	Highly colloquial
	"As if it is occurring before one's eyes"
	Etc.

prevalent—and it is overt subjects that most often participate in licensing Multi-VE in Russian. In Polish and Czech, the number of structures that show Multi-VE is so limited that listing them as collocations with variables seems to be the minimal-complexity approach. However, compiling this inventory and determining that this elliptical strategy is not very productive required using the inventory of parameters and values listed earlier.

The most productive use of Multi-VEs in Polish and Czech is for elided verbs of hitting, where each relevant category listed in the licensing strategies is a variable, and the strategies can be fully described by the patterns used for Russian.

Hitting patterns

(66) Cz: Jak mám vědět, **proč on jeho Ø**?
how have$_{1.SG}$ know$_{INFIN}$ **why he$_{NOM}$ him$_{ACC}$ Ø**
'How am I supposed to know why he hit him <beat him up, and so on>?'

(67) P: **On mnie ∅ z byka**, a **ja go ∅ w szczękę** (Maciejewski 1991).
he$_{NOM}$ me$_{ACC}$ ∅ with head and$_{CONTR}$ **I$_{NOM}$ him$_{ACC}$ ∅ in jaw$_{ACC}$**
'He butted me with his head and I belted him in the jaw.'

(68) P: **Ja ∅ go pięścią w bok**.
I$_{NOM}$ ∅ him$_{ACC}$ fist$_{INSTR}$ in side
Cz: A **já ∅ mu pěstí do boku**.
and **I$_{NOM}$ ∅ him$_{DAT}$ fist$_{INSTR}$ into side$_{GEN}$**
'(And) I jammed my fist into his side.'

(69) P: Wilk zęby wyszczerzył, tak go światło oślepiło. A **Maciuś ∅ mu kulę w samo oko** (Korczak: 193).
wolf$_{NOM}$ teeth$_{ACC}$ bared so-much him$_{ACC}$ light$_{NOM}$ blinded and **Maciuś$_{NOM}$ ∅ him$_{DAT}$ bullet$_{ACC}$ in the-very eye$_{ACC}$**
'Blinded by the light, the wolf bared its teeth. And Maciuś shot him right in the eye.'

(70) Cz: Ale **já mu ∅**!
but **I$_{NOM}$ him$_{ACC}$ ∅**
'But I'll show him <I'll give it to him, and so on>!'

There are four fixed patterns that convey a meaning straddling *speak* and *react*—much like the English slang *was like* in: *So he goes, "You've got to be kidding!" and I was like, "No, I'm totally serious."* In types 1–3 the human subject is a variable, while in 4 both the subject and the *about*-phrase are variables. It is not clear whether an empty category should be posited in these verbless clauses, since positing it might incorrectly imply both that the elliptical process was productive and that a verb could readily be inserted, which it cannot. I provisionally leave it out and, instead, place the whole phrase in boldface for emphasis.

P and Cz Speaking Pattern 1

(71) [In reference to Pavel's having learned some shocking news]
P: „A **co na to Paweł**?" — „Ach, był za . . . zaszokowany" (Olesky and Swan: 280).
Cz: „A **co na to Pavel**?" — „Byl š . . . šokován."
and$_{CONTRASTIVE}$ **what$_{ACC}$ to that$_{ACC}$ Pavel$_{NOM}$** (ach,) was$_{3.SG}$ sh . . . shocked
'"And how did Pavel react <What did Pavel say, and so on> to that?" "Oh, he was sh . . . shocked."'

P and Cz Speaking Pattern 2

(72) P: „[Adam] twierdzi, że to znajoma z pracy i że ma kłopoty i że on chciałby ją pocieszyć i dlatego właśnie pogłaskał ją po ręku." — „**A ty co**?" (Olesky and Swan: 148).
{ . . . } **and you$_{NOM}$ what$_{ACC}$**
Cz: { . . . } „**A co ty?**"

{ . . . } **and what$_{ACC}$ you$_{NOM}$**
'"Adam claims that this was an acquaintance from work and that she had problems and he wanted to console her and that's why he stroked her hand." "And what did you say to that <how did you react, what did you do, and so on>?"'

P and Cz Speaking Pattern 3

(73) P: „Ona stawia czajnik: «Zimno». **Ja nic**" (Herbert: 341).
Cz: "Ona staví konvici: 'Je zima.' **Já nic <**Nevadí**>**.[12]
she$_{NOM}$ puts-on kettle$_{ACC}$ (is) cold **I$_{NOM}$ nothing$_{ACC}$**
'She puts on the kettle: "It's cold." I don't respond.'

P and Cz Speaking Pattern 4

(74) P: „Nie mam ochoty uczestniczyć w tej sielance. Stary cymbał. Jego świat rozsypuje się, **a on tylko o swoich przepisach**" (Kruczkowski 2: 103).
NEG have$_{1.SG}$ desire$_{GEN}$ participate$_{INFIN}$ in this idyll old dolt$_{NOM}$ his world$_{NOM}$ is-going-to-pieces REFL **and$_{CONTR}$ he$_{NOM}$ only about self's$_{PREP}$ regulations$_{PREP}$**
'"I don't want to take part in this idyll. The old dolt. His world is going to pieces, and he keeps on about his regulations."'

(75) Cz: **A on pořad <jen> o svých problémech.**
and$_{CONTR}$ he$_{NOM}$ constantly <only> about self's$_{PREP}$ problems$_{PREP}$
'And he's always talking about <all he talks about is> his own problems.'

I have found evidence of only one configuration in which a verb of motion can be elided, and it applies only to Polish, not to Czech:

(76) [Someone enters a secretary's office. The secretary says:]
P: **Pan do kogo**?
you$_{NOM.MASC}$ to whom$_{GEN}$
'Who are you here for?'

Other common Multi-VE structures in Polish and Czech are listed here with examples. Russian equivalents are provided as well, since some of the meanings are slightly idiomatic and, perhaps even in Russian, the phrases are worth listing as collocations with variables. As before, it is not clear whether empty categories should be posited in these sentences. For simplicity's sake, they will not be.

Pattern 1: R: (x) *ne možet tak / P:* (x) *nie może tak / Cz:* (x) *nemůže tak*
x *can't take this <can't do* something *this way, and so on>*

(77) P: „Przede wszystkim trzeba by tu zainstalowa jakąś kuchenkę i piecyk. **Pan dyrektor nie może tak** w tym zimnie i wilgoci" (Szaniawski 2: 45).[13]
above all required$_{IMPERS}$ CONDIT here install$_{INFIN}$ some-kind-of stove$_{ACC}$ and heater$_{ACC}$ **Pan director$_{NOM}$ NEG can this-way** in this cold and dampness

‘‘‘Most important, we need to install some kind of stove and heater in here. You can’t go on like this, in this cold and dampness.’’’

R: «**Vy ne možete tak**, v ètom xolode i vlažnosti».
you$_{NOM}$ NEG **can this-way** in this cold and dampness
‘‘‘You can’t go on like this, in this cold and dampness.’’’

Cz: „To **tak nemůžete**, v té zimě i vlhkosti.”
PARTICLE **this-way cannot$_{2.PL}$** in this cold and dampness
‘‘‘You can’t go on like this, in this cold and dampness.’’’

Pattern 2: R: (x) *ne možet bol’ še / P:* (x) *nie może dłużej / Cz:* (x) *nemůže dál*
x *can’t stand this <go on like this, and so on> anymore*

(78) P: „A co do niego, Hermana . . . ci chłopcy przecież widzą, co się dzieje. Są **tacy, co już nie mogą dłużej**” (Kruczkowski 2: 135).
and as concerns (idiom) him$_{GEN}$ Herman$_{GEN}$ these boys$_{NOM}$ after-all see what PARTICLE is-happening exist$_{3.PL.}$ **those$_{NOM}$ that already** NEG **can$_{3.PL.}$ more**
‘‘‘And as for him, Herman . . . after all these boys see what’s going on. Some people can’t stand it any longer.’’’

R: «Est’ **te, kto ne možet bol’še**»
exist **those who$_{3.SG}$** NEG **can$_{3.SG}$ more**
‘‘‘Some people can’t stand it any longer.’’’

Cz: „Jsou **takoví co už dál nemůžou**.”
exist$_{3.PL.}$ **those$_{NOM}$ who** PARTICLE **more cannot$_{3.PL}$**
‘‘‘Some people can’t stand it any longer.’’’

Pattern 3: R: tak nel’zja / P: tak nie można / Cz: tak se ne dá (*one*) *mustn’t act this way <do this, and so on>*

(79) R: «Snova tost? Net, **tak nel’zja**, tol’ko vypili i snova» (Vampilov 1: 8).[14]
again toast no **in-this-way impossible$_{IMPERS}$** just drank$_{1PL}$ and again

P: „Nowy toast? Nie, **tak nie można**, przed chwilę wypiliśmy”
another toast no **in-this-way impossible$_{IMPERS}$** ago minute drank$_{1PL}$
‘‘‘Another toast? No, this isn’t the way to do it. We just had a toast and now another one.’’’

Pattern 4: R: x *ne naročno;* x *nečajanno / P:* x *nienaumyślnie / Cz:* x *nevědomky*
x *didn’t mean it <didn’t mean to do it, and so on>*

(80) P: Podstawił jej nogę . . . „Łobuzy!”— krzyczy gruba pani.— „**Ja nienaumyślnie**” (Korczak: 188).
stuck-out$_{3SG.MASC}$ her$_{DAT}$ foot$_{ACC}$ jerks$_{NOM}$ yells fat woman$_{NOM}$ **I$_{NOM}$ accidentally**
‘He stuck out his foot to trip her. . . . “Jerks!” yells the fat woman. “I didn’t mean it.”’

R: **«Ja ne naročno».**
I$_{NOM}$ NEG **intentionally**

Cz: "**Já nevědomky.**"
I$_{NOM}$ accidentally
'"I didn't mean it."'

Pattern 5: R: x *tak so vsemi / P:* x *tak ze wszystkimi / Cz:*
x *tak s každým*
x *acts like this with <does this to, and so on>*
everyone

(81) P: „ . . . Na naszym podwórku jest jedna dziewczyna, to z nią rady nie można dać sobie. Sama zaczepia, a jak jej coś zrobić, żeby ją tylko ruszyć, zaraz zaczyna wrzeszczeć i leci na skargę. I **ona tak ze wszystkimi** (Korczak: 181).[15]
{ . . . } and **she$_{NOM}$ this-way with everyone$_{INSTR}$**
'". . . In our neighborhood there's this girl who's impossible to deal with. She does something to you, and if you do something back, if you even touch her, she immediately starts screaming and goes to tell on you. And she's like this with everyone."'
R: «I **ona tak so vsemi**»
and **she$_{NOM}$ this-way with everyone$_{INSTR}$**
Cz: „To **ona tak s každým**."
PARTICLE **she$_{NOM}$ this-way with everyone**
'"And she's like this with everyone."'

This cross-linguistic comparison shows two things. First, the inventory of parameters and values is valid even for languages whose employment of Multi-VE is quite different. Second, comparing ellipsis potential in typologically similar languages can not only speed the description of a given language, if a similar one has already been described, but also suggest different ways of looking at a given structure—for example, considering more of the Russian structures to be phrases with variables than might otherwise have been done.

11

Ellipsis of Minor Parts of Speech

1. Ellipsis of conjunctions and relative pronouns

Conjunctions not only bind sentence components in a grammatical sense; they also indicate the semantic relationship between them. English, on the whole, prefers to express conjunctions overtly, so (1a) is more common than (1b):

(1) a. Yes, I pulled her hair, but she kicked me in the shin first.
b. Yes, I pulled her hair. She kicked me in the shin first.

As discussed in chapter 1, section 3.6, conjunction use is a language parameter whose value for a given language can range from practically always requiring overt conjunctions to practically never using conjunctions at all. But even within a given language, not all conjunctions are created equal with respect to their overt/covert status. For example, in English, the conjunction *that* is optional in most instances (2), coordinating conjunctions are usually overt except between the penultimate conjunct(s) in a series (3), and most other conjunctions need to be overt (4)–(5):

(2) I know (**that**) you wouldn't hold a grudge for nothing.
(3) They jumped rope ∅, had a catch, **and** played Marco Polo in the pool.
(4) My brother's sick, **so** you can't come over today.
(5) **If** you're good, you can have two desserts.

The absence of the conjunction *that* in (2) and the first *and* in (3) are quite "surfacy" phenomena that should be handled by syntactic rules. However, there are languages in which conjunctions can be elided outside of predictable grammatical constructions, which means that the semantics of the missing conjunction must be recovered from the semantics of the text entities being conjoined. Furthermore, the fact that a conjunction is missing must be detected in the first place.

Russian is one language in which conjunction ellipsis is widely possible, albeit limited to highly colloquial speech. The purpose of eliding conjunctions in Russian is to give an utterance liveliness and an emotional, colloquial flavor (Gvozdev 1965: 314). In fact, this type of ellipsis is so highly colloquial that generation grammars intended for people or computers need not include rules for it. Analysis grammars, however, must cover it—at least if those grammars are intended to process this speech register. The following examples illustrate the ellipsis of various types of Russian conjunctions.

Elided u/i *'and'*

(6) Я буду сидеть у окна ∅ смотреть когда папа с мамой придут (Лаптева 1976: 287).
Ja budu sidet' u okna ∅ smotret' kogda papa s mamoj pridut (Lapteva 1976: 287).
I_{NOM} will sit_{INFIN} at window $\varnothing_{\mathbf{AND}}$ $wait_{INFIN}$ when Dad_{NOM} and Mom_{NOM} will-come
'I'm going to sit by the window and watch for Mom and Dad to come.'

Elided u/i *'and, so' or* poètomu *'therefore'*

(7) Жаркое начало пригорать, ∅ я помешала его (Хмелевская 2: 175).
Žarkoe načalo prigorat', ∅ ja pomešala ego (Chmielewska 2: 175).
$roasted\text{-}meat_{NOM}$ started $burn_{INFIN}$ $\varnothing_{\mathbf{SO/AND}}$ I_{NOM} stirred it_{ACC}
'The roasted meat started to burn, so I stirred it.'

Elided если/esli *'if'*

(8) ∅ Позволишь — буду плакать, ∅ не позволишь — ни слёзкой я тебе не досажу (Пушкин, Русалка; quoted from Gvozdev 1965: 318).
∅ Pozvoliš' — budu plakat', ∅ ne pozvoliš' — ni slëzkoj ja tebe ne dosažu (Puškin, Rusalka).
$\varnothing_{\mathbf{IF}}$ $will\text{-}permit_{2SG}$ $will_{1SG}$ cry_{INFIN} $\varnothing_{\mathbf{IF}}$ NEG $will\text{-}permit_{2SG}$ not tear I_{NOM} you_{DAT} NEG $will\text{-}vex_{1SG}$
'If you permit me to, I'll cry; if you don't, I will not vex you with a single tear.'

(9) «∅ Я буду сто лет жить, не забуду» (Толстой 1: 286).
«∅ Ja budu sto let žit', ne zabudu» (Tolstoj 1: 286).
$\varnothing_{\mathbf{IF}}$ I_{NOM} will hundred years $live_{INFIN}$ NEG $forget_{1SG.FUT}$
'"If I live to be a hundred I'll never forget this."'

(10) «∅ Нальёшь самогон в блюдце — это ад, ∅ нальёшь в чашку — это рай. А мы вот пьём из стаканов. Это, Петька, и делает нас людьми. Понял?» (Пелевин: 353).

«∅ Nal'ëš' samogon v bljudce — èto ad, ∅ nal'ëš' v čašku — èto raj. A my vot p'ëm iz stakanov. Èto, Pet'ka, i delaet nas ljud'mi. Ponjal?» (Pelevin: 353).
∅$_{\text{IF}}$ pour$_{\text{2SG.FUT}}$ homebrew$_{\text{ACC}}$ into saucer it-is poison **∅$_{\text{IF}}$** pour$_{\text{2SG.FUT}}$ into cup it-is heaven but we$_{\text{NOM}}$ here drink from glasses$_{\text{GEN}}$ that Pet'ka PARTICLE makes us$_{\text{ACC}}$ people$_{\text{INSTR}}$ understood$_{\text{2SG}}$
'"If you pour homebrew into a saucer it's hell, and if you pour it into a cup it's heaven. [Homebrew drunk from a saucer is hell, and homebrew drunk from a cup is heaven.] But, see, we're drinking it out of glasses. *That*, Pet'ka, puts us right here on earth. Got it?"'

Elided когда/kogda *'when'*

(11) ∅ Ты пришёл они уже сели? (Лаптева 1976: 305).
∅ Ty prišël oni uže seli? (Lapteva 1976: 305).
∅$_{\text{WHEN}}$ you$_{\text{NOM}}$ came they$_{\text{NOM}}$ already sat-down
'When you arrived, had they already sat down?'

Elided relative pronoun который/kotoryj (*variable in form*) *'that'*

(12) А где мой кошелёк ∅ тут лежал? (Лаптева 1976: 291).
A gde moj košelëk ∅ tut ležal? (Lapteva 1976: 291).
but where my wallet$_{\text{NOM}}$ **∅$_{\text{THAT}}$** there was-lying
'And where's my wallet that was lying here?'

To process such sentences a system must first detect that there is a missing conjunction, then determine the sense of that conjunction—that is, the semantic relationship between conjoined text elements. Properties of the overt text elements can act as clues for creating heuristics. For example:

- if two verb phrases in the same inflectional form occur in series with no punctuation between them, there could be an elided *u/i* 'and' between them, as in (6);
- if two past-tense clauses are joined by a comma, there could be an elided *u/i* 'and,' *поэтому* 'so,' or *когда* 'when' between them, as in (7) and (11);
- if two future-tense clauses are joined by a comma, there could be an elided *если/esli* 'if' between them, as in (8)–(10);
- if a verb phrase without an overt syntactic subject follows a noun phrase belonging to a different clause, and the inflection of the verb agrees with the features of that head noun, there could be an elided relative pronoun between them, as in (12).

Heuristics like these must be further constrained, tested, and modified based on corpus trials for a given system. But even if they turn out to be useful, they do not provide a final solution to the semantic recoverability problem because of potential ambiguity. That is, there can simultaneously be several semantic relationships be-

tween clauses: in (7) there is a Cause-Effect relation (*so*, *therefore*) as well as an Actions-in-Series relation (*and*). Moreover, detecting and resolving conjunction ellipsis in examples like (9) will require not only semantic analysis but reasoning as well: the system must reason that the speaker cannot predict how long he or she will live, meaning that the utterance is unreal and suggesting that, perhaps, there is a hidden condition. Notably, the problem of ambiguous conjunctions is not limited to elliptical utterances, with polysemous English *and* being a familiar case in point.[1] So, work on recovering elided conjunctions should be incorporated into the larger work on usage restrictions for and the interpretation of overt conjunctions.

2. Ellipsis of prepositions

When prepositional phrases that use the same preposition in the same meaning occur in series, in many languages it is common to elide one or more of the prepositions. The decision of where to have prepositions overt and where not depends upon the semantics of the complement noun phrases and the rhythm of the utterance. In the following examples, for which the Russian and English variants are equally representative, an underscore is used to show where additional prepositions might be inserted (though notice that in some cases, like [14], one Russian variant—with two prepositions overt and the third elided—does not sound very good in English):

(13) С Дианой фон Фюрстенберг, имя которой вы можете обнаружить в Америку в любом доме **на** этикетках, _ платьях, _ духах, _ туфлях, мы увиделись в её нью-йоркской студии (Богуславская: 90).
S Dianoj fon Fjurstenberg, imja kotoroj vy možete obnaružit' v Amerike v ljubom dome **na** ètiketkax, _ plat'jax, _ duxax, _ tufljax, my uvidelis' v eë n'ju-jorkskoj studii (Boguslavskaja: 90).
with Diane von Furstenberg$_{INSTR}$ name$_{NOM}$ whose you$_{NOM}$ might find in America in any house **on** tags$_{INSTR}$ _ dresses$_{INSTR}$ _ perfume$_{INSTR}$ _ shoes$_{INSTR}$ we$_{NOM}$ met in her New-York studio-apartment$_{PREP}$
'We met with Diane von Furstenberg, whose name you can find in any American home on tags, dresses, perfume, shoes, in her studio apartment in New York.'

(14) За нашим столом говорят **о** шахматах, **о** бардах, _ делах академических . . . (Богуславская: 42).
Za našim stolom govorjat **o** šaxmatax, **o** bardax, _ delax akademičeskix . . . (Boguslavskaja: 42).
at our table talk$_{3PL}$ **about** chess$_{PREP}$ **about** bards$_{PREP}$ _ issues academic$_{PREP}$
'At our table they're talking about chess, bards, academic issues . . .'

(15) [About Paris]
А как там **в** изящных искусствах, **в** литературе? . . . _ политике?
A kak tam **v** izjaščnyx iskusstvax, **v** literature? . . . _ politike?
and how there **in** fine arts$_{PREP}$ **in** literature$_{PREP}$ _ politics$_{PREP}$
'And what's it like there in the fine arts, in literature? . . . politics?'

(16) . . . Для других [ленинградская блокада] превратилась в синоним детства — **с** маминой любовью, **с** бабушкой-дедушкой, **с** праздниками, _ страхами, **с** *новым*! . . . (Рачко: 31).
. . . Dlja drugix [leningradskaja blokada] prevratilas' v sinonim detstva — **s** maminoj ljubov'ju, **s** babuškoj-deduškoj, **s** prazdnikami, _ straxami, **s** *novym*! . . . (Račko: 31).
for other-people$_{GEN}$ [Leningrad blocade]$_{NOM}$ changed into synonym$_{ACC}$ childhood$_{GEN}$ **with** mother's love$_{INSTR}$ **with** Grandma-Grandpa$_{INSTR}$ _ fears$_{INSTR}$ **with** new$_{ADJ.INSTR}$
'. . . For other people, the blockade of Leningrad became synonomous with childhood—with mother's love, with Grandma and Grandpa, with holidays, fears—with all sorts of *new things*! . . .'

Not all unexpressed prepositions are, however, necessarily syntactically elided, since one can first coordinate a number of noun phrases, then incorporate them as a group into a prepositional phrase. This scenario is the least-complexity analysis of examples like (13), in which one preposition could be seen as carrying over to four coordinated noun phrases in series:

(13a) **на** [этикетках, платьях, духах, туфлях]
na [ètiketkax, plat'jax, duxax, tufljax]
on [tags, dresses, perfume, shoes]

In examples that show other combinations of overt and covert prepositions, like (14)–(16), it is not clear whether the analysis of preposition ellipsis ((a) variants) or NP conjunction ((b) variants) is more appropriate:

(14) a. **о** шахматах, **о** бардах, ∅ делах академических
o šaxmatax, **o** bardax, ∅ delax akademičeskix
about chess, **about** bards, ∅ academic issues
b. **о** шахматах, **о** [бардах, делах академических]
o šaxmatax, **o** [bardax, delax akademičeskix]
about chess, **about** [bards, academic issues]

(15) a. **в** изящных искусствах, **в** литературе . . . ∅ политике
v izjaščnyx iskusstvax, **v** literature . . . ∅ politike
in fine arts, **in** literature . . . ∅ politics
b. **в** изящных искусствах, **в** [литературе . . . политике]
v izjaščnyx iskusstvax, **v** [literature . . . politike]
in fine arts, **in** [literature . . . politics]

(16) a. **с** маминой любовью, **с** бабушкой-дедушкой, **с** праздниками, ∅ страхами, **с** *новым*
s maminoj ljubov'ju, **s** babuškoj-deduškoj, **s** prazdnikami, ∅ straxami, **s** *novym*
with mother's love, **with** Grandma and Grandpa, **with** holidays, ∅ fears, **with** *new*
b. **с** маминой любовью, **с** бабушкой-дедушкой, **с** [праздниками, страхами], **с** *новым*
s maminoj ljubov'ju, **s** babuškoj-deduškoj, **s** [prazdnikami, straxami], **s** *novym*
with mother's love, **with** Grandma and Grandpa, **with** [holidays, fears], **with** *new*

In fact, the very notion of "appropriate" depends entirely on the application of the grammatical rules used to treat such sentences. Consider what part of the routine for analyzing a Russian sentence like (14) might do (in non-application-specific terms):

1. Analyze the words *delax akademičeskix* as a noun phrase, based on their sequential placement and agreement in case marking (PREP) and number (PL).
2. Search for the preposition this noun phrase is associated with, since the PREP case can only be assigned by prepositions in Russian.
3. Use the closest preceding preposition (not in a subordinate clause, etc.) to expand the structure to *o delax akademičeskix*; or, alternatively, determine if the preceding word(s) constitute a noun phrase with the same case marking, thus making possible NP coordination and a joint search for the preposition that assigns PREP case.

This example in Russian is among the easier ones, since PREP case marking unambiguously triggers the search for a preposition to have assigned it. However, when other types of case marking are involved—like the INSTR in (16) (which can be assigned lexically or even semantically, as well as by a preposition), special rules for conjoined syntactic structures (noun phrases or prepositional phrases) might be required.

If generation were the task, semantically oriented heuristics might be incorporated as a means of deciding which prepositions to have overt and covert. In (17), for example, *dinner* and *a concert* are sufficiently similar to be subsumed under a single preposition: in the Ontological Semantics ontology they are descendants of the common ancestor SOCIAL-EVENT:

(17) «До недавнего времени женщин вообще не пускали в мужские клубы. Да-да, представьте! Нас пускали только вечером, **на** ужин или концерт, и обязательно с мужчиной!» (Богуславская: 252).
«Do nedavnego vremeni ženščin voobšče ne puskali v mužskie kluby. Da-da, predstav'te! Nas puskali tol'ko večerom, **na** užin ili koncert, i objazatel'no s mužčinoj!» (Boguslavskaja: 252).
until recent time $women_{ACC}$ totally NEG $let\text{-}in_{3PL}$ into men's clubs yes-yes $imagine_{IMPER}$ us_{ACC} $let\text{-}in_{3PL}$ only $evening_{INSTR}$ **to** dinner or concert and definitely with man_{INSTR}
'"Until recently women weren't allowed in men's clubs at all. Really, imagine that! They let us in only in the evenings, for dinner or a concert, and only when escorted by a man."'

3. Ellipsis of conditional particles

In Russian, the particle *бы/by* is used in sentences that express conditions. It requires that the verb it modifies be in the past tense, regardless of the temporal relation implied by the context. In glossing examples, however, I will translate the verb in the tense implied by the context, ignoring this surface grammatical rule.

(18) «**Я бы** приехал к тебе, если **бы** знал, что не найду Сергея Иваныча» (Толстой 1: 117).
«Ja **by** priexal k tebe, esli **by** znal, čto ne najdu Sergeja Ivanyča» (Tolstoj 1: 117).
I **CONDIT** come to you if **CONDIT** knew that NEG find Sergei Ivanych
'"I'd come see you if I could be sure that I wouldn't run into Sergei Ivanych there."'

If a conditional sentence contains just two clauses, *бы/by* must be overt in both of them: it must occur once in the *if*-clause and once in the clause that expresses what would happen if the condition was met. However, if a conditional sentence contains three or more clauses, indicating that more than one thing would happen if the condition was met, ellipsis of one of the latter instances of *бы/by* is often possible, as shown in (19). In such cases, it is the second or later instance of *бы/by* that antecedes the ellipsis, never the first (in the *если/esli*–clause), which is why that one is not in boldface in the examples.

(19) Если бы у меня было время, я поехала **бы** в Польшу и занималась (**бы**) польским языком.
Esli by u menja bylo vremja, ja poexala **by** v Pol'šu i zanimalas' (**by**) pol'skim jazykom.
if CONDIT at me was time$_{NOM}$ I$_{NOM}$ go **CONDIT** to Poland and study (**CONDIT**) Polish$_{INSTR}$ language$_{INSTR}$
'If I had time, I'd go to Poland and study Polish.'

Ellipsis is possible in some cases even if the condition is in one sentence and the outcomes are in another, as in (20):

(20) Люди испугались бы, увидев своими глазами, во что превратились их души. Они на смерть пошли **бы**, а не остались ∅ покоренным народом» (Шварц 2: 298).
«Ljudi ispugalis' by, uvidev svoimi glazami, vo čto prevratilis' ix duši. Oni na smert' pošli **by**, a ne ostalis' ∅ pokorennym narodom» (Švarc 2: 298).
people$_{NOM}$ get-horrified CONDIT having-seen own$_{INSTR}$ eyes$_{INSTR}$ into what changed their soul$_{NOM}$ they$_{NOM}$ to death go **CONDIT** and NEG remain ∅$_{\textbf{CONDIT}}$ conquered$_{INSTR}$ people$_{INSTR}$
'"People would be horrified to see for themselves what had become of their souls. They would face death rather than remain a conquered people."'

When both conjuncts contain not only *бы/by* but coreferential pronominal direct objects as well, ellipsis is promoted because repeating both the conditional marker and the direct object in the second conjunct would sound too repetitive and repeating one but not the other would be unbalanced (cf. the discussion of dependencies in ellipsis in chapter 11):

(21) «Слушай, Васька . . . Гад ты, и больше никто. Взяла **бы тебя** и убила ∅ ∅» (Вампилов 2: 84).
«Slušaj, Vas'ka . . . Gad ty, i bol'še nikto. Vzjala **by tebja** i ubila ∅ ∅» (Vampilov 2: 84).

listen$_{IMPER}$ Vas'ka creep$_{NOM}$ you$_{NOM}$ and more nobody$_{NOM}$ took$_{FEM.SG}$ **CONDIT** **you$_{ACC}$** and kill **∅$_{CONDIT}$** **∅$_{ACC}$**
'"Listen, Vas'ka . . . You're a creep, that's what you are. I could just kill you."'

When the conditional marker occurs in three or more successive conjuncts, matters of rhythm and style in large part determine preferred patterns of ellipsis. In fact, the effects of rhythm and style appear to be considerably greater in contexts with repeated *бы/by* than in contexts with repeated direct objects, meaning that it is harder to formulate an algorithm for determining ellipsis potential in a given context as well as the preferred pattern of elided and overt categories. Some multiconjunct examples, like (22), show a combination of overt and covert *бы/by*, while others, like (23), have all instances of *бы/by* overt, even though this is not grammatically mandatory:

(22) . . . И если бы не было посторонних людей и если бы такое поведение не считалось неприличным, не осуждалось бы общественным мнением, — она положила **бы** голову ему на грудь, прикрыла ∅ глаза и сказала ∅: «Я счастлива» (Токарева: 371).
. . . I esli by ne bylo postoronnix ljudej i esli by takoe povedenie ne sčitalos' nepriličnym, ne osuždalos' by obščestvennym mneniem, — ona položila **by** golovu emu na grud', prikryla ∅ glaza i skazala ∅: «Ja sčastliva» (Tokareva: 371).
{ . . . } she take$_{FEM.SG}$ **CONDIT** head$_{ACC}$ him$_{DAT}$ on chest close$_{FEM.SG}$ **∅$_{CONDIT}$** eyes$_{ACC}$ and say$_{FEM.SG}$ **∅$_{CONDIT}$** I$_{NOM}$ happy
'. . . And if there weren't any strangers there and if that sort of behavior weren't considered indecent, if it wouldn't be denounced by public opinion, she would have placed her head on his chest, closed her eyes and said: "I'm happy."'

(23) [Dolly asks Anna if she would forgive Stiva for cheating on her]
«Да, я простила **бы.** Я не была **бы** тою же, да, но простила **бы**, и так простила **бы**, как будто этого не было, совсем не было» (Толстой 1: 91).
«Da, ja prostila **by**. Ja ne byla **by** toju že, da, no prostila **by**, i tak prostila **by**, kak budto ètogo ne bylo, sovsem ne bylo» (Tolstoj 1: 91).
yes I$_{NOM}$ forgive **CONDIT** I NEG was **CONDIT** same PARTICLE yes but forgive **CONDIT** and in-such-a-way forgive **CONDIT** as if that$_{GEN}$ NEG was totally NEG was
'"Yes, I would forgive him. I wouldn't be unaffected by it, but I would forgive him—I'd forgive him so completely that it would be as if it had never happened at all."'

Perceived redundancy can be reduced not only by eliding a conditional marker but also by using its shortened form, *б/b*, in one or more instances:

(24) . . . [Левин] двигался без усилия мышц и чувствовал, что всё может сделать. Он был уверен, что полетел **бы** вверх или сдвинул **бы** угол дома, если **б** это понадобилось (Толстой 1: 516).
. . . [Levin] dvigalsja bez usilija myšc i čuvstvoval, čto vsë možet sdelat'. On byl uveren, čto poletel **by** vverx ili sdvinul **by** ugol doma, esli **b** èto ponadobilos' (Tolstoj 1: 516).

[Levin]$_{NOM}$ moved without effort$_{GEN}$ muscles$_{GEN}$ and felt that everything$_{ACC}$ can do he$_{NOM}$ was sure that fly **CONDIT** up or move **CONDIT** corner house$_{GEN}$ if **CONDIT** that was-necessary
'. . . [Levin] moved effortlessly and felt like he could do anything. He was sure that he could fly or lift the corner of a house if that were necessary.'

There is one important restriction on the ellipsis of *бы/by*: the subjects of the latter two *бы/by*-clauses must be the same, so ellipsis is not possible in sentences like (25):

(25) Если бы мы выиграли крупную сумму, *я* бросила **бы** работу, а *муж* купил **бы** себе Феррари.
Esli by my vyigrali krupnuju summu, *ja* brosila **by** rabotu, a *muž* kupil **by** sebe Ferrari.
if CONDIT we$_{NOM}$ won large sum$_{ACC}$ I$_{NOM}$ quit **CONDIT** job$_{ACC}$ and husband$_{NOM}$ buy **CONDIT** self Ferrari$_{ACC}$
'If we won a heap of money, *I'd* quit my job and *my husband* would buy himself a Ferrari.'

Recall that DO ellipsis potential is also impeded (albeit not always ruled out) when the antecedent conjunct and ellipsis conjunct contain different subjects.

In fact, patterns of *бы/by* ellipsis have much in common with patterns of object ellipsis: for example, repeating an object or conditional marker multiple times in succession may be discouraged for stylistic reasons; a sentence often permits more than one pattern of overt and covert objects or conditional markers; and patterns of overt and covert objects and conditional markers must show logical and rhythmic balance. These common features suggest that some rules that affect ellipsis cut across grammatical categories. There is yet another similarity: if a number of successive conjuncts contain a given object (26) or conditional marker (27), that category can be fronted and "carry over" to all the conjuncts that would contain it (this is arguably a "Language Strategy" of the type discussed in chapter 13 rather than ellipsis as such):

(26) «Художники нуждаются в женщинах. Нуждаются в таких женщинах, которые **их** понимают и умеют заинтересовать» (Богуславская: 97–98).
«Xudožniki nuždajutsja v ženščinax. Nuždajutsja v takix ženščinax, kotorye **ix** ponimajut i umejut zainteresovat'» (Boguslavskaja: 97–98).
artists$_{NOM}$ need PREPOSITION women$_{PREP}$ need PREPOSITION such women$_{PREP}$ who **them$_{ACC}$** understand and know-how-to interest$_{INFIN}$
'"Artists need women. They need the sort of women who understand them and can keep them interested."'

(27) [About a car chase]
Неизвестно, как долго **бы** это продолжалось и чем закончилось, если бы Деточкину не бросился в глаза дорожный знак: «Осторожно, дети!» (Брагинский и Рязанов: 117).
Neizvestno, kak dolgo **by** èto prodolžalos' i čem zakončilos', esli by Detočkinu ne brosilsja v glaza dorožnyj znak: «Ostorožno, deti!» (Braginskij i Rjazanov: 117).

unclear how long CONDIT that$_{NOM}$ last and how end if CONDIT Detochkin$_{DAT}$ NEG threw-itself into eyes street sign$_{NOM}$ caution children
'It's not clear how long this would have continued and how it would have ended if Detochkin had not caught sight of a road sign: "Caution, children at play!"'

Having a fronted *бы/by* in the first conjunct does not, however, *require* that the following *бы/by* be elided: an overt second *бы/by* is perfectly normal in some contexts, especially if there is considerable distance between the two instances of *бы/by*, as in (28):

(28) « . . . **Я бы** с вами объехала вокруг света и не соскучилась **бы**» (Толстой 1: 81).
« . . . Ja **by** s vami ob''exala vokrug sveta i ne soskučilas' **by**» (Tolstoj 1: 81).
I$_{NOM}$ CONDIT with you circle around world and NEG get-bord CONDIT
'". . . I'd travel around the world with you and wouldn't get bored."'

To summarize, the ellipsis potential of conditional *бы/by* is affected by many of the same factors as the ellipsis potential of direct objects. However, developing algorithms for detecting and generating the ellipsis of *бы/by* will be difficult because on the one hand, conditional markers are not lexically selected by a predicate word (with their absence implying that they have been elided) and on the other hand, stylistic and rhythmic considerations play a particularly important role in determining the felicity of eliding *бы/by*.

4. Ellipsis of reciprocal and reflexive particles

In Polish, the particle *się* is an enclitic that carries out a number of functions. It can be semantically full, indicating that the verb or its derived nominal is reflexive or reciprocal, or it can be semantically vacuous, sometimes called inherent s*ię*. In keeping with the free word order of Polish, *się* can occur in various positions in the clause, subject to some restrictions. The ellipsis of *się* is discussed in some detail in chapter 12, so here I present only some highlights that suggest cross-linguistic parameters and values for its ellipsis. Since most of the ellipsis-affecting properties and elliptical patterns have been encountered before, the description between illustrative examples is brief.

Polish *się*, like Russian *by*, can be elided only in parallel configurations, that is, coordinate structures or Repetition Structures. The simpler the sentence, the more strongly ellipsis is preferred, in order to offset the impression of excessive repetition. Slash notation is used to emphasize the stylistic infelicity of overt-*się* variants, with ⌋⌊ indicating greater deviance than ⌋. As always, a star indicates ungrammaticality. The nature of *się* (reflexive, reciprocal, or inherent) will generally not be indicated except in cases where a contrast is important; in the word-for-word translation, *się* will remain as *się*.

(29) Uspokój **się**, uspokój (**się**) natychmiast.
calm$_{IMPER.SG}$ **się** calm$_{IMPER.SG}$ (**się**) at-once
'Calm down, calm down at-once.'

(30) Marek i Monika pozwali **się**, zaręczyli Ø/**się** i pobrali Ø/**się** w ciągu 3 miesięcy.
Marek$_{NOM}$ and Monika$_{NOM}$ met **się** engaged Ø/**się** and married Ø/**się** in span 3 months
'Marek and Monika met, got engaged, and got married within the span of three months.'

Reducing parallelism, as by changing the word order in the conjuncts, can decrease the imperative to employ ellipsis, as shown in (31), which is taken from Kupść 1999, p. 95 (which does not, however, show the optionality of *się* that I include):

(31) Jan umył **się** i szybko (**się**) ogolił.
Jan$_{NOM}$ washed$_{MASC.SG}$ **się** and quickly (**się**) shaved$_{MASC.SG}$
'Jan washed up [lit.: washed himself] and quickly shaved [lit: shaved himself].'

In addition, Polish *się*, like direct objects and *by* in Russian, can be fronted and carry over to more than one conjunct:

(32) Jan **się** umył, a Piotr (**się**) ogolił.
John$_{NOM}$ **się** washed and$_{CONTR}$ Peter$_{NOM}$ (**się**) shaved
'John washed up and Peter shaved.'

If conjuncts with different subjects are conjoined (and *się* is not fronted), in order for ellipsis to occur, two conditions must generally hold: the conjunction must be contrastive and there must be a natural semantic correlation between the actions conveyed in the conjuncts—just as with Russian direct objects:

(33) a. Jan **się** umył, a Piotr (**się**) ogolił.
John$_{NOM}$ **się** washed and$_{CONTR}$ Peter$_{NOM}$ (**się**) shaved
b. Jan **się** umył i Piotr **się**/*Ø ogolił.
John$_{NOM}$ **się** washed and$_{COORD}$ Peter **się**/*Ø shaved
'John washed up and Peter shaved.'

However, in contrast to Russian objects, it is not the case that clausal coordinate structures with coordinating *i* 'and' *always* blocks ellipsis, as shown by (34)–(35).

(34) Gdy Jacek **się** umyje i Marek (**się**) przebierze, znow będą wyglądali jak ludzie.
when Jacek$_{NOM}$ **się** bathes$_{3.SG}$ and Marek$_{NOM}$ (**się**) changes-clothes$_{3.SG}$ again will$_{3.PL}$ look like people
'When Jacek washes up and Marek changes his clothes, they'll look like human beings again.'

(35) Żeby **się** ojciec nie denerwował i matka (**się**) nie martwiła, nie powiemy im, co się naprawdę stało.
so-that **się** Father$_{NOM}$ NEG worry$_{3SG.MASC}$ and Mother$_{NOM}$ (**się**) NEG upset$_{3SG.FEM}$ NEG will-tell$_{1PL}$ them$_{DAT}$ what się really happened
'So that Father doesn't worry and Mother doesn't get upset, we won't tell them what really happened.'

These sentences are semantically much closer to pure coordination than (33b), meaning that there is no temporal or cause-effect relationship—which appears to be the crucial parameter that accounts for the different ellipsis judgments. Here we return to one of the difficult issues that opened the chapter: the fine semantic nuances of clausal relationships and their grammatical implications.

There is yet another way for clauses to be coordinated in Polish: using paired conjunctions like *i . . . i* 'both . . . and' and *ani . . . ani* 'neither . . . nor'. Such paired conjunctions emphasize the parallel nature of the conjuncts and thereby promote ellipsis, even if the clauses contain different subjects:

(36) I Jan **się** umył, i Piotr (**się**) ogolił.
and John_{NOM} **się** bathed and $\text{Peter}_{\text{NOM}}$ (**się**) shaved
'John washed up *and* Peter shaved.'

(37) Ani **się** Piotr nie umył ani (**się**) Jan nie ogolił.
neither **się** $\text{Peter}_{\text{NOM}}$ $_{\text{NEG}}$ bathed nor (**się**) John_{NOM} $_{\text{NEG}}$ shaved
'Peter didn't wash up and John didn't shave, either.'

In examples (33)–(37), the instances of *się* within each sentence matched—both were reflexive, both were inherent, and so on. If the sources of *się* do not match, ellipsis is not permitted, as noted in Kupść (1999: 108):

(38) Jan **się** elegancko ubrał i Maria uśmiechnęła **się/*∅**, gdy go zobaczyła. (Kupść 1999: 108)
John_{NOM} **się**$_{\text{REFL}}$ elegantly dressed and $\text{Maria}_{\text{NOM}}$ laughed **się**$_{\text{INHERENT}}$**/*∅** when him_{ACC} $\text{saw}_{\text{FEM.SG}}$
'John got dressed up and Mary smiled when she saw him.'

(39) Drzewa **się** połamały, ale gałęzie szybko zazieleniły **się/*∅** od nowa.
$\text{trees}_{\text{NOM}}$ **się**$_{\text{MIDDLE-VOICE}}$ $\text{broke}_{3.\text{PL}}$ but $\text{branches}_{\text{NOM}}$ quickly $\text{greened}_{3.\text{PL}}$ **się**$_{\text{INHERENT}}$/ ***∅** again
'The trees broke but their branches quickly became green again.' (Kupść 1999: 108)

These ellipsis patterns of Polish *się* lead us to revisit many of the ellipsis patterns and ellipsis-affecting properties seen for other major and minor parts of speech, adding to the list of relevant parameters "meaning of a particle" (e.g., inherent vs. reflexive vs. reciprocal *się*) and "fine-grained semantic correlation between clauses" (e.g., pure coordination vs. temporal sequence vs. cause-effect). The overlap between elliptical patterns for different parts of speech suggests that programs developed for one could be reconfigured for another by changing the values for certain parameters.

12

Dependencies in Ellipsis: A Polish Case Study

The interdependence among overt and elided categories represents both a contentful and a stylistic balance. Failing to strike this balance can result, on the one hand, in poorly understood utterances or, on the other, in utterances "with an accent." This chapter presents a particularly clear example of dependencies in ellipsis, using as a test case three elidable categories in Polish: the multifunction particle s*ię*, the conditional particle or morpheme *by/-by* (both realizations are possible in Polish), and direct objects with default ACC case marking. This sample follows naturally from the previous chapters, since the ellipsis of *się* was just discussed in chapter 11 and the ellipsis of *by* and direct objects in Polish have much in common with their Russian counterparts.[1]

The ellipsis potential of these categories, first in isolation, then in combination, will be considered in three ellipsis-promoting syntactic configurations: two-conjunct coordinated verb phrases, three-conjunct coordinated verb phrases, and Repetition Structures.

1. Background grammatical notes

Się is an enclitic that carries out a number of functions in Polish. It can indicate that the verb or its derived nominal is reflexive or reciprocal, it can create derived impersonals (cf. chapter 13, section 3.2), and it can be a semantically vacuous appendage to a verb, called inherent s*ię*. In keeping with the free word order of Polish, *się* can occur in various positions in the clause, either closely following the verb or

preceding the verb, just about anywhere in the clause except in clause-initial position.[2] See chapter 11, section 4, for the ellipsis properties of *się*.

By is the conditional marker in Polish. It can be either attached to the main verb or detached, in which case it precedes the verb. When attached to the verb, it is placed after gender agreement morphology but before the person-number marker. For example, *pływałabym* 'I would have swum' contains:

(1)	pływał-	-a-	-by-	-m
	$swim_{SG}$	FEM.SG	CONDIT	1SG

When detached, both *by* and the person-number marker, if there is one, precede the verb, since the person-number inflection must remain attached to *by*, as in *ja bym pływała*:

(2)	ja	by-	-m	pływał-	-a
	I_{NOM}	CONDIT	1SG	$swam_{SG}$	FEM.SG

Się and *by* can be elided almost exclusively in parallel configurations, unlike many other categories, whose ellipsis can be licensed outside of parallel configurations (subject ellipsis can be licensed by verbal morphology; VP ellipsis can be licensed by an auxiliary; DO ellipsis can be licensed by a discourse topic, etc.).[3]

2. Data and discussion

Before considering ellipsis dependencies among these categories, we need baseline facts about their individual ellipsis potential. In three-conjunct VP-coordinate structures, there is a strong preference for eliding the latter two instances of *się*, *by* and a direct object, if the latter's antecedent is a pronominal. This preference is represented in the examples using slash notation: for example, ∅/⫽*się* indicates that the overt-*się* variant is moderately awkward and ∅/⫽⫽*się* indicates that it is deviant (although still not ungrammatical). Note that the type of *się* (e.g., reflexive, inherent) does not matter as long as the instances of *się* in all verb phrases match.

(3) Marek i Monika pozwali **się**, zaręczyli **∅/⫽⫽się** i pobrali **∅/⫽⫽się** w ciągu 3 miesięcy.
$Marek_{NOM}$ and $Monika_{NOM}$ met **się** engaged **∅/⫽⫽się** and married **∅/⫽⫽się** in span 3 months
'Marek and Monika met, got engaged, and got married within the span of three months.'

(4) Gdybym mieszkała nad oceanem, pływała**bym**, biegała**∅$_{\mathbf{bym}}$ /⫽⫽**biegała**bym** i opalała**∅ $_{\mathbf{bym}}$/⫽⫽**opalała**bym** się codziennie.
if_{1SG} $lived_{FEM.SG}$ near ocean $swim_{FEM.SG}$-**by$_{\mathbf{1SG}}$** $run_{FEM.SG}$**∅ $_{\mathbf{by}}$-$_{1SG}$/⫽⫽**$run_{FEM.SG}$-**by$_{\mathbf{1SG}}$** and $sunbathe_{FEM.SG}$**∅ $_{\mathbf{by}}$-$_{1SG}$/⫽⫽**$sunbathe_{FEM.SG}$-**by$_{\mathbf{1SG}}$** się daily
'If I lived near the ocean, I would swim, run, and sunbathe every day.'

(5) „Dowiedziałem się wkrótce, że człowiek siedzący naprzeciw mnie jest łowcą dzikich słoni. Łowi **je**, oswaja **Ø/⌋⌊je** i oddaje **Ø/⌋⌊je** wyedukowane w służbę ludzi (Szaniawski 1: 93).
learned$_{1SG.MASC}$ się soon that man$_{NOM}$ sitting across-from me is trapper wild$_{GEN}$ elephants$_{GEN}$ traps$_{3SG}$ **them**$_{ACC}$ tames **Ø/⌋⌊them$_{ACC}$** and gives **Ø/⌋⌊them$_{ACC}$** educated to service people$_{GEN}$
‘‘‘Before long I learned that the man sitting across from me was a trapper of wild elephants. He traps them, tames them and hands them over, educated, to people for their use.’’’

Ellipsis in this configuration is strongly preferred because of a layering of ellipsis-promoting factors: superimposed upon the syntactic parallelism is phonetic identity between the antecedent and its coreferential categories. If phonetic identity does not obtain, as when the antecedent for the direct object is a referential expression rather than a pronoun, ellipsis is not as mandatory, as shown in (6). The (c) and (d) variants are not stylistically perfect according to some speakers, but they are not quite awkard, either, as indicated by (⌋).

(6) „Możemy wykonać **słonia** z gumy w odpowiedniej wielkości,
a. napełnić **go** powietrzem i wstawić **Ø** za ogrodzenie” (Mrożek: 110).
b. napełnić **Ø** powietrzem i wstawić **Ø** za ogrodzenie.”
c. (⌋) napełnić **Ø** powietrzem i wstawić **go** za ogrodzenie.”
d. (⌋) napełnić **go** powietrzem i wstawić **go** za ogrodzenie.”
can$_{1PL}$ make$_{INFIN}$ **elephant$_{ACC}$** from rubber in right size fill$_{INFIN}$ **it$_{ACC}$** air$_{INSTR}$ and put$_{INFIN}$ **Ø** behind railing
‘‘‘We can make an elephant of the right size out of rubber, fill it with air and put it behind the railing.’’’

When a category is repeated only twice, as in two-conjunct VP-coordinate structures, ellipsis is still preferred but in some cases is felt to be less mandatory:

(7) Jak tylko Maciuś umył **się** i ubrał **Ø/⌋⌊się**, zjawił się poseł króla zagranicznego z pozdrowieniem (Korczak: 77).
as soon-as Maciuś$_{NOM}$ washed **się** and dressed **Ø/⌋⌊się** appeared się envoy$_{NOM}$ king$_{GEN}$ foreign$_{GEN.ADJ}$ with greetings
‘As soon as Maciuś had washed up and gotten dressed, the foreign king’s envoy showed up bringing greetings.’

(8) Gdyby pogoda była ładna, siedziała**bym** i czytała**Ø$_{bym}$/⌋**czytała**bym** na dworzu.
if weather$_{NOM}$ was nice sit$_{FEM.SG}$**-by**$_{1SG}$ and read$_{FEM.SG}$**Ø**$_{by\text{-}1SG}$**/⌋**read$_{FEM.SG}$**-by**$_{1SG}$ outside
‘If the weather were nice, I would sit and read outside.’

(9) Gatsby wziął **nas** pod ramiona i wprowadził **Ø/⌋nas** do restauracji . . .
(Fitzgerald: 99).
Gatsby$_{NOM}$ took **us**$_{ACC}$ by arms and took **Ø/⌋us**$_{ACC}$ to restaurant
‘Gatsby took us by the arm and led us into the restaurant . . .’

As before, in the case of direct objects, removing phonetic identity between the antecedent and the successive coreferential category decreases the impression of excessive repetition and, in this case, makes the nonelliptical variant perfectly acceptable, as shown by the contrast between (9) and (10):

(10) Doktor Mortimer złożył **gazetę** i schował (**ją**) do kieszeni (Conan Doyle: 22).[4]
Doctor Mortimer$_{NOM}$ folded **newspaper**$_{ACC}$ and put (**it**)$_{ACC}$ in pocket
'Doctor Mortimer folded up the newspaper and put it in his pocket.'

Apart from making the antecedent and potentially elided category phonetically nonidentical, which is impossible for *się* and *by*, there are other means of offsetting the high degree of parallelism typical of VP-coordinate structures, thereby making repetition permissible and sometimes even favored. For example, when substantial distance separates the verbs in the coordinated verb phrases, which is usually caused by long or numerous adjuncts, the perceived level of interconjunct parallelism is reduced and overt repetition of elements may sound perfectly natural:

(11) Spotkali **się** w zeszłym roku gdzieś we Francji i pogodzili (**się**), a teraz są zaręczeni.
bumped-into$_{3PL}$ **się** in last year somewhere in France and made-up$_{3PL}$ (**się**) and now are$_{3PL}$ engaged
'Last year they bumped into each other somewhere in France and made up, and now they're engaged.'

(12) Gdyby nie klienci i związane z nimi sprawy, [mecenas] chętnie poszukał**by** jakiejś cichej, małej uliczki i tam (**by**) zamieszkał (Szaniawski 1: 66).
if NEG clients and connected with them matters [lawyer]$_{NOM}$ gladly look-for$_{3SG.MASC.}$-**by** some quiet small street and there (**by**) live$_{3SG.MASC}$
'If it weren't for his clients and their business, the lawyer would gladly have looked for some quiet side street and would have settled there.'

(13) [Of a rock on the beach]
Podniosła **go** z piasku, wciąż wilgotnego po niedawnej burzy, i wrzuciła (**go**) do oceanu.
picked-up$_{3SG.FEM}$ **it**$_{ACC}$ from sand still wet after recent storm and threw$_{3SG.FEM}$ (**it**)$_{ACC}$ into ocean
'She picked it up from the sand, still wet from the recent storm, and threw it into the ocean.'

Another way to counterbalance the impression of extreme parallelism and thus permit stylistically neutral repetition of elements in VP-coordinate structures is to order elements differently in each conjunct. In (14a) reflexive *się* follows its verb in each conjunct and ellipsis is preferred, whereas in (14b) *się* follows the verb in the first conjunct but precedes it in the second, making the overt-*się* variant acceptable:

(14) a. Gdy dojechali do domu, Caroline pragnęła tylko zamknąć **się** w pokoju i wypłakać **∅/ się** (Garwood: 194).

when arrived$_{3PL}$ to home Caroline$_{NOM}$ wanted only lock$_{INFIN}$ **się** in room and cry-(self)-out$_{INFIN}$ **Ø/⌋się**
'When they arrived home, Caroline wanted nothing other than to lock herself in her room and cry and cry.'

b. Gdy dojechali do domu, Caroline pragnęła tylko zamknąć **się** w pokoju i (**się**) wypłakać.
when arrived$_{3PL}$ to home Caroline$_{NOM}$ wanted only lock$_{INFIN}$ **się** in room and (**się**) cry-(self)-out$_{INFIN}$
'When they arrived home, Caroline wanted nothing other than to lock herself in her room and cry and cry.'

Example (15) provides further evidence that, even in very simple sentences, where repetition might be most keenly noticed, selecting nonparallel word order can make overt repetition of categories acceptable:

(15) Siedziało **się** i (**się**) gadało (Kupść 1999, who cites Kubiński 1987).
sat$_{3SG.NEUT}$ **się** and (**się**) chatted$_{3SG.NEUT}$
'One would sit and chat.'

In sentences that contain multiple instantiations of affixal *by*, repetition of *by* can be permitted or preferred for a reason not applicable to any other part of speech: the presence of a person-number marker on *by*. In Polish, all past tense and conditional verb forms take gender-number suffixes, but only the first and second persons take person-number markers. As explained earlier, if a verb takes the conditional marker *by*, *by* comes between the gender-number suffix and the person-number marker, if there is one. Consider the pair of coordinated *by*-conjuncts in table 12.1, which mean "I would swim and run" and "she would swim and run," respectively. The symbol (⌋) indicates that the elliptical variant of the first-person structure tends to be stylistically marked, and the nonelliptical variant of the third-person structure tends to be stylistically marked, although these judgments can be influenced by the context.

If *by* in such verb forms is elided, the person-number marker must be elided as well, if there is one. However, the combined ellipsis of [*by* + person-number marker] is not a preferred strategy in many contexts, as evidenced by the following minimal pair. In (16a), there is a person-number marker and [*by* + person-number marker] is preferably overt, whereas in (16b) there is no person-number marker and *by* is preferably elided.

(16) a. Gdybym mieszkała na wsi, pływała**bym** i biegała**bym/⌋**biegała**Ø$_{bym}$**
codziennie.
if$_{1SG}$ lived$_{FEM.SG}$ in country swim$_{FEM.SG}$**-by$_{1SG}$** and run$_{FEM.SG}$**-by$_{1SG}$/⌋**run$_{FEM.SG}$**Ø$_{by\text{-}1SG}$** daily
'If I lived in the country, I would swim and run every day.'

b. Gdyby moja siostra mieszkała na wsi, pływała**by** i biegała**Ø$_{by}$/⌋**biegała**by**
codziennie
if my sister lived$_{FEM.SG}$ in country swim$_{FEM.SG}$**-by** and run$_{FEM.SG}$**Ø$_{by}$/⌋**run$_{FEM.SG}$**-by**
daily
'If my sister lived in the country, she would swim and run every day.'

TABLE 12.1 The effect of person-number markers on the ellipsis of *by* in Polish

Person	*Pers.-Num.*	*Nonelliptical variant*	*Elliptical variant*
First	*-m*	pływała**bym** i biegała**bym**	(⌋) pływała**bym** i biegała$\varnothing_{\textbf{bym}}$
Third		(⌋) pływała**by** i biegała**by**	pływała**by** i biegała$\varnothing_{\textbf{by}}$

The obvious question is, what causes the ellipsis of [*by* + person-number marker] to be not preferred in (16a)? One possible explanation derives from an extended notion of parallelism. In the third-person pair *pływała**by*** ~ *biegała*(***by***), which has no person-number marker, when *by* is elided from the second verb form what remains is a typical third-person form. By contrast, in the first-person pair *pływała**bym*** ~ *biegała*(***bym***), which contains a person-number marker, when -(*bym*) is elided what remains is *not* a typical first-person form—in fact, it looks just like a third-person form. Therefore, there is a certain lack of parallelism created when the person-number marker is elided along with *by*. This may in part account for why the ellipsis of *by* is more prevalent in structures whose verb forms do not contain a person-number marker.

The presence of a person-number marker does not, however, always preclude the ellipsis of *by*, which leads us to the matter of ellipsis dependencies. One factor that strongly promotes the ellipsis of *by*, even if it has an attached person-number marker, is for the conjuncts to contain coreferential pronominal direct objects as well. In such instances, *by*, the person-number marker, and the direct object should all be elided, since pronouncing them all would be too repetitive and eliding a subset of them would be unbalanced:

(17) Gdybyś napisał mi list miłosny, porwała**bym go** i wyrzuciła$\varnothing_{\textbf{bym}}$ $\varnothing_{\textbf{go}}$/⌋⌊wyrzuciła-**bym go**.
if$_{\text{2SG}}$ wrote$_{\text{MASC.SG}}$ me$_{\text{DAT}}$ letter$_{\text{ACC}}$ love$_{\text{ACC.ADJ}}$ tear$_{\text{FEM.SG}}$-**by$_{\textbf{1SG}}$** **it**$_{\text{ACC}}$ and throw-away$_{\text{FEM.SG}}$$\varnothing_{\text{by-1SG}}$ $\varnothing_{\text{it}}$/⌋⌊throw-away-**by$_{\textbf{1SG}}$** **it**$_{\text{ACC}}$
'If you wrote me a love letter, I would tear it up and throw it away.'

However, there are exceptions to this "all or nothing" rule. For example, if the antecedent for the potentially elided direct object is a referential expression, ellipsis of that direct object is optional, making both variants of (18) acceptable. For ease of comparison, the (a) variant is presented twice: first in the full context, then as an excerpt that corresponds to the (b) variant:

(18) „Gdybym był radnym miasta, kupił**bym** to obrzydlistwo i spalił$\varnothing_{\textbf{bym}}$ $\varnothing_{\textbf{je}}$!" (Galsworthy: 11).
if$_{\text{1SG}}$ was$_{\text{1SG}}$ mayor$_{\text{INSTR}}$ city$_{\text{GEN}}$ buy-**by$_{\textbf{1SG}}$** this monstrosity$_{\text{ACC}}$ and burn$\varnothing_{\text{by-1SG}}$ $\varnothing_{\textbf{DO}}$
'"If I were the mayor of this city, I'd buy this monstrosity and burn it to the ground."'
a. kupił**bym** to obrzydlistwo i spalił$\varnothing_{\textbf{bym}}$ $\varnothing_{\textbf{je}}$
b. kupił**bym** to obrzydlistwo i spalił$\varnothing_{\textbf{bym}}$ **je**
buy-**by$_{\textbf{1SG}}$** this monstrosity$_{\text{ACC}}$ and burn$\varnothing_{\text{by-1SG}}$ $\varnothing_{\textbf{je}}$/**je**

Other combinations of overt and covert, however, are quite unacceptable. Splitting up conditional *by* and the person-number marker *–m* such that one was overt and the other elided would be absolutely ungrammatical, since they always function as a single entity, and eliding the direct object while having [*by* + person-number marker] overt would be similarly unacceptable.

One way of avoiding the issue of what to elide and in what combination is to place all repeated categories before the verb in the first conjunct, such that they carry over to later conjuncts. This strategy is interesting for two reasons. First, it makes the nonexpression of all the relevant elements in the second conjunct practically obligatory. Second, it erases the problematic mismatch in participial forms found in examples like (16a), which resists ellipsis because the first verb form contains a 1SG affix while the second one does not, creating a lack of parallelism between the surface forms of the verbs. When, however, the person-number marker precedes *both* participles, the surface form of the participles is identical—both lack a person-number marker. This makes structures like (19) fully acceptable:

(19) Gdybyś wykąpała psa, ja **bym go** nakarmiła i wyprowadziła na dwór.
if$_{2SG}$ wash$_{FEM.SG}$ dog$_{ACC}$ I$_{NOM}$ **by**$_{1SG}$ **it$_{ACC}$** feed$_{FEM.SG}$ and take$_{FEM.SG}$ outside
'If you had washed the dog, I would have fed and walked it.'

Such structures could be analyzed variously: (1) the verb forms could first be coordinated, then jointly select a direct object and be marked conditional; (2) both conjuncts could contain [*by* (+ person-number marker) + DO], and the preverbal clustering of these elements in the first conjunct could trigger a strong preference for ellipsis in the second conjunct; (3) such structures could display some form of across-the-boards extraction, such that there is a set of elliptical slots associated with each conjunct.

Apart from coordinate structures, another configuration that often practically requires the ellipsis of *się*, *by*, and direct objects is Repetition Structures. These combine syntactic parallelism and phonetic parallelism with lexico-semantic parallelism, realized by use of the same verb in both clauses. When the repetition is unmodified, the absolute parallelism between conjuncts makes ellipsis highly preferred, whereas when the repeated element is modified, that parallelism is reduced and ellipsis becomes optional. This contrast is shown for each part of speech by the contrast between the (a) and (b) variants of examples (20)–(22):

(20) a. „Uspokój **się**, uspokój **Ø**/⌋**się**" (Kruczkowski 1: 75).
calm-down$_{IMPER.SG}$ **się** calm$_{IMPER.SG}$ **Ø**/⌋**się**
'"Calm down, calm down."'
b. Uspokój **się**, uspokój (**się**) natychmiast.
calm$_{IMPER.SG}$ **się** calm$_{IMPER.SG}$ (**się**) at-once
'Calm down, calm down at once.'

(21) a. Gdybyś stracił ukochana, płakał**byś**, oj płakał**Ø$_{byś}$**/⌋płakał**byś**!
if$_{2SG}$ lost beloved$_{ACC}$ cry-**by**$_{2SG}$ oh cry**Ø$_{by\text{-}2SG}$**/⌋cry-**by**$_{2SG}$
'If you lost your beloved you'd cry; oh, how you'd cry!'

b. Gdyby mąż się z nią rozwiódł, umarła**by**, umarła(**by**) na pewno.
if husband$_{NOM}$ się with her$_{INSTR}$ divorced$_{MASC.SG}$ die$_{FEM.SG}$-**by** die$_{FEM.SG}$(-**by**) for sure
'If her husband divorced her, she'd die; she'd just die.'

(22) a. Powiedz coś śmiesznego, błaznie. Rozbaw **mnie**, rozbaw **Ø/⌋mnie**!
say$_{IMPER.SG}$ something$_{ACC}$ funny$_{GEN}$ jester$_{VOC}$ amuse$_{IMPER}$ **me$_{ACC}$** amuse$_{IMPER}$ **Ø/⌋me$_{ACC}$**
'Say something funny, jester. Amuse me; amuse me!'
b. Kocham **cię**, kocham (**cię**) tak bardzo!
love$_{1SG}$ **you**$_{ACC}$ love$_{1SG}$ (**you**)$_{ACC}$ so much
'I love you; I love you so much!'

When combinations of *się*, *by*, or a direct object occur in Repetition Structures, ellipsis of all coreferential elements in the second conjunct is preferred—sometimes strongly so—because having all overt would be too repetitive and eliding some but not the others would be unbalanced. In (23), which contains *się, by*, and a person-number marker, the only good variant is (a), in which all three are elided. The (b) variant, with all overt, is too repetitive, and the other variants, which show various combinations of overt and covert, are ungrammatical:

(23) Gdybyś długo nie miał od niej listu,
a. denerwował**byś się**, oj denerwował**Ø$_{byś}$ Ø$_{się}$**.
b. ⌋ denerwował**byś się**, oj denerwował**byś się**.
c. * denerwował**byś się**, oj denerwował**byś Ø$_{się}$**.
d. * denerwował**byś się**, oj denerwował**Ø$_{byś}$ się**.
if for-a-long-time NEG had$_{MASC.SG}$ from her letter$_{GEN}$ worry$_{MASC.SG}$-**by**$_{2SG}$ **się** oh worry$_{MASC}$ (**by$_{2SG}$**) **się/Ø$_{się}$**
'If he didn't get a letter from her for a long time, he'd worry; oh, how he'd worry.'

In (24), which contains *by* and a direct object, the only good variant has *by* and the direct object elided. Having both overt, as in (b), is stylistically marked, and having any combination of overt and elided is ungrammatical, as in (c) and (d):

(24) a. Kochał**by ją** do końca swego życia, naprawdę kochał**Ø$_{by}$ Ø$_{ją}$**.
b. ⌋ Kochał**by ją** do końca swego życia, naprawdę kochał**by ją**.
c. ⌋⌋ Kochał**by ją** do końca swego życia, naprawdę kochał**by Ø$_{ją}$**.
d. * Kochał**by ją** do końca swego życia, naprawdę kochał **Ø$_{by}$ ją**.
love$_{MASC.SG}$-**by her**$_{ACC}$ to end self's life really love$_{MASC.SG}$(**by**) **her/Ø$_{her}$**
'He'd love her till the end of his days, really love her!'

In sum, this Polish data shows that, when determining dependencies among elided categories, notions of overrepetition and balance are key. Both of these factors actually apply not only to the concurrent elision of categories but also to the elision of an individual category.

13

More Elliptical Phenomena

Resolving syntactic ellipsis—that is, the nonexpression of syntactically obligatory categories, which has been the focus of most of the book thus far—generally begins by seeking a coreference relation between the elided category and some linguistic referent. But establishing a *co*reference relation does not ensure resolving reference, that is, connecting the elided category to some person, thing, or event in the real or imagined world. After all, the antecedent could itself be a pronoun or some other semantically underspecified category that requires reference resolution. Moreover, the referent need not be syntactically overt; it could be understood from the real-world context. So, processing syntactic ellipsis can be described as a three-step process, with the second step either required or not, depending on the context:

Processing Syntactic Ellipsis: Preliminary Algorithm
detect syntactic gap → (establish coreference relation→) link to real-world referent

Most of the elliptical phenomena discussed so far have represented syntactic ellipsis. This chapter, in part, describes some more such phenomena that present particular challenges for machine processing.

There are, however, syntactic gaps that do not have a specific textual or contextual referent, for example, gaps referring to generalized people that function as agents, themes, or experiencers. Such gaps are unlike elliptical gaps in two ways: in some cases there is no equivalent overt-category variant, and no contextual referent need be sought. For these reasons, they are usually excluded from ellipsis studies, even

though they present similar reference resolution issues. Thus, the algorithm for resolving syntactic ellipsis is actually more complex:

Processing Syntactic Gaps: Extended Algorithm
detect syntactic gap → determine if coreference is necessary → (establish coreference relation →) link to real-world referent

Another category of phenomena that needs to be treated by a comprehensive theory of ellipsis is semantic ellipsis, since semantically elided categories are just as necessary for a full semantic representation as syntactically elided categories are for a full syntactic representation. Although with semantic ellipsis there is no syntactic gap to trigger the search for unexpressed information, other triggers can be exploited. For example, some syntactic structures are known to potentially leave arguments unexpressed (like agent-adjuncts in passive constructions). In addition, if Ontological Semantic processing is available, empty slots in instances of ontological concepts can trigger the search for their fillers. With semantic ellipsis, there may or may not be a textual referent, making the second step of the processing algorithm facultative.

Processing Semantic Ellipsis
detect missing semantic elements → (establish coreference relation →) link to real-world referent

Yet another category of elliptical phenomena that must be covered by this theory is unexpressed morphemes, which can arise from haplology, ellipsis as such, or as a side effect of the grammatical process called incorporation, none of which occur in English.

Processing Unexpressed Morphemes
detect missing morpheme → recover its content

A final class of phenomena that can productively be subsumed under a theory of ellipsis is what I will call Language Strategies, which include things like question-answer strategies and conventions for writing headlines. No matter how one chooses to classify the missing elements in such structures—as syntactic ellipsis, semantic ellipsis, or not ellipsis at all—each one requires a special set of routines, its own "rules of the game," as it were.

Processing Language Strategies
detect a given language strategy → access special processing algorithm for it

What follows is a list of classes of ellipsis along with those representative types that will be discussed in this chapter (this list is not intended to be comprehensive; it will grow along with the theory). The numbers in brackets indicate the relevant subsections.

Syntactic ellipsis with coreference [1]
- Subject ellipsis [1.1]
- Object ellipsis with an extralinguistic antecedent [1.2]
- Nonfinite clauses [1.3]

Syntactic gaps without coreference [2]
- Generalized Third-Person Plural (Indefinite Personal) Construction
- Generalized Second-Person Construction

Semantic ellipsis [3]
- Unexpressed agents in passives [3.1]
- Agentive impersonals [3.2]
- Unexpressed experiencers and possessors [3.3]
- Unexpressed arguments in derived nominals and embedded verb forms [3.4]

Unexpressed morphemes [4]
- Haplology
- Morpheme Ellipsis
- Morpheme loss during incorporation

Language Strategies[1] [5]
- Dialogue strategies [5.1]
- Sentence fragments [5.2]
- Nominal sentences [5.3]
- Unagentive impersonals [5.4]

In the following, I briefly describe each phenomenon, posit a preliminary inventory of parameters and values for describing it cross-linguistically, note salient processing issues, and build a bridge between ellipsis and the larger topic of reference resolution.

1. Syntactic ellipsis with coreference: More phenomena

Most of the types of syntactic ellipsis discussed in previous chapters readily lent themselves to description by a combination of syntactic, lexico-semantic, and Ontological Semantic rules and routines, which are much easier to formulate than rules intended to capture pragmatic (discourse) factors. The latter were called upon only as a last resort, to permit a maximally comprehensive description of ellipsis irrespective of the intended application. However, the resolution of some prevalent types of syntactic ellipsis requires significant reliance on an understanding of the discourse context. How the discourse is represented and what a given system can know about it in large part determines the types and interactions of rules used to resolve such ellipsis.

1.1. Subject ellipsis

Subject ellipsis, well known and widely studied, is found in language groups including but not limited to Romance (e.g., Italian), Slavic (e.g., Russian, Polish, Czech),

Sino-Tibetan (e.g., Chinese), and sign languages (e.g., American Sign Language). Here I use the traditional definition of subject ellipsis, which puts English among the languages that do not permit it (that is, the missing subjects in highly colloquial *Got milk?* and in the second conjunct of conjoined predicates like *The dog grabbed the shoe and ran out of the house* do not count).

In languages that have inflectional verbal morphology, the inflectional form of the matrix verb usually provides clues for the subject's referent but often not an unambiguous indication. For example, whereas in (1) the verbal morphology restricts the referent to 1SG (an unambiguous referent, if we know who the speaker is), in (2) the verbal morphology restricts the referent to 3PL, but the preceding linguistic context must provide its referent:

(1) ∅ Страшно хочу есть.
∅ Strašno xoču est'.
∅ terribly $want_{1SG}$ eat_{INFIN}
'I'm dying for something to eat.'

(2) «В нашей, говорит, семье **эти игрушки** живут лет, наверное, девяносто. ∅ Переходят от матери к дочке» (Шварц 5: 467).
«V našej, govorit, sem'e **èti igruški** živut let, navernoe, devjanosto. ∅ Perexodjat ot materi k dočke» (Švarc 5: 467).
in our $says_{3SG}$ family **these$_{NOM}$ toys$_{NOM}$** $live_{3PL}$ years probably ninety ∅ $pass_{3PL}$ from mother to daughter
'"These toys," he says "have been in our family for what must be a good ninety years. They get passed down from mother to daughter."'

In languages that permit subject ellipsis, overtly repeating a subject too many times causes stylistic infelicity. However, how much repetition is too much varies from language to language. For example, although both Polish and Russian widely employ subject ellipsis, Polish requires it in more contexts than Russian—which might be a reflex of its richer verbal morphology in certain inflectional forms.[2]

This preliminary description of subject ellipsis suggests a number of parameters and values for describing it cross-linguistically, listed in table 13.1.

This inventory covers the easily discernible aspects of subject ellipsis but does not scratch the surface of what one needs to know to write rules for its resolution or generation. What is needed is heuristics for when and when not to elide subjects and for how to find referents for subjects that are not unambiguously recoverable by verbal morphology. This task, it turns out, is trickier for subjects than for objects because there are fewer configurations in which the placement of the antecedent—and its role as antecedent—can be predicted with confidence. In fact, this is as much a matter of reference resolution as of coreference resolution.

Much work has been done cross-linguistically on subject ellipsis, but most of it is either (1) not intended for practical applications, for example, theoretical syntactic treatments, for which usage rules are less important than the possibility or impossibility of the phenomenon to begin with; (2) not suitable for practical applications, for example, discourse-theory treatments, in which the rules tend to be quite abstract;

TABLE 13.1 A short inventory of parameters and values to describe subject ellipsis

Parameters	*Values*
Is subject ellipsis employed?	Yes No
What can the referents be?	Speaker Hearer Third person (person or thing)
Does verbal morphology provide clues for recovering the elided subject?	Yes No
(If yes above . . .) Which subject features are reflected by verbal morphology? (possibly different for different tense/mood, etc., combinations)	Gender Person Number Etc.
What can an elided subject refer to?	Concrete person or thing Generalized person or thing
What can the status of the referent be?	Syntactically accessible Understood extralinguistically Understood by world knowledge
How "necessary" is subject ellipsis when the subject can be recovered by morphological or contextual means?	Virtually obligatory Very common Optional

or (3) not generalizable to anything but a given application, for example, knowledge-poor, statistical treatments for the machine processing of some specific language.[3] It is likely that, for any given application, a combination of well-specified selectional restrictions, heuristics, and statistical methods will best handle subject ellipsis.

1.2. Object ellipsis with an extralinguistic antecedent

Object ellipsis with an extralinguistic antecedent presents the same scope of challenges as subject ellipsis and can be described using the same inventory of parameters and values (barring, of course, the status of the antecedent, since the delineation of this class of phenomena takes care of that: object ellipsis *with an extralinguistic antecedent*). For languages that show object agreement morphology on the verb, the resolution challenge is precisely the same as for elided subjects. But there are languages, like Russian, that permit object ellipsis despite the fact that objects do not agree with their selecting verbs. This means that the only clues for the referent of that object are the verb's selectional restrictions and an understanding of the overall context. In the following Russian examples, the notes in brackets describe the salient contextual information:

(3) [The king is upset]
«Вон! Все пошли вон! Расстроили ∅! Обидели ∅! Всех переколю! Заточу! Стерилизую! Вон!» (Шварц 1: 104).

«Von! Vse pošli von! Rasstroili ∅! Obideli ∅! Vsex perekolju! Zatoču! Sterilizuju! Von!» (Švarc 1: 104).
out everyone$_{NOM}$ go out upset$_{3PL}$ ∅$_{\mathbf{ACC}}$ hurt-feelings$_{PL}$ ∅$_{\mathbf{ACC}}$ everyone$_{ACC}$ will-massacre$_{1SG}$ will-incarcerate$_{1SG}$ will-sterilize$_{1SG}$ out
‘‘“Get out of here! Everyone—out of here! You’ve upset me! You’ve hurt my feelings! I’ll massacre you all! I’ll incarcerate you! I’ll sterilize you! Out with you!”’

(4) **Сильва** (*задерживает вторую девушку*). Дай хоть поцелую ∅ на сон грядущий (Вампилов 2: 72).
Sil’va (*zaderživaet vtoruju devušku*). Daj xot’ poceluju ∅ na son grjaduščij (Vampilov 2: 72).
Sil’va (*holds-back second girl*) let at-least kiss$_{1SG.PL}$ ∅$_{\mathbf{ACC}}$ [idiom: at bedtime]
Sil’va (*holding back the second girl*). Let me at least kiss you good night!

(5) [Having seen someone drop money]
«Стойте-стойте! Пан потерял ∅!» (Хмелевская 4: 150).
«Stojte-stojte! Pan poterjal ∅!» (Chmielewska 4: 150).
stop$_{IMPER}$ stop$_{IMPER}$ sir$_{NOM}$ dropped$_{3SG}$ ∅$_{\mathbf{ACC}}$
‘‘“Stop, stop! You dropped something!”’

Some generalizations about elided objects with extralinguistic antecedents can be made that might help in ultimately developing resolution routines. The following, although derived so far only from Russian, coincide with more general linguistic principles and therefore might be expected to apply to other languages as well:

- Object ellipsis with an extralinguistic antecedent tends to occur much more frequently in colloquial dialogues than in any other speech register, so it should be a strong candidate analysis for syntactic gaps only in such registers; in other registers, it should be a fallback if no other analysis is viable.
- When the unexpressed object must be human, based on the verb’s selectional restrictions, the elided object is very often the speaker or the hearer—which reflects the discourse property that the interlocutors have a privileged status in any speech event.[4]
- There appears to be some correlation between certain verbs used in certain forms and objects with certain types of referents. For example, when *dat’* ‘give’ is used in the imperative and the indirect object is elided, it tends to have the referent ‘me’: *Daj knigu* ‘Give (me) the book’. Similarly, when the verb *ponravit’sja* ‘please’ is used in the past tense in an interrogative sentence and the indirect object is elided, the referent tends to be ‘you’: *Ponravilsja film?* ‘Did (you) like the film?’ If such tendencies hold up under larger corpus analysis and if they can be discovered for at least the verbs most commonly used with an unexpressed extralinguistic antecedent, they could serve as heuristics for positing candidate referents for objects.

Whereas resolving subject ellipsis must be on the agenda of any practical system intended to process a language that uses it, resolving object ellipsis with an extralinguistic antecedent need not necessarily be. Since this type of ellipsis is most prevalent in colloquial dialogues, those text-processing systems that target more formal language need not consider it a priority. In the event that such ellipsis does occur, the usual type of stopgap could be implemented, like accepting the syntactic structure as valid, then using selectional restrictions and other heuristics to posit one or more possible referents, and finally disambiguating later if knowledge permits.

1.3. Nonfinite clauses

In languages with verbal inflection, it is typical for each clause to contain an inflected verb form, like *swims* or *dived*. When nonfinite verb forms are used, they are most often the complement of an inflected verb form: *I like **to swim*** and *I plan **to dive** off the high board*. Sometimes, however, nonfinite verb forms can be used independently, as in ***To be** or not **to be**, that is the question.* Apart from Shakespeare, however, independent nonfinite clauses are practically never used in English.[5] Not so, however, in some other languages. Russian, for example, widely permits clauses whose sole verb is in the infinitive. But if the expected subject (often an agent) is not expressed, what does it refer to and how can it be recovered? And why would one use a nonfinite clause to begin with?

In Russian, nonfinite clauses imply modality—what *kind* of modality depends upon the larger context, as shown by (6)–(11).[6] As for the agent, it can have a specific overt referent (realized as NP_{DAT}), a specific elided referent, a generalized human referent, or an imprecise combination of specific and nonspecific referents.

Imperative

(6) [The implied agent is the addressee, a dog]
Сидеть!
Sidet'!
sit$_{INFIN}$
'Sit!'

Shall

(7) **Налить** тебе кофе?
Nalit' tebe kofe?
pour$_{INFIN}$ you$_{DAT}$ coffee?
'Shall I pour you some coffee?'

Should

(8) «Может быть, **показать** его психиатру?» (Войнович 2: 80).
«Možet byt', **pokazat'** ego psixiatru?» (Vojnovič 2: 80).
may be **show$_{INFIN}$** him psychiatrist$_{DAT}$
'"Maybe we should take him to <he should be seen by> a psychiatrist?"'

It would be nice

(9) «Больше всего, Саня, я любил ночные привалы. Пылает огонь, трещит хворост, и искры уносятся в синюю тьму. Сейчас бы **пойти** в поход» (Войнович 1: 56).
«Bol'še vsego, Sanja, ja ljubil nočnye privaly. Pylaet ogon', treščit xvorost, i iskry unosjatsja v sinjuju t'mu. Sejčas by **pojti** v poxod» (Vojnovič 1: 56).
{ . . . } and sparks fly-away into blue darkness now CONDIT **go$_{INFIN}$** on hike
'"More than anything, Sanya, I liked nighttime stops. The fire blazes, the brushwood crackles, and sparks fly off into the blue darkness. It'd be great [or: Oh, how I'd love] to go hiking right now."'

It would be impossible

(10) Как раз через два дня у Томми и Анники начинались летние каникулы, а их мама и папа должны были на несколько дней уехать. Короче — лучшего случая поиграть в робинзонов **не найти** (Линдгрен: 122).
Kak raz čerez dva dnja u Tommi i Anniki načinalis' letnie kanikuly, a ix mama i papa dolžny byli na neskol'ko dnej uexat'. Koroče — lučšego slučaja poigrat' v robinzonov **ne najti** (Lindgren: 122).
{ . . . } in short a better$_{GEN}$ opportunity$_{GEN}$ play$_{INFIN}$ at Robinson-Crusoe NEG **find$_{INFIN}$**
'In just two days Tommy and Annika's summer vacation would start and their mother and father were going away for a few days. In short, you'd never find a better opportunity to play Robinson Crusoe.'

Must

(11) «Пойдём, что ли? Завтра рано **вставать**» (Войнович 1: 35).
«Pojdëm, čto li? Zavtra rano **vstavat'**» (Vojnovič 1: 35).
let's-go PARTICLE$_{INTERROG}$ tomorrow early **get-up$_{INFIN}$**
'"Shall we go? We've got to get up early in the morning."'

Given sufficient examples, heuristics for guessing the modality of infinitival clauses can be developed. Some worth testing on a Russian corpus are:

- If an infinitival clause ends in an exclamation point, the modality is imperative.
- If an infinitival clause ends in a question mark, the modality is *shall?*
- If an infinitival clause is modified by an adverb like *maybe*, *perhaps*, the modality is *should*.
- If an infinitival clause has a conditional particle, the modality is *would be nice*.
- If an infinitival clause is negated and the object, if any, is in the Genitive (of negation) case, the modality is *would be impossible*.
- If none of the preceding conditions applies, the modality is (tentatively) *must*.

The more difficult problem, however, is recovering the experiencer if it is unexpressed—a topic taken up in section 3.3. The parameters and values for describing nonfinite clauses include those listed in table 13.2.

2. Syntactic gaps without coreference

Humans have a special status in language because of its anthropocentric nature. Generalized or nonspecific humans have yet a different status, which can have various grammatical realizations.[7] In English, for example, *they*, *you*, and *one* can all be used with generalized meaning, as in ***They*** *say it will rain tomorrow*, ***You*** *can never trust the forecast*, and ***One*** *can never be too careful!* Nonlexical methods, like passivization (*My car's been stolen!*) and the use of nonfinite verbal forms (*Painting the house would be a good idea*), also permit the expression of generalized or nonspecific human agents. In all of these structures, there are no gaps in the syntax, but a complete semantic representation requires specification of the agent as generalized or unspecified. What a human knows and a machine must be taught is that in such contexts no *co*reference relations need be established; instead, a reference link to generalized/nonspecific people must be established.

Similar analysis is required for languages in which references to generalized/nonspecific people can or must be elliptical. In Russian, for example, the 3PL subject pronoun in sentences like *They say it will rain tomorrow* must be elided in order

TABLE 13.2 Parameters and values used to describe nonfinite clauses

Parameters	*Values*
Can the main verb in a clause be nonfinite?	Yes No
How can the experiencer be expressed?	As an NP (with some case marking) As a PP Using ellipsis
What can the referent for the experiencer be?	Some specific person(s) or animal(s) Generalized people A hybrid of specific and generalized people
What semantic implications does the clause have?	Modality None Other
(If modality) What types of modality can be expressed?	Shall Must Should It would be nice Etc.
(For each type of modality) What clause features suggest this type of modality?	Exclamation point Question mark Conditional particle Etc.

for the generalized meaning to be conveyed (*Говорят, что завтра будет дождь / Govorjat, čto zavtra budet dožd'*). This is called the Indefinite Personal Construction, and it precludes the overt realization of the subject. If the subject were overt, it would necessarily imply the existence of a textually or contextually determined specific referent. So, whereas an overt 3PL pronoun unambiguously indicates that a coreference relation must be found, an elided 3PL pronoun leaves open the same two options as overt *they* in English—either there is a textually bound referent or generalized/nonspecific people are implied. The following examples show the correlation between the Russian Indefinite Personal Construction and some of its English counterparts, with the distribution of variants in English being tangential:

(12) Из его собственного гаража выезжала его собственная «Волга»! «**Угоняют** машину!» — беспомощно закричал Филипп (Брагинский и Рязанов: 26).
Iz ego sobstvennogo garaža vyezžala ego sobstvennaja «Volga»! «**Ugonjajut** mašinu!» — bespomoščno zakričal Filipp (Braginskij i Rjazanov: 26).
from his own garage$_{GEN}$ pulled-out$_{FEM.SG}$ his own Volga$_{FEM.SG}$ **drive-away$_{3PL}$** car$_{ACC}$ helplessly cried Philip
'His very own Volga was pulling out of his very own garage! "My car's **being stolen**!" cried Philip helplessly.'

(13) Увидев голого человека, милиционер с нескрываемым любопытством высунулся из своего стакана и сочувственно спросил: «Вас **раздели**?» — «У меня **угнали** машину!» — «И **раздели**?» — «Нет, я сам!..» (Брагинский и Рязанов: 28–29).
Uvidev gologo čeloveka, milicioner s neskryvaemym ljubopytstvom vysunulsja iz svoego stakana I sočuvstvenno sprosil: «Vas **razdeli**?» — «U menja ugnali mašinu!» — «I **раздели**?» — «Net, ja sam!..» (Braginskij i Rjazanov: 28–29).
{ . . . } you$_{ACC}$ **undressed$_{3PL}$** at me **drove-away$_{3PL}$** car$_{ACC}$ and **undressed$_{3PL}$** no I$_{NOM}$ myself$_{NOM}$
'At the sight of a naked man, the policeman, with unconcealed curiosity, leaned out of his booth and asked sympathetically: "Has someone undressed you?" "No, they've stolen my car!" "And undressed you?" "No, I did that myself!"'

In some instances the interpretation of the generalized humans can be limited by contextual clues; for example, in the sentence *In Paris, **they** drink a lot of wine*, the drinkers are not all of humanity but people living in or perhaps visiting Paris.[8] Similarly, in (14) the eaters must be cannibals—certainly *not* all of humanity!

(14) «Ее муж был миссионером. Жуткая история. **Его *съели***, представляете себе? Буквально **съели**» (Кристи: 20).
«Ee muž byl missionerom. Žutkaja istorija. **Ego *s''eli***, predstavljaete sebe? Bukval'no ***s''eli***» (Christie: 20).
her husband$_{NOM}$ was missionary$_{INSTR}$ horrific story$_{NOM}$ **him$_{ACC}$ ate$_{3PL}$** imagine$_{IMPER}$ self$_{DAT}$ literally **ate$_{3PL}$**
'"Her husband was a missionary. A horrific story. They ate him, can you imagine? They actually ate him."'

Finally, while the verb form must be plural, the implied agent need not refer to more than one person:

(15) Меня **прервали** ещё раз. На этот раз мой помощник, Хоуз (Кристи: 16).
Menja **prervali** eščë raz. Na ètot raz moj pomoščnik, Xouz (Christie: 16).
me$_{ACC}$ **interrupted$_{3PL}$** again PREPOSITION this time my assistant$_{NOM}$ Howes$_{NOM}$
'I was interrupted again, this time by my assistant, Howes.'

Another common realization of generalized/nonspecific people is an overt or elided second-person pronoun, as illustrated in the Russian and English variants of (16):

(16) «Когда **хочешь** очень хорошо говорить, то, как нарочно, мысли путаются в голове и все нужные слова разбегаются» (Шварц 4: 172).
«Kogda **xočeš'** očen' xorošo govorit', to, kak naročno, mysli putajutsja v golove i vse nužnye slova razbegajutsja» (Švarc 4: 172).
when **want$_{2SG}$** very-much well speak$_{INFIN}$ then as-if intentionally thoughts$_{NOM}$ get-mixed-up in head and all$_{NOM}$ necessary$_{NOM}$ words$_{NOM}$ scatter
'"When you want to speak really well, it always happens—as if to spite you—that your thoughts get mixed up in your head and the words you need just won't come."'

Using *you* in this manner not only makes a generalization; it also serves to draw the listener into the generalization. Whereas in English the pronoun must be overt, in Russian it may be overt or elided and singular or plural, though singular is more common for both the overt and elliptical variants.

The Generalized Second-Person Construction is commonly found in proverbs, like "As you sow, so shall you reap" (Russian: *Что посеешь, то и пожнёшь*) and "You can't please everyone" (Polish: *Wszystkim ne dogodzisz*). However, its most widespread usage is as a productive process, applicable to any verb and posing the familiar challenge for resolution: determining whether or not to seek/use a contextual referent or to assume generalized meaning.

The expression of generalized meanings can be viewed as a language parameter, one of whose values is the expression of generalized/nonspecific people. Possible realizations of the latter include:

- overt 3PL pronoun (*they* in English),
- obligatorily elided 3PL pronoun (Indefinite Personal Construction in Russian),
- overt second-person pronoun (*you* in English; *ty/vy* in Russian),
- elided second-person pronoun (in Russian, with singular or plural verbal agreement),
- passivization with an unexpressed agent-adjunct (English and Russian), and
- overt *one* (important in English because it is ambiguous with 'one' as a number).

3. Semantic ellipsis

Semantic ellipsis covers many phenomena, each of whose processing requires different knowledge sources and resolution routines. For example, all lexical collocations that imply some additional information must be listed as patterns with variables, and the implied semantics must be encoded in the associated semantic representation—supplemented, if necessary, by contextual information. For example, in *I forgot my keys*, the lexical item *forget* taking a PHYSICAL-OBJECT complement should trigger instantiation of a concept like TAKE or BRING. Similarly, when a lexical item presupposes some contextual knowledge, that contextual knowledge must be recovered. For example, in *Jim is a sophomore*, a full understanding of *sophomore* requires knowledge of whether Jim is in high school or college. Such lexically triggered types of semantic ellipsis must be prepared for in case-by-case or group-by-group fashion: for example, the same algorithm used for disambiguating *sophomore* could be used for *freshman*, *junior*, and *senior*. (The treatment of these and other types of semantic ellipsis in the Ontological Semantics environment is described in McShane, Nirenburg, and Beale 2004.)

The following subsections focus on nonlexical types of semantic ellipsis—namely, the nonexpression of "players" in the clause. These cannot be called arguments because they are not syntactically required, but they represent case roles like agent and experiencer, which are crucial bits of information in the semantic representation of any event. One issue that runs throughout and whose solution will not be attempted here is how to disambiguate between the potential types of referents for unexpressed agents and experiencers, be they context specific, completely generalized, or restricted to some implied group.

3.1. Unexpressed agents in passives

The passive voice is used to defocus the agent of an action, even if that agent is overtly expressed as the so-called agent-adjunct (*The shed was painted by the gardener*).[9] When the agent is not expressed, its implied referent can be range from very specific to very generalized, as illustrated by the following examples:

- If someone says to a mother who is breast-feeding her child *The baby should be fed soon*, the implied agent is that particular mother.
- If someone says to a group of breast-feeding mothers *Your babies should be fed at regular intervals*, the implied agents are each mother for her child.
- If someone asks a home owner *Will your house be repainted soon?* the implied agent is that home owner or someone acting on his or her behalf.
- If someone asks *Are houses in America repainted every year?* the implied agent is each respective home owner or someone acting on his or her behalf.
- If someone says *Many languages are spoken in the world*, the implied agent is all of humanity.

- If someone asks *Was this play well received?* the implied agent is the subset of people who saw the play.

In short, a significant amount of world knowledge and case-by-case interpretation is required to specify or constrain to the smallest possible subset the interpretation of unexpressed agents in passive constructions.

The parameters and values for the description of passive constructions include the means of realizing the agent-adjunct (there may be more than one in a given language), whether impersonal passives exist (they do in Ukrainian, for example), the verb form(s) used in passive constructions (often a participle used with or without an auxiliary), and any agreement patterns between the subjects of passive constructions and the verb forms.[10]

3.2. Agentive impersonals

Impersonal clauses have no syntactic subject. In many cases, they lack an agent as well, in which case there is no semantic ellipsis (cf. section 5.4). In other cases, however, the absence of a syntactic subject does not make the clause semantically unagentive: there is either (1) a specific agent overtly expressed as something other than the syntactic subject, (2) a specific agent implied by the context, or (3) a generalized-human agent—the same semantic state of affairs as for unexpressed agents in passive constructions.

English does not have agentive impersonals, but Polish does—several types, in fact. One common strategy for creating impersonals from fundamentally personal verbs is to add the particle *się* to the 3SG (Neuter) form. Impersonals thus formed can add two different semantic nuances, depending on the context. The first is 'ability', in which case the agent may be specific (17) or generalized (18). If it is specific and overt, it is realized as a DAT-case noun phrase. The second semantic nuance is 'applicable to all people at all times', in which case the agent is always unexpressed (19):

(17) a. Joannie **tańczy się** bardzo dobrze (Topolińska, n.d.: 141).
Joanna$_{DAT}$ **dances$_{3SG}$ się** very well
'Joanna dances very well.'
b. Tutaj mi **się będzie** dobrze **pracowało** (Kaleta 1995: 328).
here me$_{DAT}$ **się will$_{3SG}$** well **work$_{3SG.NEUT}$**
'I'll be able to work well here.'

(18) a. Dobrze **się** tu **pracjue** (Topolińska, n.d.: 149).
well **się** here **work$_{3SG}$**
'One can work well here.'
b. Po pracy **śpi się** doskonale (Topolińska n.d.: 149).
after work **sleep$_{3SG}$ się** great
'After work one sleeps great.'

(19) a. Przyjemnie **podróżjue się** po Francji (Topolińska n.d.: 149).
nice **travel$_{3SG}$ się** through France
'It's nice to travel through France.'

b. Nie **myśli się** o tym, że wkrótce będzie zima (Kaleta 1995: 328).
not **think$_{3SG}$ się** about that$_{PREP}$ that soon will-be winter$_{NOM}$
'One doesn't think about the fact that winter will soon be here.'

Another Polish strategy for creating impersonals out of personal verbs is to use the *–no* or *–to* participial form as the main verb in the clause. The referent may be completely unspecified people (20) or some contextually implied type or group of people (21). The time frame is always past.

(20) a. U sąsiadów nad nami **tańczono** i **śpiewano** (Topolińska n.d.: 150).
at neighbors' above us **danced$_{PARTICIPLE}$** and **sang$_{PARTICIPLE}$**
'There was singing and dancing going on in the apartment of our upstairs neighbors.'
b. **Mówiono** wielo o tym wydarzeniu (Kaleta 1995: 327).
spoken$_{PARTICIPLE}$ much about that happening
'That happening was talked about a lot.'

(21) a. Było późno i na seans już nas nie **wpuszczono** (Topolińska n.d.: 150).
was late and to show already us NEG **let-in$_{PARTICIPLE}$**
'It was late and they (whoever had the power to do so) didn't let us into the show.'
b. Dawniej **robiono** meble z drzewa (Kaleta 1995: 328).
in-the-pat **made$_{PARTICIPLE}$** furniture from wood
'They (whoever knew how to work with wood) used to make furniture out of wood.'

A complete description of agentive impersonal clauses in a language would include descriptors like the ones in table 13.3, which is a summary of the state of affairs in Polish.

In addition, the description of impersonal constructions should include:

1. Semantic Restrictions. For example, one cannot use an ability-impersonal for the verb *play* in Polish, as in **Joannie bawi się dobrze* (Topolińska n.d.: 141), meaning "Joanna can play (games) well."
2. Contrasts in Usage with Other Constructions. For example, whereas the Indefinite Personal Construction in Polish can include the speaker as an agent, the 3SG + *się* construction cannot (Topolińska n.d.: 150–151).

TABLE 13.3 Descriptors for impersonal sentences in Polish

Referent of agent	*Syntactic status of agent*	*Impersonalization strategy*	*Semantic nuances of construction*	*Example number*
Specific referent	NP_{DAT}	3SG verb + *się*	Ability	17
Generalized person	Unexpressed	3SG verb + *się*	Ability	18
Generalized person	Unexpressed	3SG verb + *się*	For all people at all times	19
Generalized person	Unexpressed	*-no/-to* participle	Past tense	20
Generalized person contextually constrained	Unexpressed	*-no/-to* participle	Past tense	21

3. Comparisons in Usage with Other Constructions. For example, the neuter past + *się* can be used with the same meaning as the *-no/-to* constructions, as shown in (22a–b) (Kaleta 1995: 328):

(22) a. Do tego domu **wchodzono** z drugiej strony.
to this house **go-in**$_{\text{PARTICIPLE}}$ from other side
'One entered this house from the other side.'
b. Do tego domu **wchodziło się** z drugiej strony.
to this house **go-in**$_{\text{3SG.NEUT.PAST}}$ **się** from other side
'One entered this house from the other side.'

3.3. Unexpressed experiencers and possessors

One of the most difficult aspects of semantic ellipsis is determining where it occurs. Obviously, on the one hand, it would serve no purpose to say that all the encyclopedic and world knowledge that might have relevance for each component of the sentence be considered semantically elided. On the other hand, especially if we look at cross-linguistic means of expressing things, there *are* elements that can regularly be left out when understood in one language but not in others. The scope of this phenomenon is most easily shown by examples.

Body parts, unless referred to in general or medical terms, must belong to a specific person or animal. In English, the person must be specified (*I broke **my** leg*), but in other languages, not always. For example, body-part possessors in Russian and Polish are generally unexpressed if they can be recovered either by syntactic means (the missing possessor of *nose* in [23] must be Masha) or by extralinguistics means (the missing possessor of *nose* in [24] could be anyone understood from the context).

(23) У Маши царапина на щеке.
U Maši carapina na ščeke.
at Masha scratch on cheek
'Masha has a scratch on her cheek.'

(24) Мама с бабушкой говорили о царапина на щеке.
Mama s babuškoj govorili o carapine na ščeke.
Mom and Grandma talked about scratch on cheek
'Mom and Grandma were talking about the scratch on *someone's* cheek.
[need context to recover who *someone* is]

Apart from body parts, many other possessed things need context-specific semantic linkings. For example, in Russian sentence (25) the room is the child's (*decorate his room*), in (26) the life is Diana's (*first time in her life*), and in (27) the desire is the speaker's:

(25) Меня очень привлекало в американках отношение к маленьким детям, желание доставить ребенку радость: везти его в Диснейленд, украшать **комнату**, мчаться за игрушками (Богуславская: 266).

Menja očen' privlekalo v amerikankax otnošenie k malen'kim detjam, želanie dostavit' rebenku radost': vesti ego v Disnejlend, **ukrašat' komnatu**, mčat'sja za igruškami (Boguslavskaja: 266).
{ . . . } take$_{INFIN}$ him to Disneyland **decorate$_{INFIN}$ room** run-around$_{INFIN}$ for toys
'As regards American women's relationship to children, I was really drawn by their desire to make the child happy: taking him to Disneyland, **decorating his room** [lit.: decorating room], running around buying him toys.'

(26) Он отвез ее домой рано утром. Мистер Нилл звонил, но Дайана солгала — первый раз в жизни, — что она спит (Минчин. 66).
On otvez ee domoj rano utrom. Mister Nill zvonil, no Dajana solgala — pervyj raz **v žizni**, — čto ona spit (Minčin: 66).
he$_{NOM}$ drove her$_{ACC}$ home early in-the-morning Mr. Neal$_{NOM}$ called but Diana$_{NOM}$ lied first time **in life** that she$_{NOM}$ sleeps
'He drove her home early in the morning. Mr. Neal had called, but Diana had lied for the first time in her life—she had said she was sleeping.'

(27) [The author describes with despair her own struggles with her grandmother.]
Испытание—это знак даже некоторой избранности. Вот ведь! Все еще **охота** в избранные. Стыдно (Рачко: 99).
Ispytanie—eto znak daže nekotoroj izbrannosti. Vot ved'! Vse ešče **oxota** v izbrannye. Stydno (Račko: 99).
ordeal$_{NOM}$ it-is sign$_{NOM}$ even some$_{GEN}$ chosenness PARTICLE after-all still **desire$_{NOM}$** to chosen$_{NOM.PL}$ shameful$_{IMPERS}$
'An ordeal is even a sign of being chosen, in a way. There, you see! I still have a desire to be among the chosen. It's shameful.'

In all of these examples, an extended notion of possessor covers the type of element that needs to be recovered. Another large category of elements that need to be recovered to create a full semantic representation is experiencers—those humans or animals (and perhaps even plants) who experience a physical or psychological state. Detecting a missing experiencer is not trivial, as evident from analysis of the variants of (28):

(28) a. Xolodno.
cold$_{IMPERS}$
'It's cold <I'm cold>.'
b. Zdes' xolodno.
here cold$_{IMPERS}$
'It's cold here.'
c. Mne xolodno.
me$_{DAT}$ cold$_{IMPERS}$
'I'm cold.'
d. Mne zdes' xolodno.
me$_{DAT}$ here cold$_{IMPERS}$
'I'm cold here <It's too cold for me here>.'

Saying that *it is cold* expresses a real or potential judgment or experience. However, whether that judgment or experience is attributed to a specific person, a class of people, or generalized people or animals (even plants) is often less clear, leading to a special set of problems. In English, for example, the seemingly complete sentence *It's cold* generally reflects the typical human perspective, but if one was narrating a piece about Eskimos the standard of judgment would be different, and even more so for animals living at the South Pole. Therefore, any inferencing that might be done on the basis of such a statement must take into consideration the understood experiencer. Moreover, one can walk into a room where everyone is sitting comfortably in short sleeves and say *It's cold in here!*, reflecting a purely subjective state. A human-oriented semantic analysis of experiencers would include at least the subtypes listed here. How this level of analysis might be achieved automatically remains a research issue.

Specific experiencer recoverable syntactically

(29) *Тригорин.* Когда пишу, приятно. И корректуру читать приятно . . . (Чехов 2: 415).
Trigorin. Kogda pišu, prijatno. I korrekturu čitat' prijatno . . . (Čexov 2: 415).
Trigorin. when write$_{1SG}$ nice$_{IMPERS}$ and proofs$_{ACC}$ read$_{INFIN}$ nice$_{IMPERS}$
Trigorin. 'When I write, it's nice <I enjoy myself>. And reading proofs is nice . . .'

Definite experiencer recoverable extralinguistically

(30) Что-то не **спится**: здесь слишком шумно.
Čto-to ne **spitsja**: zdes' sliškom šumno.
somehow NEG **sleeps$_{3SG}$** here too noisy
'Somehow I can't sleep: it's too noisy in here.'

Definite referent extended to anyone else in his or her shoes

(31) Экзюпери очень мудрый писатель. Даже страшно становится, до чего мудрый (Золотые: 88).
Ekzjuperi očen' mudryj pisatel'. Daže **strašno stanovitsja**, do čego mudryj (Zolotye: 88).
Exupery$_{NOM}$ very wise writer$_{NOM}$ even **scary it-becomes** to what exent wise
'Exupery is a very wise writer. It's even scary how wise.'

Generalized experiencer constrained by place, membership, participation, and so on

(32) На концерте было очень приятно.
Na koncerte bylo očen' prijatno.
at concert was very nice
'It was very nice at the concert.'

Fully generalized experiencer

(33) Надо есть, а то умрешь.
Nado est', a to umreš'.

necessary$_{IMPERS}$ eat$_{INFIN}$ otherwise will-die$_{2SG}$
'You have to eat or else you'll die.'

Intended ambiguity can add semantic richness to an utterance. For example, in (34) Medvedenko is trying to convince his wife, Masha, to go home with him to see their daughter, who misses them. In saying *žalko* '(it's a)pity$_{IMPERS}$' he is expressing not only that *he* feels sorry for her but also that it's a sad situation that she's home alone. In addition, he avoids saying *mne žalko* 'I feel sorry for her' because he doesn't want to offend his wife by making the implied comparison *I* do but *you* obviously don't:

(34) *Медведенко* (умоляюще). Маша, поедем! Наш ребеночек небось голоден.
Маша. Пустяки. Его Матрена покормит.
Медведенко. **Жалко**. Уже третью ночь без матери (Чехов 2: act 4).
Medvedenko (umoljajušče). Maša, poedem! Naš' rebenoček nebos' goloden.
Maša. Pustjaki. Ego Matrena pokormit.
Medvedenko. **Žalko**. Uže tret'ju noč' bez materi. (Čexov 2, act 4).
{ . . . } **a-pity$_{IMPERS}$** already third night without mother
Medvedenko (*emploringly*). Masha, let's go! Our child must be hungry.
Masha. That's no problem. Matryona will feed him.
Medvedenko. It's a pity. This is already the third night without his/her mother.

Other semantic nuances can also be conveyed by unexpressed possessors and experiencers. Wierzbicka points out that there are crucial semantic distinctions between Polish sentences in which possessors of body parts and clothing are and are not expressed overtly. She introduces the concept "simple bodily action," explaining that this category "comprises both actions involving a part of the agent's body and simple routine actions involving objects worn on the body, and thus indirectly involving a part of the body" (1979: 338). Among "simple bodily actions" are opening one's eyes, brushing one's teeth, putting on one's gloves, taking off one's hat, and so on. When these actions are carried out in the normal fashion, the DAT possessor is not indicated and no modifiers may be added.[11] (The notation (*) indicates that realizing the category overtly is ungrammatical.)

(35) Piotr otworzył (*niebieskie) oczy (Wierzbicka 1979: 338).
Piotr opened (*blue) eyes

(36) Piotr zdjął (*czarny) kapelusz (Wierzbicka 1979: 338).
Piotr took off (*black) hat

However, when such actions are *not* carried out in the typical fashion or when a modifier is necessary, the DAT possessor must be expressed. As Wierzbicka explains: "The essence of the 'simple ordinary bodily action' consists, it seems, in its integral character: the action is viewed as a simple, unitary event; if an action is not viewed as simple and ordinary it is no longer viewed as unitary either, and its different components (which are no longer predictable) have to be described separately" (1979:

338). In (37) Peter's eyes are being opened in a nontraditional way, by his hand, so both the hand and the possessor of the eyes must be expressed. Similarly, in (38) Peter does not just take off his hat in any old way; he takes it off in a specific way—with one finger—so the possessor and the place the hat is taken from must be indicated for clarity's sake:

(37) Piotr otworzył sobie (*Adj.) oczy ręką. (Wierzbicka 1979: 341)
Peter opened to-himself eyes (*Adj.) with hand

(38) Piotr jednym palcem zdjął sobie kapelusz z głowy (Wierzbicka 1979: 342).
Peter with one finger took off to-himself hat from head

Wierzbicka (1979: 338) suggests that sentences like (35) and (36) represent "a construction which grammaticalizes the concept of 'simple bodily action.'"

When experiencers and possessors are unexpressed, determining the referent is a challenge because, even if an antecedent is located in the directly preceding context, its location can vary widely. For example, in (39) the antecedent is the subject of the most recent main clause; in (40)–(41) it is another object in the same clause; in (42) it is the object of the preceding clause. In short, syntactic relations do not provide even heuristics, no less confident predictive power.

(39) Она меняет уже третью страну, начав все заного, — каким жизненным иммунитетом **надо было обладать**? (Богуславская: 109).
Ona menjajet uže tret'ju stranu, načav vse zanovo, — kakim žiznennym immunitetom **nado bylo obladat'**? (Boguslavskaja: 109).
she$_{NOM}$ changes already third country$_{ACC}$ having-started everything from-scratch what life immunity **necessary$_{IMPERSONAL}$ have$_{INFIN}$**
'This is already the third country she's come to, having started everything from scratch. What immunity to life would one need to do that?'

(40) Uciekła mu **żona**.
escaped$_{FEM.SG}$ him$_{DAT}$ **wife$_{NOM.FEM}$**
'His wife ran out on him.'

(41) Ее хоронили рядом с **мамой** (Минчин: 314).
Ee xoronili rjadom s **mamoj** (Minčin: 314).
her$_{ACC}$ buried$_{3PL}$ next to **mother**
'They buried her next to her mother.'

(42) Меня очень привлекало в американках отношение к маленьким детям, желание доставить ребенку радость: везти его в Диснейленд, украшать **комнату**, мчаться за игрушками (Богуславская: 266).
Menja očen' privlekalo v amerikankax otnošenie k malen'kim detjam, želanie dostavit' rebenku radost': vesti ego v Disnejlend, ukrašat' **komnatu**, mčat'sja za igruškami (Boguslavskaja: 266).

{ . . . } take$_{INFIN}$ him to Disneyland decorate$_{INFIN}$ **room** run-around$_{INFIN}$ for toys
'As regards American women's relationship to children, I was really drawn by their desire to make the child happy: taking him to Disneyland, decorating his room, running around buying him toys.'

Stepping back from a narrow definition of possessor and experiencer, we find more realms in which people and things must be linked to something else that is unexpressed. For example, any emotion needs some person or animal to experience it (43); birthdays, anniversaries, weddings, and so on, are holidays for someone (44); and matters affect or are salient to someone (45):

(43) Высшее образование ей досталось с таким трудом, что особой **радости** это не принесло.
Vysšee obrazovanie ej dostalos' s takim trudom, čto osoboj **radosti** èto ne prineslo.
higher education$_{NOM}$ her$_{DAT}$ came with such difficulty that special **joy$_{GEN}$** it$_{NOM}$ NEG brought
'Getting higher education was such a struggle that she didn't even get any particular pleasure from having done it.'

(44) У них была пересадка в Чикаго, и они решили провести ночь там. Чтобы прилетеь прямо в день рождения (Минчин: 141).
U nix byla peresadka v Čikago, i oni rešili provesti noč' tam. Čtoby priletet' prjamo **v den' roždenija** (Minčin: 141).
at them was transfer in Chicago and they decided spend$_{INFIN}$ night there in-order-to arrive$_{INFIN}$ exactly on **birthday**
'They had a layover in Chicago and they decided to spend the night there. In order to arrive right on the birthday [whose birthday is implied by the larger context; it is neither of theirs].'

(45) «Počemu ty prišël? Est' **delo**?»
«Počemu ty prišël? Est' **delo**?»
why you$_{NOM}$ came is-there **matter$_{NOM.SG}$**.
"Why did you come? Do you need something?"

The semantic ellipsis of possessors, experiencers, and other related categories can be detected using Ontological Semantic analysis. For example, an arm is part of a body, and a body must be associated with some person or animal; therefore, any reference to a body part will trigger a search for the associated animate being.

A sufficient description of a language in terms of the expression or nonexpression of possessors and experiencers should include the following:

- how possession and experiencers can be expressed,
- whether possessors and experiencers need to be expressed and in which contexts,
- what semantic nuances the expression/nonexpression of possessors and experiencers carries, and

- heuristics (if possible) for disambiguating between an elided definite possessor/experiencer and a generalized one.

However, the truly difficult problems of reference resolution for experiencers and possessors are not language specific, instead representing semantic and pragmatic issues that cut across languages. As such, approaches to their resolution, when developed, should be applied cross-linguistically, with the necessary parametric variation instantiated.

3.4. Unexpressed arguments in derived nominals and embedded verb forms

Derived nominals represent events, and they presuppose the same inventory of arguments as their respective verbs. For example, *invasion,* like *invade,* presupposes an agent who invades and a place invaded, which can be represented in English by *of-* and *by-*clauses, respectively: *the invasion of Poland by the Germans* (*in 1939*). Detection of implied arguments in nominalizations can be done in a variety of ways. If a system productively analyzes nominalizations as derived from their respective verbs, the subcategorization frame of the verb provides the template of expected arguments. If a system employs an Ontological Semantic lexicon—i.e., a lexicon whose entries are linked to ontological concepts—all nominalizations should be linked to the corresponding event, in which case the concept itself should suggest the expected semantic roles. Detecting the semantic ellipsis must then trigger a search for fillers for these roles.

A complete description of nominalizations includes how the verb's arguments are realized if they are expressed overtly. Highlights from Babby's (1997) description of Russian derived nominals can serve as an example:

- The head of the derived nominal must precede whichever of its arguments is expressed overtly, but the latter can occur in any order.
- The agent (i.e., would-be subject of the corresponding verb in a declarative sentence) is realized as NP_{GEN} for intransitive derived nominals, NP_{GEN} or NP_{INSTR} for monotransitive derived nominals, and NP_{INSTR} for ditransitive derived nominals.
- The ACC direct object of a transitive verb must be realized as NP_{GEN} in the derived nominal.
- Direct object–like entities that would receive lexical case marking by the corresponding verb receive the same case marking in the derived nominal.
- Objects other than direct objects retain the case marking they would have in a verbal construction.

This suggests that an inventory of parameters and values for the description of derived nominals should include those listed in table 13.4.

In analysis, this information will help to determine what case roles the overt categories (if any) in a clause with a nominalization could represent. Accounting for those should help to determine which roles have been semantically elided in the nominalization, and recovery procedures can be launched.

TABLE 13.4 Parameters and values to describe derived nominals. All parameters except the first need to be considered separately for intransitive, monotransitive, and ditransitive verbs, if there are differences in realization.

Parameters	*Values*
Order of overt elements	Free order of all elements Derived nominal + arguments in any order Derived nominal + agent + base direct object + other argument Etc.
Realization of agent	NP with some case marking PP with certain preposition and NP case marking
Realization of base direct object	As above
Realization of lexically case-marked direct objectlike entity	As above
Realization of other objects (split more finely, if necessary)	As above

There is a useful parallel between searching for the implied semantic roles for nominalizations and searching for the implied semantic roles of nonmatrix verbs. In both cases, syntactic heuristics, supplemented by a check of selectional restrictions, can in large part restrict—sometimes unambiguously—the potential referents for at least some semantic roles. For example, reconstructing the semantic roles for participial *invading* and nominalized *invasion* in (46a–b)

(46) a. The United States looked grimly on their invading Poland.
b. The United States looked grimly on their invasion of Poland.

might be done by writing rules like the following: if the participle or nominalization is "possessed" by a semantically valid agent, assume the possessor is the agent; if the participle takes a complement that semantically suits the verb's object, assume that complement is the object; if the nominalization takes an *of*-complement that semantically suits the verb's object, assume that complement is the object; and so on. Each language will require no small inventory of such rules, since they must cover not only nominalizations and participles but also things like infinitival complements of subject-control verbs (*I hate **to lose***) and object-control verbs (*I beg you **to stay***).

Languages that use subject ellipsis add yet another step to the analysis task, as illustrated by Polish examples (47)–(48):

(47) Zaprzestałem palenia papierosów.
stopped$_{1SG}$ smoking$_{GEN.SG}$ cigarettes$_{GEN.PL}$
'I stopped smoking.'

(48) Liczy na dostanie nagrody (Fisiak, Lipińska-Grzegorek, and Zabrocki 1978: 159).
counts$_{3SG}$ on receiving reward
'He/she counts on getting the reward.'

In both examples, the subject of the matrix verb must be the agent of the nominalization, based on syntactic rules. However, whereas in (47) the subject can be unambiguously recovered as 'I' based on verbal morphology, in (48) it could refer to any person. Therefore, filling out of the nominalization's semantic roles requires prior resolution of the elided subject pronoun—which, in actuality, is not very different from resolving an overt subject pronoun in English.

The suggestion that methods for resolving implied semantic roles should be approached in a unified way regardless of whether the category is a nominalization or a nonmatrix verb form is supported by a comparison for Polish made by Topolińska (n.d.: 80), who points out the connection between obligatory suppression of the subject in active clauses and the corresponding argument in nominalizations, as shown in (49):

(49) a. Janek zdecydował się na wyjazd do Francji.
Janek decided się on trip to France
b. Janek zdecydował się, że wyjedzie do Francji.
Janek decided się that will-go to France
c. *Janek zdecydował się, że on wyjedzie do Francji.
Janek decided się that he (= Janek) will-go to France

In (49a) the implied agent of nominalized *trip* must be *Janek.* In (49b), a subordinate clause with the finite verb form *will go* expresses the same meaning. The important comparison is between (49b) and (49c). In (49b) the subject of *will go* is elided, as it must be in order to establish the coreference relation between the decider and the goer. If the subject of *will go* were overt, the goer could not be Janek; it would have to be someone else.

A final consideration in analyzing nominalizations and nonmatrix verbs forms is that the referent could be a generalized person, as in (50):

(50) Jego imię jest łatwe do zapamiętania (Fisiak, Lipińska-Grzegorek, and Zabrocki 1978: 151).
his name is easy for remembering
'His name is easy to remember.'

The sum of this language evidence suggests that the resolution of semantic roles for nominalizations and nonmatrix verb forms should be approached in a unified, semantically supported (not strictly syntactic) manner. Various delineations of parameters, values, and realizations might be posited, among those the ones listed in table 13.5

4. Unexpressed morphemes

There are at least three circumstances under which morphological markers in some languages can be elided: haplology, morpheme ellipsis, and affixal loss during incorporation.

TABLE 13.5 Parameters and values for describing "participants" in nominalizations and nonmatrix verb forms

Parameters	*Values*
Referent for the agent of a nominalization	"Possessor" of nominalized form Subject-agent of matrix verb Generalized-human referent Etc.
Referent for the agent of a nonmatrix verb form	"Possessor" of nominalized form Subject-agent of matrix verb Generalized-human referent Etc.
Referent for the theme of a nominalization	NP complement of the nominalization PP complement of the nominalization Etc.

Haplology, the first source of missing morphemes, describes a situation in which a morpheme simultaneously carries out several different functions in a clause as a means of avoiding its repetition. In the following cross-linguistic examples, the morphemes that are collapsed into one surface realization are noted.[12]

Polish: Two instances of the "inherent" use of the particle się *(for* bać się *'fear' and* roześmiać się *'laugh')*

(51) Boję **się** głośno roześmiać (Kupść 1999: 21).
fear$_{1SG}$ **się** loudly laugh$_{INFIN}$
'I'm afraid to laugh loudly.'

Polish: One instance of inherent się *(for* modlenie się *'praying') and one instance of impersonal* się (*arising when the verb* zakazać *undergoes impersonalization*)

(52) W naszym kraju zakazało **się** modlenia.
in our country forbade$_{3SG}$ **się** praying$_{GEN}$
'Praying was forbidden in our country.'

French: Comparative que *and complementizer* que

(53) Je préfère que tu restes, plutôt **que** (***que**) tu t'en ailles
I prefer that you remain rather **than** (***that**) you go-away
'I prefer that you stay rather than that you go away.'

Spanish: Reflexive clitic se *and impersonal clitic* se (*haplology is possible at least in some dialects*)

(54) **Se** (***se**) lava.
one oneself washes
'One washes oneself.'[13]

Mandarin: Co-referential third-person pronoun ta

(55) Wo wen ta (*ta) mingtian lai bu lai.
I ask he_i (*he_i) tomorrow come not come
'I asked him_i whether he_i would come tomorrow.'[14]

Close descriptions of haplology in various languages (e.g., Kupść 1999 for Polish, Neeleman and van de Koot 2001 for Dutch, and Yip 1998 for Mandarin Chinese, Classical Greek, and Hindi) suggest that a full description of this phenomenon must start with questions like the following, which could easily be recast into parameters with language-specific values: Which morphemes, in what meanings, are subject to haplology? Does haplology occur only if the morphemes are adjacent? Does haplology require that the morphemes represent the same meaning? Does haplology always or optionally occur in instances where it is grammatically possible? How many meanings or instances of a given morpheme can be rendered by a single instance of that morpheme?

Morpheme Ellipsis, the second source of missing morphemes, differs from haplology in that the morphemes in question originate in different verb phrases or clauses. For example, plural person-number markers can be elided in Polish VP-coordinate structures when word stress is antipenultimate, indicated in (56) by capitalized vowels (Franks and Bański 1999: 126):

(56) CzytAli-**śmy** i pisAli-$\emptyset_{\textbf{śmy}}$
$read_{1PL}$ and wrote$\emptyset_{\textbf{1PL}}$
'We read and wrote.'

A description of Morpheme Ellipsis potential must be correlated with a description of morpheme-hopping potential. For example, the person-number marker that was elided in (56) can occupy practically any preverbal, non-sentence-intial position, as shown in (57):

(57) a. My-**śmy** znowu wczoraj poszli do parku.
we_{1PL} again yesterday went to park
b. My znowu-**śmy** wczoraj poszli do parku.
c. My znowu wczoraj-**śmy** poszli do parku.
d. My znowu wczoraj poszli-**śmy** do parku.
e. *My znowu wczoraj poszli do-**śmy** parku.
f. *My znowu wczoraj poszli do parku-**śmy**. (Franks and Bański 1999: 125)

What is important from the point of view of processing is that the verb form *poszli* is perfectly valid, representing the 3PL past form of the verb. However, in the second verb phrase of (56) and in (57a–c) that analysis of the token is not correct. The "missing" 1PL morphology must be recovered, in the former case by detecting and resolving ellipsis and in the latter case by linking the mobile affix back to its host verb.

The third source of missing morphemes is the loss of inflection during the grammatical process called incorporation, which is a subset of compounding. In noun incorporation, the verb and one of its arguments (usually the subject or object or just

the head noun of the subject or object) either occur as a single word or occur in series with certain morpho-syntactic modifications that indicate that incorporation has occurred. As Baker (1988: 26) reports, incorporated nouns generally lose their inflectional morphology, which would make their grammatical role opaque to an NLP system. The repercussions of this, and ways to recover the missing morphological information, must be subsumed under broader approaches to the machine analysis of incorporating languages.

5. Language Strategies

I separate Language Strategies from other instances of missing elements based on the hypothesis that certain types of missing elements in certain types of contexts can be most easily recovered by not using general principles but by accessing specialized sets of rules. As such, I remain theoretically neutral with respect to whether the missing elements represent syntactic or semantic ellipsis but do assert that some type of semantic reconstruction is necessary for a full semantic interpretation.

5.1. Dialogue strategies

Much of what happens in dialogues does not, I would suggest, represent syntactic ellipsis. When syntactic ellipsis occurs, there is a basically complete sentence structure with some slot(s) unfilled, and the major questions are what licenses the ellipsis and how to reconstruct the referent. With dialogues, the relevant sentence structures and questions to be asked are entirely different. *Yes-no* questions are a good case in point. If one is asked *Would you like a café au lait?* and answers *Yes!* the process is more like filling in the blanks, a sort of anti-ellipsis. By asking the question, the speaker creates a slot for a response, whose semantic and pragmatic constraints are quite clear. Whether one answers directly *yes* or *no* or provides an explanation that can be interpreted as one of those—*Unfortunately, I'm off caffeine*—the response is a specific type of rejoinder to the question itself. On the basis of Occam's razor, it is much simpler to view dialogues this way than to assume that the entire response is elided and has to be reconstructed.

Put another way, what is considered a complete *utterance* in the context of a dialogue would not be considered a complete *sentence* in isolation. As Vardul' (1969: 59–63) explains, the independent utterance *Yesterday* can be the answer to many different questions, like *When did you meet with him?* and *When did he arrive in Moscow?* so in one context *yesterday* by itself will be interpreted as *I met with him yesterday* whereas in another context it will be interpreted as *He arrived in Moscow yesterday*.

Different languages permit different dialogue strategies. For example, in Russian, *yes-no* questions are commonly answered by repeating the key word of the question, either with or without negation, as in (58):

(58) «Ты купила молоко?» — Купила». / «Не купила».
«Ty kupila moloko?» — «Kupila». / «Ne kupila».

you$_{NOM}$ bought milk$_{ACC}$ — bought / NEG bought
‘“Did you buy milk?” “Yes.” / “No.”’

(59) «Значит, говоришь, будут чудеса?» — «А как же! Вы — король сказочный? Сказочный! Живём мы в сказочном королевстве? В сказочном!» (Шварц 3: 535–536).
«Značit, govoriš’, budut čudesa?» — «A kak že! Vy — korol’ skazočnyj? Skazočnyj! Živëm my v skazočnom korolevstve? V skazočnom!» (Švarc 3: 535–536).
that-means say$_{2SG}$ will-be miracles COMPOUND PARTICLE you$_{NOM}$ king$_{NOM}$ fairytale$_{NOM}$ fairytale$_{NOM}$ live we$_{NOM}$ in fairytale$_{PREP}$ kingdom$_{PREP}$ in fairytale$_{PREP}$
‘“So, you say there will be miracles?” “But of course! You’re a fairytale king, aren’t you? Yes! And we’re living in a fairytale kingdom, aren’t we? Yes!”’

Repeating the key word(s) is used in Russian not only when answering questions but also when reacting to, responding to, or querying the previous statement—which is another type of dialogue strategy:

(60) «Ты же завалишь экзамены». — Не завалю». — «А я тебе говорю — завалишь» (Войнович 1: 15).
«Ty že zavališ’ èkzameny». — «Ne zavalju». — «A ja tebe govorju — zavališ’» (Vojnovič 1: 15).
you$_{NOM}$ PARTICLE will-fail$_{2SG}$ exams$_{ACC}$ NEG will-fail$_{1SG}$ but I$_{NOM}$ you$_{DAT}$ say$_{1SG}$ will-fail$_{2SG}$
‘“You’ll fail your exams.” “No, I won’t.” “I tell you, you will.”’

One restriction on the process is that only key elements can be overt in the reply, rejoinder, or query. For example, in (61) both the verb *считаю/sčitaju* ‘consider’ and the adverb *иногда/inogda* ‘sometimes’ are key elements and are overt. It would be incorrect, however, to respond *Иногда считаю тебя ∅ / Inogda sčitaju tebja ∅* ‘Sometimes I consider you ∅’ or *Иногда считаю ∅ дурой / Inogda sčitaju ∅ duroj* ‘Sometimes I consider ∅ a fool’.[15]

(61) «Ну что ты врёшь?» — «Я не вру». — «Значит, считаешь меня дурой . . . » — «Иногда считаю» (Токарева: 141).
«Nu čto ty vrëš’?» — «Ja ne vru». — «Značit, sčitaeš’ menja duroj . . . » — «Inogda sčitaju» (Tokareva: 141).
{ . . . } that-means consider$_{2SG}$ me$_{ACC}$ fool$_{INSTR}$ sometimes consider$_{1SG}$
‘“Why are you lying?” “I’m not lying.” “Then you consider me a fool . . .” “Sometimes.”’

There is, however, one exception to the repeat-only-key-elements rule: when the deletion of the theme in a Russian reply would contain only an overt subject, especially a pronominal one, partial repetition of the theme may be preferred or even required (as noted by Launer 1974, whose example is presented as [62]):

(62) «Кто стирает бельё?» — **«Стирает он»**.
«Kto stiraet bel’ë?» — «**Stiraet on**».

who washes laundry **washes he$_{NOM}$**
'"Who's doing the laundry?" "He is."'

(63) «И неужели не сказал, кто он такой? — поинтересовалась я. — Человек, когда знакомится, называет свою фамилию». — «**Он и назвал** — пробормотал что-то себе под нос, я и не разобрал» (Хмелевская 5: 68).
«I neuželi ne skazal, kto on takoj? — pointeresovalas' ja. — Čelovek, kogda znakomitsja, nazyvaet svoju familiju». — «**On i nazval** — probormotal čto-to sebe pod nos, ja i ne razobral» (Chmielewska 5: 68).
and really NEG said who he was (idomatic) asked-out-of-interest I$_{NOM}$ person$_{NOM}$ when meets(-someone) names self's surname **he$_{NOM}$ PARTICLE named** mumbled something self$_{DAT}$ under nose I$_{NOM}$ NEG made-it-out
'"Did he really not say who he was?" I asked. "A person usually says his name when he meets someone." "He did say it—he mumbled something to himself that I couldn't make out."'

To account for which elements can be deleted in answers to questions, and in which order, Kuno (1982) developed what he calls the "pecking order of deletion." For example, he explains that the (b) answer of Speaker B in (64) is unacceptable in this context (indicated by */) because the speaker's question really means "*When* did you get your Ph.D.?" making the time the crucial and therefore an unelidable element in the response:

(64) Speaker A: Did you get your Ph.D. *last year*?
Speaker B: a. Yes, I got it *last year.*
b. */ Yes, I got it ∅

Kuno further explains that this pecking order must be correlated with other aspects of grammar: for example, in (65), *I* is overt not because of the pecking order but because tensed verbs in English require overt subjects:

(65) Speaker A: Did you buy this watch in Switzerland?
Speaker B: Yes, I did. (Kuno 1982: 64)

Dialogues also permit reductions that are not canonically possible in a language. For example, as Thomas (1979) points out, in informal spoken English one can say, *Got the tickets?* and *Fancy a beer?* even though English is not a *pro*-drop language. Such types of noncanonical reductions should be covered not by the basic grammar, since they are register dependent, but by a supplement to be accessed only for dialogues.

Categories of dialogues that have been treated in NLP include: dialogues in restricted domains, like making hotel reservations or buying tickets (Tsujii and Nagao 1988);[16] questions answered by bare noun phrases (*Who did she marry? Leonard.*); and statements responded to by *wh*-questions (*Someone beat up Billy. Who?*) (Ginzburg, Gregory, and Lappin 2000).

Since division of dialogues into categories like this is valid cross-linguistically, with parameterization for things like the yes-no question-answering strategies

discussed earlier, approaches to processing dialogues should be applicable across languages.

5.2. Sentence fragments

In some instances, a full thought is presented in parts, either cooperatively among speakers or by a single speaker pausing for various stylistic or other effects. This strategy is used in many, if not all, languages:

(66) «Давай съедим **мороженого**». — «**С клубникой**».
«Davaj s"edim **moroženogo**». — «**S klubnikoj**».
let's eat **ice-cream with strawberries**
'"Let's have some **ice cream**." "**With strawberries**."'

(67) «Не выпьете ли стаканчик **шерри-бренди? Домашний**! Я сама делала, по рецепту моей бабушки» (Кристи: 62).
«Ne vyp'ete li stakančik **šerri-brendi? Domašnij**! Ja sama delala, po receptu moej babuški» (Christie: 62).
NEG will-drink$_{2PL}$ PARTICLE$_{INTERROG}$ glass **sherry-brandy homemade** { . . . }
'"Would you like a glass of sherry-brandy? Homemade! I made it myself using my grandmother's recipe."'

Most often, the added fragment is an adverbial and/or stands alone in an utterace that could not otherwise represent a full sentence. As such, it should be relatively easy to detect. Determining its relation to the previous utterance, however, will require syntactic and semantic analysis.

5.3. Nominal sentences

Nominal sentences are much like sentence fragments except that they tend to be used to set the stage or interject some evaluation or opinion. The "setting the stage" subtype can be used in literary (68) or journalistic works (69):

(68) The dead of night. A deserted street corner.

(69) Paris. June 18.

Such nominal sentences should be processed as adverbials attached to the following proposition or text.

When used as an expressive way of conveying an opinion, impression, and so forth —as in examples (70)–(72)—nominal sentences should be turned into a proposition with an existential verb assumed:

(70) A real lifesaver, your secretary!

(71) An amazing country, Israel.

(72) The greatest food on earth, chocolate!

Such reconstruction loses the stylistic flavor for human consumption but conveys the full meaning for machine processing.

There has been much debate, at least in the Russian school, about the status of nominal sentences. Although the specifics of the debate are not of concern here, they do shed some light on the nature of what is implied by such sentences. For example, to explain the difference between the word *winter* and the statement *Winter*, Fedosjuk suggests that one can pose a question for a statement but not for a word. Therefore, to *Winter* we could ask *What is going on in the speaker's sphere of vision?* So *Winter*, here, is the rheme for the implicit theme *in the speaker's sphere of vision* (1988: 9). Zolotova analyzes such sentences differently, saying that nominal sentences can be thought of as two-part sentences in which the named object(s) is/are predicated of the place of action. So, *Railroad car. Two passengers* is akin to *In the railroad car are two passengers.* (1982: 114). Using nominal sentences that indicate place and time, Zolotova writes, serves to characterize the given person as being "here" and "now" (1982: 115).

5.4. Unagentive impersonals

Impersonal clauses have no subject. Agentless impersonals correspondingly lack a semantic agent (cf. section 3.2). Agentless impersonals can be either inherent impersonals, whose subcategorization always lacks a subject, or derived impersonals, which can be used personally with a subject or impersonally without one. Semantically, unagentive impersonal sentences assert that the action is outside of any agent's control. Accordingly, many of the verbs that are inherently impersonal or are commonly used impersonally represent uncontrollable physical or mental states of people and weather phenomena. If one were to say there was an agentlike entity responsible for such events, it would have to be God or natural forces.

Inherent impersonals should pose no processing problems, since any language that has them should have a set of special syntactic rules that accepts as complete those subjectless sentences that use the associated inventory of verbs. An example of an inherent impersonal in Russian and Polish is the verb *nauseate*:

(73) R: Mamu **tošnilo/** P: Matkę **mdliło**.
mom$_{ACC}$ made-nauseous$_{3SG.NEUT.PAST}$
'Mom felt nauseous.'

Derived impersonals, however, will create analysis ambiguity, since a neuter singular verb that lacks a syntactic subject in Russian or Polish could represent either impersonal usage or personal usage with an elided subject. In order to reduce this ambiguity, either the semantic classes of verbs that commonly permit unagentive impersonals could be flagged in the ontology or the actual verbs that belong to those classes could be flagged in the lexicon—which would result in an increased preference for the impersonal analysis, barring evidence to the contrary (which could come only from a high confidence level for a candidate elliptical analysis).

The meanings that are expressed by impersonal clauses should largely overlap in languages that use impersonals, since the semantics that underlie impersonalization reflect real-world phenomena that are not language bound. Therefore, a list of meanings known to be realized by impersonal clauses in some languages would be one tool to support the description of this phenomenon cross-linguistically. A short sample list with examples follows.

Uncontrollable physical states

(74) Сразу после оглавления следовал перечень имущества — длинный список. **Михала начало трясти** после ознакомления с первым же пунктом “Перечня” (Хмелевская 5: 44).
Srazu posle oglavlenija sledoval perečen’ imučšestva – dlinnyj spisok. **Mixala načalo trjasti** posle oznakomlenija s pervym že punktom “Perečnja” (Chmielewska 5: 44).
immediately after title$_{GEN}$ followed inventory$_{NOM}$ property$_{GEN}$ long list **Michael$_{ACC}$ began$_{3SG.NEUT}$ shake$_{INFIN}$** after becoming-familiar with first PARTICLE point inventory$_{GEN}$
‘Immediately after the title came the inventory of property—a long list. Michael started to shake after reading the very first point of the “Inventory.”’

(75) Ознакомился Михал с пунктом вторым — **и у него дыхание перехватило** (Хмелевская 5: 44–45).
Oznakomilsja Mixal s punktom vtorym – i **u nego dyxanie perexvatilo** (Chmielewska 5: 44–45).
became-familiar Michael$_{NOM}$ with point second and **at him breathing intercepted$_{3SG.NEUT}$** [= idiom]
‘Michael read the second point and gasped.’

(76) **Михала** опять **бросило в жар** (Хмелевская 5: 47).
Mixala opjat’ **brosilo v žar** (Chmielewska 5: 47).
Michael$_{ACC}$ again **threw$_{3SG.NEUT}$ into heat**
‘Michael started to sweat again.’

(77) «Я сам себе не говорю правды уже столько лет, что и забыл, какая она, правда-то. **Меня от неё воротит, отшвыривает**» (Шварц 2: 289).
«Ja sam sebe ne govorju pravdy uže stol’ko let, čto i zabyl, kakaja ona, pravda-to. **Menja ot neë vorotit, otšvyrivaet**» (Švarc 2: 289).
{ . . . } **me$_{ACC}$ from it sickens$_{3SG}$ repels$_{3SG}$**
‘“I’ve been deceiving myself for so many years already that I’ve forgotten what truth is. It makes me sick, repels me.”’

(78) «[Наш старикан] заявился домой только на рассвете. От него ужасно **несло** рыбой, что с ним случается всегда, когда он озабочен» (Шварц 2: 288).
«[Naš starikan] zajavilsja domoj tol’ko na rassvete. **Ot nego** užasno **neslo ryboj**, čto s nim slučaetsja vsegda, kogda on ozabočen» (Švarc 2: 288).
{ . . . } **from him** horribly **smelled$_{3SG.NEUT}$ fish$_{INSTR}$** which with him happens always when he$_{NOM}$ worried

'Our old fella didn't get home till dawn. He smelled like fish—and how!—which always happens when he's worried.'

Uncontrollable or unexplainable emotional states or urges

(79) «Может, он опять нашёл худую и рыжую, его к таким **тянет**, можете мне поверить» (Хмелевская 2: 175).
«Možet, on opjat' našël xuduju i ryžuju, ego k takim **tjanet**, možete mne poverit'» (Chmielewska 2: 175).
perhaps he$_{NOM}$ again found skinny$_{ACC}$ and red-headed$_{ACC}$ **him$_{ACC}$ to such-ones pulls$_{3SG}$** can$_{2PL}$ me$_{DAT}$ believe$_{INFIN}$
'"Maybe he's found another skinny redhead, he's attracted to that kind, believe you me."'

(80) В разгаре игры (разумеется, пряток) **меня занесло в глухой лабиринт** с дубовыми виноградными гроздьями по стенам (Рачко: 35).
V razgare igry (razumeetsja, prjatok) **menja zaneslo v gluxoj labirint** s dubovymi vinogradnymi grozd'jami po stenam (Račko: 35).
in heat$_{PREP}$ game$_{GEN}$ naturally hide-and-seek **me$_{ACC}$ drew$_{3SG.NEUT}$ into dead-end maze** with oaken grape bunches on walls
'Right in the middle of the game (hide-and-seek, of course) I was drawn into a dead-end maze with oaken grape bunches on the walls.'

(81) **У Михала отлегло от сердца** (Хмелевская 5: 56).
U Mixala otleglo ot serdca (Chmielewska 5: 56).
at Michael$_{GEN}$ relieved$_{3SG.NEUT}$ from heart [= idiom]
'Michael calmed down.'

Weather phenomena

(82) P: Na leśnej polanie **bielało** od śniegu (Topolińska: 158).
on forest clearing **was-white$_{3SG.NEUT}$** from snow$_{GEN}$
'The forest clearing glistened white with snow.'

(83) После захода солнца **подморозило,** и копыта лошади звонко цокали по затвердевшему грунту (Линдгрен: 66).
Posle zaxoda solnca **podmorozilo**, i kopyta lošadi zvonko cokali po zatverdevšemu gruntu (Lindgren: 66).
after setting$_{NOM}$ sun$_{GEN}$ **froze-over$_{3SG.NEUT}$** and hooves horse$_{GEN}$ loudly clicked on frozen ground
'After sunset a frost set in and the horse's hooves clicked loudly on the frozen ground.'

While these semantically predictable instances of derived impersonalization cover a large percentage of impersonal sentences, the process is actually quite productive in Russian, meaning that as long as the action or event can be interpreted as being unagentive, the construction can be used. So, although flagging ontological concepts or lexical entries can act as a heuristic for disambiguation, it will not cover all the

cases, as shown by the following examples, which use quite a range of verbs (Babby 1994 provides many more examples):[17]

(84) [Two girls are playing at getting electric shocks]
И в какой-то момент **нас** довольно сильно **дёргало током** (Рачко: 91).
I v kakoj-to moment **nas** dovol'no sil'no **dërgalo tokom** (Račko: 91).
and at some moment **us$_{ACC}$** relatively hard **yanked$_{3SG.NEUT}$ shock$_{INSTR}$**
'And at one point we got a pretty good shock.'

(85) «У нас очень тихий город. Здесь никогда и ничего не случается». — «Никогда?» — «Никогда. На прошлой неделе, правда, был очень сильный ветер. У одного дома едва не **снесло крышу**» (Шварц 2: 274).
«U nas očen' tixij gorod. Zdes' nikogda i ničego ne slučaetsja». — «Nikogda?» — «Nikogda. Na prošloj nedele, pravda, byl očen' sil'nyj veter. U odnogo doma edva ne **sneslo kryšu**» (Švarc 2: 274).
{ . . . } at one house almost **blew-off$_{3SG.NEUT}$ roof$_{ACC}$**
'"Our town is very quiet. Nothing ever happens here." "Ever?" "Never. Well, true, last week there was a very strong wind and one house almost had its roof blown off."'

(86) Поверхность поля, по которому неслась наша коляска, сделалась куда менее ровной, чем в начале нашей скачки — иногда **нас подбрасывало** высоко в воздух вместе с коляской, и я уже стал опасаться, что для кого-то из наших это кончится сломанной шеей (Пелевин: 252).
Poverxnost' polja, po kotoromu neslas' naša koljaska, sdelalas' kuda menee rovnoj, čem v načale našej skački — inogda **nas podbrasyvalo** vysoko v vozdux vmeste s koljaskoj, i ja uže stal opasat'sja, čto dlja kogo-to iz našix èto končitsja slomannoj šeej (Pelevin: 252).
{ . . . } sometimes **us$_{ACC}$ threw$_{3SG.NEUT}$** high into air together with carriage { . . . }
'The surface of the field that our carriage was speeding along became a whole lot bumpier than it was at the beginning of this race—from time to time we were thrown high into the air along with the carriage, and I had already started to fear that one of us would end up breaking his neck.'

All of the preceding sentences showed the impersonal use of *verbs*. Agentless impersonal sentences can, however, also be built around adverbs and other predicate words that express the same types of events as many pleonastic-*it* sentences in English: *It is cold, It is good* (*that . . .*), *It is impossible* (*for . . .*), and so on. A description of these types of sentences includes the following three properties: the inventory of such lexical entities, what types of complements they take (if any) in impersonal usage, and whether or not they can occur as predicates in personal sentences (since, if so, a sentence in which they are used might be personal with an elided subject, which then needs to be disambiguated). Some samples from Russian follow:

(87) Возможно <Хорошо, Плохо, Очевидно, Жаль, Неисключено, . . . >, что . . .
possible <good, bad, obvious, too-bad, not-impossible . . . >, that . . .
'It's possible <good, bad, obvious, too bad, not impossible . . . > that . . .'

(88) Холодно <жарко> (кому)
cold <hot> (whom)
'It's cold <hot>' or 'I'm <you're, etc.> cold, hot.'

These words are not always used impersonally; they can also be used as predicates, as shown in (90). Dismbiguating between their impersonal use and their use with an elided subject will be require syntactic and, perhaps, semantic analysis:

(89) Это возможно.
that(-is) possible
'That's possible'

The parameters and values that can orient the description of agentless impersonals include those in table 13.6.

6. Closing thoughts

Ellipsis has earned its reputation as a complex phenomenon whose complexity derives from its reliance on many aspects of the language system. And cloaked under the obvious cross-modular issues are many component issues that, in and of themselves, are among the most difficult in all of linguistics. Take, for example, the related topics of transitivity, thematic roles, and the argument-adjunct distinction, all of which are crucial for the interpretation of unexpressed objects. I used some simplifying assumptions here in order to push forward work on ellipsis, like a simple definition of transitivity, and reference to an idealized lexicon for each language in question that specifies both argument structure and selectional restrictions. However, such simplifications will not hold up in practical applications. The problem is that the lexical specification of verbs presents enormous challenges from the semantic, syntactic, and practical-lexicographic perspectives, as shown vividly, for example,

TABLE 13.6 Parameters and values that can orient the description of agentless impersonals

Parameters	*Values*
Nature of "selecting" predicate	Inherent impersonal verb (syntactic limitations make it clear that there is no gap)
	Derived impersonal verb (the lack of a subject needs to be determined, as the subject could be elided rather than absent in the structure)
	Impersonal predicate word (the lack of a subject needs to be determined, since the subject could be elided rather than absent)
Semantic class of agentless state/action	Uncontrollable physical state of a person
	Uncontrollable or unexplainable emotional state or urge of a person
	Weather phenomenon
	Other (i.e., productive use of the strategy)

in the body of work by Levin and Rappaport Hovav.[18] Although work on ellipsis should not be postponed until some imagined stage of development has been achieved in all the contributing fields, a sober recognition of the interdependencies is essential for developing realistic expectations for ellipsis processing.

Any attempt to tackle a big problem, especially one comprised of many difficult and largely unresolved issues, runs the risk of being deemed too ambitious. NLP, however, is that rare field that lends such ambition a natural context since, in most cases, its object of interest is unrestricted text. This means that, in a short randomly selected text, one could conceivably encounter practically every type of ellipsis discussed in this book, not to mention hundreds of other linguistic phenomena equally deserving of attention. This, in itself, provides motivation for taking a bird's-eye view of related phenomena. Moreover, a common tactical approach in NLP is first handling the most frequent, typical cases, then using failure-driven methodologies for incremental system improvement. The theory developed here adopts that approach, proposing *some* parameters and values for *some* elliptical phenomena in *some* languages and explicating, as part of the theory, how those inventories can and should be expanded based on future cross-linguistic research.

In a perfect-for-NLP world, parallel stepwise development in all the fields that contribute to our understanding of ellipsis would be informed by the current or imagined future needs of NLP, with research directions and goals catered accordingly. Of course, this state of affairs will never obtain. However, NLP as a field can only benefit from encouraging the types of linguistic research that could eventually be applied to it. Highly specific descriptions of ellipsis, like those supported by this theory, are just one case in point. Until one has the goal of processing natural language, there is little need to specify patterns of ellipsis to the grain size attempted here: syntactic and pragmatic theories can work at a more conceptual level, and grammars of languages can present larger generalizations, relying on people's intuitions (for native speakers) or ability to subconsciously internalize patterns (for non-native speakers) to fill in the gaps. But the needs of NLP have set a task to descriptive linguists. One can choose to attribute this task to the field of descriptive computational linguistics or simply to descriptive linguistics, since it need not to be carried out by researchers trained in language processing. In either case, the description is independent both of any given application and from the state of the art across applications at any given time.

The most difficult aspects of ellipsis resolution are actually problems of reference resolution, meaning that they are as prevalent in languages where a given type of ellipsis is not permitted as in languages where it is. For example, *it* in English can function as a semantically vacuous subject (*It's raining*), a pronoun that refers to a person or object (*He kicked the door and it slammed back in his face*), a pronoun that refers to an event or situation (*The wind messed up Mattie's curls and it made her really mad*), or a pronoun that refers to practically any object, event, or real or potential utterance. The latter usage presents particular challenges for analysis that are identical to the challenges presented by elliptical realizations of corresponding meanings. Consider the following examples, in which Russian ellipsis corresponds to English *it*:

(90) «Ах! – живо воскликнула мисс Уэзерби. — **∅ Совсем из головы вылетело**! У меня есть для вас новость» (Кристи: 20).

«Ах — živo voskliknula miss Uèzerbi. — **∅ Sovsem iz golovy vyletelo**! U menja est' dlja vas novost'» (Christie: 20).
oh excitedly cried Miss Weatherby$_{NOM}$ **∅ totally from head$_{GEN}$ flew-out$_{3SG.NEUT}$** at me is for you news
'"Oh!" said Miss Weatherby excitedly. "**It completely slipped my mind**! I have news for you."'

(91) «Вам не скучно, мадемуазель Стамп?» Кларисса не ответила, **сочла ∅ ниже своего достоинства** (Акунин: 109).
«Vam ne skučno, mademuazel' Stamp?» Klarissa ne otvetila, **sočla ∅ niže svoego dostoinstva** (Akunin: 109).
{ . . . } Clarissa$_{NOM}$ NEG responded **considered ∅ below self's dignity**
'"You're not bored, Mademoiselle Stamp?" Clarissa didn't respond, **she considered it beneath her**.'

Whether a system is induced to search for a real-world referent by the presence of a syntactic gap or by the use of a practically contentless pronoun, the same recovery procedures will be necessary. Therefore, although identifying elliptical gaps is, certainly, an extra bit of work, it is not the greatest or most daunting task. Connecting any and all instances of objects and events to real-world referents is.

There are many ways to discover all the elliptical patterns in a language. For example, one could build a system that starts with only the coarse-grained basics and then use failure-driven methodologies to improve it. Or one could pay close attention to natural speech and jot down interesting examples to serve as the basis for generalizations. The theory and description in this book will, I hope, facilitate the task, but it will not directly tease out as yet unaccounted for phenomena and interactions of factors. My preferred discovery method is reading my way into the details of ellipsis. In fact, choosing ellipsis as a subject of study presents the perfect opportunity to consider reading literature part of a day's work.

This page intentionally left blank

NOTES

1. Getting Started

1. Syntactic gaps are indicated by ∅ and their antecedents, if syntactically accessible, are in boldface.

2. Collins English Dictionary (Glasgow: Harper Collins, 1992), 1598.

3. My use of Nirenburg and Raskin's work naturally reflects some degree of interpretation of their ideas and intent.

4. I will say little about phonology at this time, although it should be incorporated if speech-processing systems are to be covered.

5. VP-ellipsis usage in English is explored in detail in Johnson 2001.

6. The material in this subsection stems from joint work with Sergei Nirenburg, published in part in McShane and Nirenburg 2002.

7. Somers 1984, p. 508, describes adjuncts as "essentially optional elements which can be said to complete the meaning of the central predication as a whole."

8. See Williams 1977 for the distinction between sentence grammar and discourse grammar.

9. See Morgan 1982 for discussion of this debate.

10. Lambrecht 1994, p. 2, notes that there have been various names given to the study of the relationship between discourse and grammar, including "discourse pragmatics," "information structure," "functional sentence perspective," "information packaging," and "informatics." He provides a nice overview of the issues involved in integrated approaches to language.

11. Word order in such languages is not, strictly speaking, free—it depends upon the interaction of intonation and discourse structure.

12. Another option for discourse orientation is topic/comment in the way these terms are used to describe languages such as Chinese.

13. These are terms borrowed from the Russian literature: *prjamaja* and *obratnaja valentnost'*. See, for example, Zemskaja 1987 and Zemskaja, Kitajgordskaja, and Širjaev 1981.

14. Yokoyama 1986, p. 312, discusses a similar topic—the semantic relationship between verbs—using the concept of set membership. According to her analysis, sets can be of

the following three types: universal (*live/die*), culturally dependent (*pitch/hit*), and limited to a certain group of individuals who share some common knowledge.

15. For theoretical approaches to coordination and conjunction, see, e.g., Gazdar 1981, Gazdar et al. 1982, Goodall 1987, and Kayne 1994.

16. See Carston 1993, Mithun 1988, and Quirk et al. 1972 for discussion.

17. Sag et al. (1985: 135) propose to include NIL (i.e., syntactic null) among English conjunctions for the earlier clauses in multiclause coordinate structures.

18. Mithun 1988, p. 336, explains that in some languages two different types of intonation are used in coordinate structures with no overt conjunction: *no intonational break*, which signals that the clauses being conjoined are conceived of as "conceptually unitary," and *comma intonation*, which signals that they are conceived of as being "conceptually distinct."

19. They write: "Not all juxtaposed words, phrases or clauses are manifestations of asyndetic coordination. The possibility of inserting the coordinator *and* is evidence that the construction is asyndetic coordination" (Quirk et al. 1972: 550).

20. The coherence relations between clauses that I am collectively calling elaboration might need to be further split, but only if evidence suggests that such a split is relevant *for the instances of ellipsis that refer to A and E structures in this theory.* The research programs of other scholars have different requirements for the categorization of clause complexes and use the term "elaboration" with slightly different semantic implications (e.g., Halliday 1994, Kehler 2002).

21. See Halliday 1994 for a discussion of clause complexes and syndetic/asyndetic coordination.

22. For computational treatments of semantic ellipsis, see McShane, Nirenburg, and Beale 2004 and Viegas and Nirenburg 1995.

23. See especially Schank and Abelson 1977 for early work on scripts.

24. See, e.g., Beale, Nirenburg, and Mahesh 1996, Nirenburg and Levin 1992, and Nirenburg and Raskin 2004.

2. Object Ellipsis: Preliminaries

1. There are other parameters as well, e.g., whether or not an object is incorporated into the verb in an incorporating language. When the theory is applied to such languages, that and any other necessary parameters and values can be added.

2. See Massam and Roberge 1989 for a discussion of recipe contexts.

3. This coarse-grained approach to verb types and object status, while sufficient for the current stage of work on theory building, represents a significant simplification of the state of affairs. For example, Hopper and Thompson 1980 distinguishes 10 parameters of transitivity, each with high and low transitivity values. They suggest that the transitivity of a clause can be ranked by the combined values of these parameters and show that many languages exhibit morpho-syntactic reflexes of transitivity. Herbst and Roe 1996, p. 183, adds to the obligatory/optional split one other gradation: contextually optional complements, which account for the grammaticality of a sentence like *Steffi won* in an appropriate context. They say that Valency Theory "has devoted more attention [than other theories] to the actual criteria that can be employed to establish the distinction between complements and adjuncts . . . and is concerned with the problem of specifying the conditions under which certain complements have to be realized in a text and under which conditions they are purely optional" (180). Other works on various aspects of transitivity and valency include Apresjan 1986, Günther 1978, Holvoet 1991, Panevova 1992, Ružička 1978, Thomas 1979, and Vater 1978. Levin and Rappaport Hovav have made extensive contributions to this field as well (e.g., Levin 1993, 1999, Levin and Rappaport Hovav 1995, Rappaport Hovav and Levin 2002). Some of the issues related to the lexically determined status of an object are discussed in chapter 4.

4. Indirect objects in English can be expressed with or without *to*: *She gave the key to me. She gave me the key*. This is an idiosyncrasy of English.

5. See Bailyn 1995 for a discussion of configurational GEN, DAT, and INSTR.

6. Leonard Babby developed the notion of semantic case in Russian; see, e.g., Babby 1993.

7. Bailyn formulates these rules as follows: "Assign Genitive case to any case-bearing argument in SpecV under the scope of negation (whenever the Pr^0 that selects VP is associated with the quantificational feature [+q]" (1995: 114) and "Assign partitive genitive case to any case-bearing argument in SpecV whenever the Pr^0 that selects VP is associated with the quantificational partitive feature [+ 'p']" (116).

8. Rothstein 1980, pp. 79–82, discusses NP reference.

9. Gladney 1986, p. 142, discusses prepostional phrases of measure.

10. Chinese has a syntactic topic position, such that one can present the discourse topic followed by a full or elliptical sentence that comments on that topic. The following example, presented in English words, conveys the basic idea: *Fish$_{TOPIC}$ I like to go there*. 'I like to go there becase of the fish.' Topics in Chinese are not selected by the verb like arguments are—they are independent and limit the scope of the main predication. They may or may not corefer with a selected argument in the sentence they introduce. In topic-prominent languages, the discourse topic is sentence initial, is often signaled by a topic marker, and freely co-occurs with a subject. Huang 1984 notes that while Chinese has two topic positions and can have two arguments elided simultaneously, German has only one topic position and permits only one argument to be elided per sentence. The word *topic* is actually somewhat problematic, because it has the syntactic meaning used in Chinese and the discourse meaning used to refer to what the discourse is about.

11. Raposo 1986, p. 385, suggests that "the representation of this 'zero topic' in a syntactic structure . . . is nothing but an adequate metaphor to represent the interaction between grammar and pragmatics."

3. Direct Object Ellipsis with a Like Antecedent

1. This will be discussed further in chapter 11, section 2.

2. It is often difficult to tell whether a conjunctionless structure represents asyndetic coordination or the A and E Strategy. Heuristics can help: e.g., having a colon or dash between clauses rather than a comma tends to indicate A and E, as does having an overt subject of the second clause that is coreferential with the subject of the first clause. It is not clear that semantic analysis could be automated to aid in this determination.

3. In Russian, coordinated clauses with coreferential subjects may not have the subject overt in the latter conjunct(s).

4. In English, *take someone's raincoat* can mean either take it off his back (as here) or take it from his hands.

5. I adopt the term *gerund phrase* from Babby and Franks 1998.

6. The conjunction *čtoby* can also be used in the (a) and (b) variants, yielding the same ellipsis judgments.

7. The definite and discrete parameters are discussed for Russian in Channon 1983.

8. The source text implies *the* covered bridge and has direct object ellipsis.

9. This example is quoted from Neidle 1988, p. 15. The translation is mine.

10. The conjuncts may be split by parenthetical material.

11. A more detailed cross-linguistic analysis can be found in McShane 1999a.

12. The direct object was overt in the source text.

13. It is impossible to posit a Czech variant for this sentence because the Czech conjunction *i* 'and$_{COORD}$' is not used in such contexts.

14. The direct object was overt in the source text.

15. Polish informants gave different judgments for this example: one considered DO ellipsis optional, another considered it highly preferred, and still another considered it impossible. Accepting the majority opinion, I mark it optional.

16. The conjunction *i* 'and' is technically possible in this Polish example but significantly worse than the conjunction *więc* 'so'.

4. Direct Object Ellipsis with a Nominative Antecedent

1. The punctuation ellipsis included here to shorten the example represents narrative; no speech has been omitted.

2. The literature on this topic can be found under the key word *unaccusativity*. The book *Unaccusativity* (Levin and Rappaport Hovav 1995) contains an in-depth treatment.

3. Any time that pairs of compared sentences are more than minimally different there is the chance that factors other than the one being tested for play a role: for example, in (12b) the imperative verb in the second clause might facilitate ellipsis. Despite this problem, it is preferable to present natural, realistic contexts rather than to force sentences into a predetermined pattern.

4. The empty category in Polish is case-marked GEN here because, in negated clauses, Polish imposes GEN of negation on all direct objects that would have been configurational ACC in an affirmative clause. This surface case-marking wrinkle does not, however, affect the point in question.

5. The Czech would require the existential verb in the first sentence: „*To* ***je*** *zima*!" „*Opravdu? Já* ***to/*∅*** *vůbec necítím.*"

5. Direct Object Ellipsis with an Oblique Antecedent

1. In this section I am not attempting to cover every use of every case (e.g., lexical GEN vs. GEN-Neg, lexical DAT vs. DAT recipient); I am simply pointing out the unmistakable increase in ellipsis potential associated with a pronominal antecedent.

2. It is not unprecedented to consider form a major determining factor in a grammatical process: e.g., Franks 1995, p. 63, concludes that "it is morphological realization, rather than the case per se, that is relevant in licensing ATB [across-the-boards] dependencies."

3. In VP-coordinate structures there is no second subject; in A and E structures the second subject is generally elided.

4. In some instances, ambiguity that derives from ellipsis is permitted—even welcomed—but this applies only to certain types of ambiguity, such as ambiguity between a definite human object and a generalized-human one, as in: *Housework never gives you a rest!*

5. The ellipsis-blocking effects of verbs with wide selectional restrictions even hold when the antecedent is an ideal configurational ACC noun phrases, as in: *Ja xoču znat'* ***vsju pravdu****, ja dolžna* ***ee*** *znat'*! 'I want to know **the whole truth**$_{\text{ACC}}$, I must know **it**$_{\textbf{ACC}}$!' If the direct object is not expressed overtly, the interlocutor is left waiting for a complement (perhaps a clausal one) to follow.

6. McShane 1998 proposes a theoretical explanation of why A and E structures permit case mismatches more liberally than coordinate structures do. Coordinate structures belong to sentence grammar, where there is a direct link between antecedent and elided category, whereas A and E structures belong to discourse grammar, where the link between the antecedent and elided category is mediated by an intervening null discourse topic of the type proposed in Huang 1984.

7. This example is from the article "Kino dlja Tamary" 'A Movie for Tamara'. The preceding context indicates that the infinitival *ne zabyt'* has the implied modality *must* and the subject *I*.

6. Elided Lexically Case-Marked Objects

1. It is practically impossible to study the ellipsis of configurationally case-marked objects with Genitive-of-Negation or Partitive-Genitive case marking because, at least in Russian, these are semantically enhanced and often optional case-marking variants of default ACC. Therefore, one cannot say definitively whether an elided object has ACC or GEN case marking.

8. Head Noun Ellipsis . . . or Not?

1. See Lobeck 1995 for a discussion of head noun ellipsis in English.
2. See McShane 2002b for further discussion of bare adjectives in Russian.
3. Head noun ellipsis can also be licensed by a number or quantifier in Russian, with the same usage patterns.
4. The last clause of this example has an elided verb as well.
5. Actually, there's a bit of ambiguity in English as well, discussed in chapter 1, section 4.1.

9. Verbal Ellipsis with One Licensor

1. Some other patterns of verbal ellipsis have been discussed in the literature (albeit less widely) and should ultimately be included in this theory of ellipsis. One example is pseudo-Gapping, as in *Mary opened up the large presents before she did the small ones*.
2. Such an overview of the literature could constitute a book in itself. Among the most influential sources are Hudson 1989, Jackendoff 1971, Johnson 2001, Kehler 2002, Lobeck 1995, Merchant 2001, Neijt 1979 and 1981, Rooryck 1985, Ross 1967 and 1969, and Stillings 1975.
3. Lobeck 1995, pp. 24–25, follows Hankamer and Sag 1976 and Williams 1977 in stating that Gapping cannot occur across utterance boundaries. Merchant 2001 follows Sag 1976 in considering such structures grammatical as long as the speakers are considered to be collaborating. The latter position is supported by example (12).
4. This example is based on Neijt 1981.
5. Although English has morphological case marking on pronouns, it is systematically too weak to permit a Gapping interpretation of sentences like (18). Therefore, the only way to interpret a sentence like the following is as a structure composed of coordinated verb phrases with *she* incorrectly used in place of *her*: *I asked him to sing and she to play the piano*. It is absolutely impossible in English to interpret *she* as the subject of Gapped *asked*.
6. In order to explain the impossibility of a Gapping interpretation of English examples like (18), Hankamer 1973, p. 29, proposes the No-Ambiguity Condition: "Any application of Gapping which would yield an output structure identical to a structure derivable by Gapping from another source, but with the 'gap' at the left extremity, is disallowed."
7. Lobeck cites a similar example from Chao 1987.
8. Lobeck's [e] is changed to ∅ for consistency.
9. These examples are based on (1)–(3) from Gardent 1993, in turn drawn from Klein and Stainton-Ellis 1989.
10. See especially Fiengo and May 1994, Hardt 1999, Hobbs and Kehler 1997, and Kehler and Shieber 1997.

11. Kehler follows up this generalization with some borderline cases where a strict interpretation is not completely ruled out because of semantic clues: e.g., *? The alleged murderer$_i$ defended himself$_i$, and his lawyer$_j$ did too* (2002: 58). It is the interpretation by which the lawyer defends the murderer that is marked as questionable. Borderline cases, while theoretically of interest, are of less interest for my practical aims since, in a practical environment, the game of dealing with examples like this would likely not be worth the candle. That is, whereas on ontological grounds one might favor the reading where the lawyer defended the criminal, on syntactic grounds—which might even be supported by textual cues (e.g., if the lawyer himself was accused of fraud while defending the criminal)—one would have to choose the reading where the lawyer defended himself.

10. Verbal Ellipsis with a Combination of Licensors

1. The Russian-language literature provides various descriptions of such sentences (e.g., Dun 1982, Zemskaja 1973, Zolotova 1982), which have served as a starting point for the current treatment.

2. See McShane 2000 for a different treatment of the data, not oriented toward NLP.

3. *Go* here is the highly colloquial American slang for "say." This sentence is highly colloquial in Russian as well.

4. The description of this type of ellipsis and many of the Russian examples are drawn from Zemskaja 1973.

5. It should be *черт-те*.

6. In highly colloquial English, one could also say *Those poor kids . . . music, foreign language . . . and God knows what else!*

7. I leave the possible ellipsis-affecting properties of *èto* used as a particle to future study.

8. In English, we can say *Look with your eyes!* or just *With your eyes!*, but there can be no overt subject as there is in Russian: **You ∅ with your eyes*.

9. See Zemskaja 1973 for a discussion of default interpretations in Multi-VE sentences.

10. Russian's use of *head* where English would use *hair* is language-specific detail that should be handled during the acquisition of the Russian ontological lexicon.

11. The intonation of this last remark implies "that'll do him good; it serves him right; that'll sober him up."

12. One Czech informant accepts the verbless *já nic*, while another prefers the verbal form *nevadí*. Both Polish *zimno* and Czech *je zima* are ambiguous; English *it's cold* (i.e., either the weather or the kettle/water) retains this ambiguity.

13. *W tym zimnie i wilgoci* 'in this cold and dampness' is a restatement of *tak* and thereby does not contribute to licensing the ellipsis.

14. In Czech, this would be conveyed by a nonelliptical structure.

15. This example includes several idiomatic turns of phrase, so the word-for-word translation only loosely represents the structure of the Polish variant.

11. Ellipsis of Minor Parts of Speech

1. See Carston 1993 for a discussion of the semantics of conjunctions.

12. Dependencies in Ellipsis: A Polish Case Study

1. See McShane 1998 and 1999a,b for DO-ellipsis potential in Polish. See McShane 2002a for more extensive discussion of the ellipsis potential of all three of these categories in Polish.

2. Clitic placement is discussed widely in the literature, e.g., Franks 1998, Franks and Bański 1999, Kupść 1999.

3. Here we are considering instances of *się* in different verb phrases. Within a single verb phrase, instances of *się* can also be collapsed, creating haplology (see chapter 13, section 4).

4. The direct object was overt in the source text.

13. More Elliptical Phenomena

1. Wilson 2000 discusses a number of additional Language Strategies—what he calls "telegraphic ellipsis"—that require special descriptive and processing approaches: e.g., headlines, lists, headings, signs.

2. Whereas, in the past tense, Polish verbs inflect for person, number, and gender, Russian verbs inflect only for person and gender.

3. For syntactic treatments see, e.g., Franks 1995, Huang 1984, Law 1993, and Lillo-Martin 1986 and 1991. For discourse treatments see, e.g., Nilsson 1982. For computational treatments, see, e.g., Yamamoto and Sumita 1998.

4. I base this and the next next generalization on about one hundred Russian examples collected manually. Yokoyama 1986, pp. 31–32, explains that the awareness of {I, you, here, now} is a common (though not obligatory) prelinguistic set of referential knowledge, which she refers to as {DEIXIS}. Firbas 1992, p. 24, presents a similar idea:

> The chief representative of the first group [of referents] is the speaker/writer. As producer, he or she is permanently present in the ever-changing situational context. Another important representative is the listener/reader, to whom the sentence is addressed. When the sentence is perceived, he or she is even present in the flesh. Of all the other notions, it is that of people in general and that of nature in general which, in regard to permanent obviousness, come close to the notions of speaker/writer and listener/reader. These four notions are frequently referred to by such pronouns as *I*, *you*, the impersonal *it*, *one*, German *man*, French *on*, etc. . . . It is worth noting that these pronouns can be introduced into the discourse directly, not pronominalizing any antecedents (predecessors). . . . This bears out the permanent places that the notions involved occupy in the ever-changing immediately relevant situational contexts.

5. Nonfinite verb forms are used in English as parentheticals (*to be frank, to generalize*) and as arguments of other predicates (*Being good is important for a child, To be invited to speak there is an honor*), but these are quite different issues.

6. I avoid the quagmire of naming each mood, since many naming and classification schemes have been proposed. See, e.g., Jespersen 1963, Palmer 1986, and Quirk et al. 1972.

7. For some applications, it might be necessary to distinguish between generalized people and some unspecified person or people.

8. For discussion of this in Russian see Ružička 1978 and in Polish see Topolińska, n.d. p. 148.

9. See Grimshaw 1992 for argument-adjuncts.

10. The large literature on passives should be consulted for more descriptive parameters and values.

11. I present the translations without Wierzbicka's indices.

12. Yip 1998, p. 220, notes that English *cats'* rather than **cats's* is an example of the avoidance of repetition of homophonous morphemes. However, the *s'* morpheme, although perhaps diachronically divisible, is now an established morpheme that does not need to be broken down into component parts and does not indicate Morpheme Ellipsis.

13. The French and Spanish examples were drawn from Neeleman and van de Koot 2001, modified to the glossing conventions used here. Neelman and van de Koot attribute the French example to Radford 1977 and 1979, and the Spanish example to Grimshaw 1997.

14. Example from Yeh 1994, quoted in Yip 1998.

15. Saloni 1974, pp. 77–78, supports this generalization with comparable examples from Polish.

16. Tsujii and Nagao 1988 have suggested that certain types of dialogue translation systems for restricted domains are architecturally more feasible than currently available text translation systems, since one can define by the purpose of the dialogue which information should be conveyed in the translation. So, "results of *understanding* [a text] can be represented in a language independent (but task dependent) way" (690).

17. Babby 1994 argues that Russian verbs that can be used as derived impersonals (which he calls "adversity impersonals" because they tend to convey adverse physical events) have an optional external agentive thematic role; therefore, when they are used impersonally, that thematic role is simply not selected. This suggests one possible way of lexically specifying those verbs that are open to adversity impersonals in a language.

18. See Levin 1993, 1999, Levin and Rappaport Hovav 1995, and Rappaport Hovav and Levin 2002, among many others.

REFERENCES

Linguistic Sources

Apresjan, Jurii D. 1986. Tipy sintaksičeskogo otsutstvija v russkom jazyke. In *Problemy semantiki predloženija: Vyražennyj i nevyražennyj smysl*. Krasnojarsk: Krasnojarskij ordena "Znak početa," gos. ped. Institut, 111–115.

Babby, Leonard H. 1993. Case theory. In Carlos P. Otero, ed., *Noam Chomsky: Critical Assessments.* London: Routledge, 630–652.

Babby, Leonard. H. 1994. A theta-theoretic analysis of adversity impersonal sentences in Russian. In Sergey Avrutin, Steven Franks, and Ljiljana Progovac, eds., *Formal Approaches to Slavic Linguistics, the MIT Meeting.* Ann Arbor: Michigan Slavic Publications, 25–67.

Babby, Leonard H. 1997. Nominalization in Russian. In Wayles Brown (ed.), *Formal Approaches to Slavic Linguistics, the Cornell Meeting, 1995.* Ann Arbor: Michigan Slavic Publications, 54–83.

Babby, Leonard H., and Steven Franks. 1998. The syntax of adverbial participles in Russian revisited. *Slavic and East European Journal* 42, 483–515.

Bahan, Benjamin, Judy Kegl, Robert G. Lee, Dawn MacLaughlin, and Carol Neidle. 2000. The licensing of null arguments in American Sign Language. *Linguistic Inquiry* 31.1, 1–28.

Bailyn, John F. 1995. *A Configurational Approach to Russian "Free" Word Order*. Ph.D. Dissertation, Cornell University.

Baker, Mark C. 1988. Morphology and syntax: An interlocking independence. In Martin Everaet, Arnold Evers, Riny Huybregts, and Mieke Trommelen, eds., *Morphology and Modularity*. Dordrecht: Foris, 9–32.

Beale, Stephen, Sergei Nirenburg, and Kavi Mahesh. 1996. Semantic analysis in the Mikrokosmos machine translation project. In *Proceedings from the Symposium on NLP, Kaset Sart University, Bangkok, Thailand, 1995.*

Carston, Robyn. 1993. Conjunction, explanation and relevance. *Lingua* 90(1–2), 27–48.

Channon, Robert. 1983. A comparative sketch of certain anaphoric processes in Russian and English. In Michael S. Flier, ed., *American Contributions to the Ninth International Congress of Slavists: Kiev*, September 1983, vol. 1: Linguistics. Columbus, OH: Slavica Publishers, 51–69.

Chao, W. 1987. *On Ellipsis.* Ph.D. Dissertation, University of Massachusetts, Amherst. Published by the Graduate Linguistics Students Association, UMass, Amherst.

Chvany, Catherine V. 1996a. A continuum of lexical transitivity: Slightly-transitive verbs. In Olga T. Yokoyama and Emily Kleinin, eds., *Selected Essays of Catherine V. Chvany*. Columbus, Ohio: Slavica Publishers, 161–168. First published in 1990 in Björklund et al., eds., *Carina Amicorum, Carina Davidsson septuagenariae 28.3.1990*, Åbo: Pargas: Åbo akademis förlag, 51–63.

Chvany, Catherine V. 1996b. Hierarchies in the Russian case system: For N-A-G-L-D-I, against N-G-D-A-I-L. In Olga T. Yokoyama and Emily Kleinin, eds., *Selected Essays of Catherine V. Chvany*. Columbus, Ohio: Slavica Publishers, 175–187. First published in 1982 in *Russian Language Journal* 36/125, 133–147.

Comrie, Bernard, and Norval Smith. 1977. Lingua descriptive questionnaire. *Lingua* 42, 1–72.

Dun, Čžènmin. 1982. *Grammatičeskaja priroda, klassifikacija i stilističeskaja funkcija èlliptičeskix predloženij.* Ph.D. Dissertation, Shanghai Institute of Foreign Languages.

Evgen'eva, A. P., ed. 1981. *Slovar' russkogo jazyka v četyrex tomax, izd. vtoroe*. Moscow: Russkij jazyk.

Fedosjuk, M. Ju. 1988. *Nejavnye sposoby peredači informacii v tekste*. Moskva: Mosk. gos. ped. inst. im. V. I. Lenina.

Fiengo, Robert, and Robert May. 1994. *Indices and Identity*. Cambridge, Mass.: MIT Press.

Firbas, Jan. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.

Fisiak, Jacek, Maria Lipińska-Grzegorek, and Tadeusz Zabrocki. 1978. *An Introductory English-Polish Contrastive Grammar*. Warsaw: Państwowe wydawnictwo naukowe.

Fominyx, V. I. 1965. Tipy èlliptičeskix predloženij v russkom jazyke i èkvivalenty v češkom. In *Voprosy sovremennogo russkogo jazyka i dialektologii*. Tjumen': Sverdlovskij gos. ped. institut, 99–114.

Fowler, George. 1996. Oblique passivization in Russian. *Slavic and East European Journal* 40, 519–545.

Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. Oxford: Oxford University Press.

Franks, Steven. 1998. Clitics in Slavic. Unpublished paper from the Workshop on Comparative Slavic Morphosyntax, Spencer, Indiana. Available at http://www.indiana.edu/~slavconf/linguistics/frnks.pdf.

Franks, Steven, and Piotr Bański. 1999. Approaches to "schizophrenic" Polish person agreement. In *Proceedings from the Annual Workshop on Formal Approaches to Slavic Linguistics*, Seattle, 1998. Ann Arbor: Michigan Slavic Publications, 123–143.

Gardent, Claire. 1993. A unification-based approach to Multiple VP Ellipsis resolution. In *Proceedings of the 6th European Meeting of the Association for Computational Linguistics*. Utrecht, The Netherlands: OTS-Research Institute for Language and Speech, Utrecht University, April 21–23, 1993, 139–148.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structures. *Linguistic Inquiry* 12, 155–184.

Gazdar, Gerald, Geoffrey Pullum, Ivan Sag, and Thomas Wasow. 1982. Coordination and transformational grammar. *Linguistic Inquiry* 13, 663–676.

Ginzburg, Jonathan, Howard Gregory, and Shalom Lappin. 2000. A dialogue system for interpreting fragments. In *Proceedings of the ESSLLI-2000 Workshop on Linguistic Theory and Grammar Implementation*. Birmingham, UK Aug. 6–12, 2000: 187–199.

Gladney, Frank Y. 1986. Prepositions and case government in Russian. In Richard D. Brecht and James S. Levine, eds., *Case in Slavic*. Columbus, Ohio: Slavica Publishers, 130–151.

Goodall, Grant. 1987. *Parallel Structures in Syntax: Coordination, Causatives and Restructuring*. Cambridge: Cambridge University Press.

Grimshaw, Jane. 1992. *Argument Structure*. Cambridge, Mass.: MIT Press, Linguisitic Inquiry Monographs.

Grimshaw, Jane. 1997. The best clitic: Constraint conflict in morphosyntax. In Liliane Haegeman, ed., *Elements of Grammar: Handbook of Generative Syntax*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 169–196.

Günther, Hartmut. 1978. Valence in categorial syntax. In Werner Abraham, ed., *Valence, Semantic Case, and Grammatical Relations*, Workshop studies prepared for the 12th International Congress of Linguists, Vienna, August 29 to September 3, 1977. Amsterdam: John Benjamins, 127–156.

Gvozdev, A. N. 1958. *Sovremennyj russkij literaturnyj jazyk. Čast' II, Sintaksis*. Moscow: Gosudarstvennoe učebno-pedagogičeskoe izd., Ministerstva prosveščenija RSFSR.

Gvozdev, A. N. 1965. *Očerki po stilistike russkogo jazyka, Izd. tret'e.* Mosow: Prosveščenie.

Halliday, Michael Alexander Kirkwood. 1994. *An Introduction to Functional Grammar*. New York: Arnold.

Hankamer, Jorge. 1973. Unacceptable ambiguity. *Linguistic Inquiry* 4, 17–68.

Hankamer, Jorge, and Ivan A. Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7, 391–426.

Hardt, Daniel. 1999. Ellipsis and the structure of discourse. In *Workshop on Ellipsis and Information Structure*. Berlin: ZAS.

Herbst, Thomas, and Ian Roe. 1996. How obligatory are obligatory complements? An alternative approach to the categorization of subjects and other complements in valency grammar. In *English Studies* 2, 179–199.

Hobbs, Jerry R., and Andrew Kehler. 1997. A theory of parallelism and the case of VP Ellipsis. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*. Madrid, Spain, Morgan Kaufmann Publishers July 1997, 394–401.

Holvoet, A. 1991. *Transistivity and Clause Structure in Polish.* Warsaw: Slawistyczny Osrodek Wydawniczy, Omnitech Press.

Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, 251–299.

Huang, C. T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15, 531–574.

Hudson, Richard A. 1989. Gapping and grammatical relations. *Journal of Linguistics* 25, 57–94.

Jackendoff, Ray S. 1971. Gapping and related rules. *Linguistic Inquiry* 2, 21–35.

Jakobson, Roman. 1984. Morphological observations on Slavic declension. In L. R. Waugh and M. Halle, eds., *Russian and Slavic Grammar, Grammatical Studies 1931–1981*. Berlin: Mouton de Gruyter, 105–133.

Jespersen, Otto. 1963. *The Philosophy of Grammar*. London: George Allen and Unwin.

Johnson, Kyle. 2001. What VP Ellipsis can do, what it can't, but not why. In Mark Baltin and Chris Collins, eds., *The Handbook of Contemporary Syntactic Theory*. Malden, Mass.: Blackwell Publishers, 439–479.

Kaleta, Zofia. 1995. *Gramatyka języka polskiego dla cudzoziemców*. Krakow: Nakładem Uniwersytetu Jagiellońskiego.

Kayne, Richard. 1994. *The Asymmetry of Syntax*. Cambridge, Mass.: MIT Press.

Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, 63–99.

Kehler, Andrew. 2002. *Coherence, Reference and the Theory of Grammar*. Stanford, Cal.: CSLI Publications.

Kehler, Andrew, and Stuart Shieber. 1997. Anaphoric dependencies in ellipsis. *Computational Linguistics* 23, 457–466.

Klein, Ewan, and Kate Stainton-Ellis. 1989. *A Note on Multiple VP Ellipsis*, Research Paper EUCCS/RP-30. Centre for Cognitive Science, University of Edinburgh.

Kuno, Susumu. 1982. Principles of discourse deletion. *Journal of Semantics* 1, 120–154.

Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse, and Empathy.* Chicago: University of Chicago Press.

Kupść, Anna. 1999. Haplology of the Polish reflexive marker. In Robert D. Borsley and Adam Przepiórkowski, eds., *Slavic in Head-Driven Phrase Structure Grammar*. Stanford, Cal.: CSLI Publications, 91–124.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representation of Discourse Referents.* Cambridge: Cambridge University Press.

Lambrecht, Knud, and Kevin Lemoine. 1996. Vers une grammaire des compléments zéro en français parlé. In Jean Chuquet and Marc Fryd, eds., *Absence de Marques et Représentation de L'absence*, *Actes du Neuvième Colloque du CerLiCO.* Rennes: Presses Universitaires de Rennes, 279–309.

Lappin, Shalom. 1999. An HPSG account of antecedent contained ellipsis. In S. Lappin and E. Benmamoun, eds., *Fragments*: *Studies in Ellipsis and Gapping*. New York: Oxford University Press, 1999, 68–97.

Lappin, Shalom, and Herbert J. Leass. 1994. A syntactically based algorithm for pronominal anaphora resolution. *Computational Linguistics* 20, 535–561.

Lapteva, Ol'ga A. 1976. *Russkij razgovornyj sintaksis.* Moscow: Nauka.

Launer, Michael K. 1974. *Elementary Russian Syntax.* Cambridge, Mass.: Slavica Publishers.

Law, Paul. 1993. On null subjects and null arguments. *Canadian Journal of Linguistics* 38, 1–41.

Levin, Beth. 1993. *English Verb Classes and Alternations.* Chicago: University of Chicago Press.

Levin, Beth. 1999. Objecthood: An event structure perspective. In *Proceedings of CLS 35*, volume 1: *The Main Session.* Chicago: Chicago Linguistic Society, University of Chicago, 223–247.

Levin, Beth, and Malka Rappaport Hovav. 1995. *Unaccusativity*. Cambridge, Mass.: MIT Press.

Levin, Nancy, and Ellen Prince. 1986. Gapping and causal implicature. *Papers in Linguistics* 19, 351–364.

Lillo-Martin, Diane C. 1986. Two kinds of null arguments in American sign language. *Natural Language and Linguistic Theory* 4, 415–444.

Lillo-Martin, Diane C. 1991. *Universal Grammar and American Sign Language: Setting the Null Argument Parameter*. Dordrecht: Kluwer Academic Publishers.

Lobeck, Anne C. 1995. *Ellipsis: Functional Heads, Licensing, and Identification*. New York: Oxford: Oxford University Press.

Longacre, Robert E. 1964. *Grammar Discovery Procedures*. The Hague: Mouton.

Maciejewski, Witold. 1991. Ellipsis in Polish and Swedish: Some typological considerations. *Lingua Posnaniensis: Czasopismo poświęcone językoznawstwu porównawczemu i ogólnemu* 32–33 (1989–1990), 169–179.

Massam, Diane, and Yves Roberge. 1989. Recipe context null objects in English. *Linguistic Inquiry* 20, 134–139.

McShane, Marjorie J. 1998. *Ellipsis in Slavic: The Syntax-Discourse Interface*. Ph.D. Dissertation, Princeton University.

McShane, Marjorie J. 1999a. The ellipsis of accusative direct objects in Russian, Polish and Czech. *Journal of Slavic Linguistics* 7, 45–88.

McShane, Marjorie J. 1999b. Predictive rules of direct object ellipsis in Russian. In Katarzyna Dziwirek and Herbert S. Coats, eds., *Proceedings from the Annual Workshop on Formal Approaches to Slavic Linguistics*, Seattle, 1998. Ann Arbor: Michigan Slavic Publications, 329–348.

McShane, Marjorie J. 2000. Verbal ellipsis in Russian, Polish and Czech. *Slavic and East European Journal* 44, 195–233.

McShane, Marjorie J. 2002a. Hierarchies of parallelism in elliptical Polish structures. *Journal of Slavic Linguistics* 8, 83–117.

McShane, Marjorie. 2002b. So Where's That Noun? *Rusistika* 26, 7–14.

McShane, Marjorie J. 2002c. Unexpressed objects in Russian. *Journal of Slavic Linguistics* 10, 291–328.

McShane, Marjorie, and Sergei Nirenburg. 2002. Reference and ellipsis in ontological semantics. In *Memoranda in Computer and Cognitive Science,* MCCS-02-329. The Computing Research Laboratory, New Mexico State University.

McShane, Marjorie, Sergei Nirenburg, and Stephen Beale. 2004. OntoSem methods for processing semantic ellipsis. In *Proceedings from the HLT/NAACL Workshop on Computational Lexical Semantics*, East Stoudsburg, PA: ACL, 1–8.

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands and the Theory of Ellipsis*. New York: Oxford University Press.

Mithun, Marianne. 1988. The grammaticalization of coordination. In John Hainman and Sandra A. Thompson, eds., *Clause Combining in Grammar and Discourse.* Amsterdam: John Benjamins, 331–359.

Morgan, Jerry L. 1982. Discourse theory and the independence of sentence grammar. In Deborah Tannen, ed., *Analyzing Discourse: Text and Talk, Proceedings of the Georgetown University Round Table on Languages and Linguistics 1981.* Washington, D.C.: Georgetown University Press, 196–204.

Neeleman, Ad, and Hans van de Koot. 2001. Syntactic haplology. Unpublished paper available at http://www.phon.ucl.ac.uk/home/hans/haplology.html.

Neidle, Carol J. 1988. *The Role of Case in Russian Syntax.* Dordrecht: Kluwer Academic Publishers.

Neijt, Anneke. 1979. *Gapping: A Contribution to Sentence Grammar*. Dordrecht: Foris.

Neijt, Anneke. 1981. Gaps and remnants: Sentence grammar aspects of Gapping. *Linguistic Analysis* 8(1), 69–93.

Nillson, Barbro. 1982. *Personal Pronouns in Russian and Polish: A Study of Their Communicative Function and Placement in the Sentence*. Stockholm: Almqvist and Wiksell International. Translated from the Swedish by Charles Rougle.

Nirenburg, Sergei, and Lori Levin. 1992. Syntax-driven and ontology-driven lexical semantics. In James Pustejovsky and Sabine Bergler, eds., *Lexical Semantics and Knowledge Representation*, First SIGLEX Workshop Proceedings, Berkeley, Cal., June 17, 1991. Berlin, Germany: Springer-Verlag, 5–20.

Nirenburg, Sergei, and Victor Raskin. 2004. *Ontological Semantics*. Cambridge, Mass.: MIT Press.

Palmer, Frank R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

Panevova, Ja. 1992. O nekotoryx tipax obobščennyx aktantov. In *Festschrift Für Viktor Jul'evič Rozencvejg*. Vienna: 73–78. Wiener Slawistischer Almanach, Sonderband 33.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1972. *A Grammar of Contemporary English*. London: Longman.

Radford, Andrew. 1977. Counter-filtering rules. *York Papers in Linguistics* 7, 7–45.

Radford, Andrew. 1979. The functional basis of transformations. In *Transactions of the Philological Society*, B. Blackwell, 1–42.

Raposo, Eduardo. 1986. On the null object in European Portuguese. In Osvaldo Jaeggli and Carmen Silva-Corvalan, eds., *Studies in Romance Linguistics*. Dordrecht: Foris, 373–390.

Rappaport Hovav, Malka, and Beth Levin. 2002. Change of state verbs: Implications for

theories of argument projection. In *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society*, 269–280.

Rögnvaldsson, Eirikur. 1990. Null objects in Icelandic. In J. Maling and A. Zaenen, eds., *Syntax and Semantics*, volume 24 (Modern Icelandic Syntax). San Diego: Academic Press, 367–379.

Rooryck, Johan. 1985. Gapping-zeugma in French and English: A non deletion analysis. *Linguistic Analysis* 15, 187–229.

Ross, John R. 1967. *Constraints on Variables in Syntax*. Ph.D. Dissertation, MIT, Cambridge, Massachusetts.

Ross, John R. 1969. Guess who? In *Proceedings of the Sixth Regional Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press, 252–286.

Ross, John R. 1970. Gapping and the order of constituents. In Manfred Bierwisch and Karl E. Heidolph, eds., *Progress in Linguistics*. The Hague: Mouton, 249–259.

Rothstein, Robert A. 1980. Gender and reference in Polish and Russian. In Catherine V. Chvany and Richard D. Brecht, eds., *Morphosyntax in Slavic*. Columbus, Ohio: Slavica Publishers, 79–97.

Ružička, Rudolf. 1978. Three aspects of valence. In Werner Abraham, ed., *Valence, Semantic Case, and Grammatical Relations*, Workshop studies prepared for the 12th International Congress of Linguists, University of Vienna, August 29 to September 3, 1977. Amsterdam: John Benjamins, 47–53.

Sag, Ivan. 1976. *Deletion and Logical Form*. Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Sag, Ivan A., Gerald Gazdar, Thomas Wasow, and Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3, 117–172.

Saloni, Zygmunt. 1974. O możliwościach elipsy określenia czasownika w polskim tekście wielozdaniowym. In M. R. Mayenowa, ed., *Tekst i język: problemy semantyczne*. Wrocław: Wydawnictwo polskiej akademii nauk, 59–81.

Schank, Roger, and Robert Abelson. 1977. *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. New York: L. Erlbaum.

Sjoblom, Todd. 1980. *Coordination*. Ph.D. Dissertation, MIT, Cambridge, Massachusetts.

Somers, H. L. 1984. On the validity of the complement-adjunct distinction in valency grammars. *Linguistics* 22, 507–530.

Stillings, Justine T. 1975. The formulation of Gapping in English as evidence for variable types in syntactic transformations. *Linguistic Analysis* 1, 247–273.

Suñer, Margarita, and Maria Yépez. 1988. Null definite objects in Quiteño. *Linguistic Inquiry* 19, 511–519.

Thomas, Andrew L. 1979. Ellipsis: The interplay of sentence structure and context. *Lingua* 47, 43–68.

Topolińska, Zuzanna, ed. N.d. *Gramatyka współczesnego języka polskiego: Składania, Morfologia, Fonologia*. Warsaw: Państwowe wydawnictwo naukowe.

Tsujii, Jun-ichi, and Makoto Nagao. 1988. Dialogue translation vs. text translation: Interpretation based approach. In *Proceedings of the 12th International Conference on Computational Linguistics*. Budapest, August 22–27, 1988. Budapest: John von Neumann Society for Computing Sciences, 688–693.

Vardul', I. F. 1969. K voprosu o javlenii ellipsisa. In *Invariantnye sintaksičeskie značenija i struktura predloženija*. Moscow: Nauka, 59–70.

Vater, Heinz. 1978. On the possibility of distinguishing between complements and adjuncts. In Werner Abraham, ed., *Valence, Semantic Case, and Grammatical Relations*, Workshop studies prepared for the 12th International Congress of Linguists, University of Vienna, August 29 to September 3, 1977. Amsterdam: John Benjamins, 21–45.

Viegas, Evelyne, and Sergei Nirenburg. 1995. The semantic recovery of event ellipsis: Its computational treatment. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montreal, Quebec, Canada, Aug. 20–25, 1995. San Mateo, Calif.: International Joint Conferences on Artificial Intelligence, Inc.: Distributed by M. Kaufmann Publishers.

Wierzbicka, Anna. 1979. Ethno-syntax and the philosophy of grammar. *Studies in Language: International Journal Sponsored by the Foundation "Foundations of Language"* 3, 313–385.

Williams, Edwin S. 1977. Discourse and logical form. *Linguistic Inquiry* 8, 101–139.

Wilson, Peter. 2000. *Mind the Gap.* London: Longman.

Yamamoto, Kazuhide, and Eiichiro Sumita. 1998. Feasibility study for ellipsis resolution in dialogues by machine-learning technique. In *Proceedings of the 17th International Conference on Computational Linguistics* and *the 36th Annual Meeting of the Association for Computational Linguistics* (*COLING-ACL '98*), Aug. 10–14, 1998, Montreal, Quebec, Canada. New Brunswick, NJ: ACL; Morgan Kaufmann, 1428–1435.

Yeh, Ling-hsia. 1994. Haplology in Mandarin. Ms., UMass, Amherst.

Yip, Moira. 1998. Identity avoidance in phonology and morphology. In Steven G. LaPointe, Diane K. Brentari, and Patrick M. Farrell, eds., *Morphology and Its Relation to Phonology and Syntax*. Stanford, Cal.: CSLI Publications, 216–246.

Yokoyama, Olga T. 1986. *Discourse and Word Order*. Amsterdam: John Benjamins.

Zemskaja, Elena A., ed. 1973. *Russkaja razgovornaja reč'*. Moscow: Nauka.

Zemskaja, Elena A. 1987. *Russkaja razgovornaja reč': Lingvističeskij analiz i problemy obučenija*. Moscow: Russkij jazyk.

Zemskaja, Elena A., M. V. Kitajgorodskaja, and E. N. Širjaev, eds. 1981. *Russkaja razgovornaja reč'*. Akademija nauk SSSR, Institut russkogo jazyka. Moscow: Nauka.

Zolotova, Galina A. 1982. *Kommunikativnye aspekty russkogo sintaksisa*. Moscow: Nauka.

Literary Sources

Following the citation, in square brackets, is the means of referencing each work and, as applicable, its transliteration, the source language, and the original title.

Russian

Акунин, Борис. 2000. *Левиафан*. Изд. Захаров. [Акунин/Akunin]

Антология русского советского рассказа 60-е годы. Москва: Современник, 1989. [*Антология/Antologija*]

Азимов, Айзек. 1993. *Песчинка в небе*. Перевод с английского А.К. Андреева. Москва: Мир. [Азимов/Asimov; translation of the English *Pebble in the Sky*]

Богуславская, Зоя. 1991. *Американки*. Москва: Прогресс.[Богуславская/Boguslavskaja]

Брагинский, Эмиль, и Элдарь Рязанов. 1967. *Берегись автомобиля!* Москва: Искусство. [Брагинский и Рязанов/Braginskij i Rjazanov]

Вампилов, Александр. 1984. Прощание в июне. *Избранное*. Москва: Искусство, 5–69. [Вампилов 1/Vampilov 1]

Вампилов, Александр. 1984. Старший сын. *Избранное*. Москва: Искусство, 70–147. [Вампилов 2/Vampilov 2]

Вампилов, Александр. 1984. Утиная охота. *Избранное*. Москва: Искусство, 148–241. [Вампилов 3/Vampilov 3]

Вампилов, Александр. 1984. Прошлым летом в Чулимске. *Избранное*. Москва: Искусство, 301–377. [Вампилов 4/Vampilov 4]

Вампилов, Александр. 1984. Дом окнами в поле. *Избранное*. Москва: Искусство, 378–390. [Вампилов 5/Vampilov 5]

Войнович, Владимир. 1963. *Мы здесь живём*. Москва: Советский писатель. [Войнович 1/Vojnovič 1]

Войнович, Владимир. 1988. Шапка. London: Overseas Publications Interchange. [Войнович 2/Vojnovič 2]

Волков, Александр. 1985. *Волшебник изумрудного города*. Москва: Детская литература. [Волков/Volkov]

Достоевский, Фёдор М. 1987. Белые ночи. *Сочинения в двух томах, том I.* Москва: Правда, 115–162. [Достоевский 1/Dostoevskij 1]

Достоевский, Фёдор М. 1985. Мужик Марей. In Igor S. Mihalchenko, ed., *Russian Intermediate Reader.* Lincolnwood, Ill.: National Textbook Company, 53–58. [Достоевский 2/Dostoevskij 2]

Золотые слова. 1988. Москва: Русский язык. Составитель и автор комментария: Л. А. Белая. [Золотые/Zolotye]

Кристи, Агата. 1996. *Мисс Марпл. Детективы Агаты Кристи в сорока томах. Том 21.* Рига: Полярис. [Кристи/Christie, translation of the English *Murder at the Vicarage* and *The Body in the Library*]

Линдгрен, Астрид. 1992. *Пеппи Длинныйчулок*. Перевод со шведского Л. Лунгиной. Москва: Изд. Советско-Британского совместного предприятия Слово/Slovo. [Линдгрен/Lindgren; translation of the Swedish *Pippi Långstrump*]

Минчин, Александр. 1996. *Юджиния*. Москва: Новости. [Минчин/Minčin]

Пелевин, Виктор. 1999. *Чапаев и пустота*. Москва: Вагриус. [Пелевин/Pelevin]

Рачко, М. 1990. *Через не могу*. Tenafly, N.J.: Ермитаж. [Рачко/Račko]

Токарева, Виктория. 1994. *День без вранья*. Москва: СП ‘Квадрат’. [Токарева/Tokareva]

Толстой, Лев Н. 1993. *Анна Каренина,* Москва: Терра-Terra. [Толстой 1/ Tolstoj 1].

Толстой, Лев Н. 1973. Детство. *Детство, отрочество, юность.* Москва: Детская литература, 5–104. [Толстой 2/Tolstoj 2]

Уоллер, Роберт Джеймс. 1994. *Мосты округа Мэдисон.* Перевод с английского Е. Богдановой. Москва: РИПОЛ. [Уоллер/Waller, translation of the English *The Bridges of Madison County*]

Фолсом, Франклин. 1977. *Книга о языке*. Москва: Прогресс. [Фолсом/Folsom]

Фрэнсис, Дик. 1993. *Испытай себя*. Москва: Новости. [Фрлнсис/Francis, translation of the English *Longshot*]

Херриот, Джеймс. 1993. *Из воспоминаний сельского ветеринара*. Перевод с английского И. Гуровой. Москва: Мир. [Херриот/Herriot, translation of the English *If Only They Could Talk*, *It Shouldn’t Happen to a Vet*, *Let Sleeping Vets Lie*, *Vet in Harness*, *Vets Might Fly*, *Vet in a Spin*, *The Lord God Made Them All*]

Хмелевская, Иоанна. 1997. *2/3 успеха*. Перевод с польского В.С. Селивановой. Москва: Фантом Пресс Интер В.М. [Хмелевская 1/Chmielewska 1, translation of the Polish *2/3 sukcesu*].

Хмелевская, Иоанна. 1998. *Версия про запас*. Перевод с польского Е.А. Полецкой, Москва: Фантом Пресс. [Хмелевская 2/Chmielewska 2, translation of the Polish *Drugi watek*].

Хмелевская, Иоанна. 1997. *Клин клином*. Перевод с польского Л. Ермиловой. Москва: Фантом Пресс Интер В.М. [Хмелевская 3/ Chmielewska 3, translation of the Polish *Klin*].

Хмелевская, Иоанна. 1997. *Свистопляска*. Перевод с польского Н.В. Селивановой.

Москва: Фантом Пресс Интер В.М. [Хмелевская 4/ Chmielewska 4, translation of the Polish *Duża polka*].
Хмелевская, Иоанна. 2001. *Колодцы предков*. Москва: Фантом Пресс Интер. [Хмелевская 5/ Chmielewska 5, translation of the Polish *Studnie przodków*]
Чехов, Антон П. 1967. Вишнёвый сад. *Извранные произведения в трёх томах*, *том III.* Москва: Художественная литература, 555–606. [Чехов 1/Čexov 1]
Чехов, Антон П. 1967. Чайка. *Извранные произведения в трёх томах, том III*. Москва: Художественная литература, 391–442. [Чехов 2/Čexov 2]
Шварц, Евгений. 1998. Голый король. *Дракон, Клад, Тень, Два клёна, Обыкновенное чудо и другие произведения*. Москва: Гудьял пресс, 55–130. [Шварц 1/Švarc 1]
Шварц, Евгений. 1998. Дракон, 269–339. *Дракон, Клад, Тень, Два клёна, Обыкновенное чудо и другие произведения.* Москва: Гудьял пресс. [Шварц 2/Švarc 2]
Шварц, Евгений. 1998. Золушка. *Дракон, Клад, Тень, Два клёна, Обыкновенное чудо и другие произведения.* Москва: Гудьял пресс, 533–571. [Шварц 3/Švarc 3]
Шварц, Евгений. 1998. Снежная королева. *Дракон, Клад, Тень, Два клёна, Обыкновенное чудо и другие произведения.* Москва: Гудьял пресс, 131–194. [Шварц 4/Švarc 4]
Шварц, Евгений. 1998. Повесть о молодых супругах. *Дракон, Клад, Тень, Два клёна, Обыкновенное чудо и другие произведения*. Москва: Гудьял пресс, 463–531. [Шварц 5/Švarc 5]
Шинов, Чавдар. 1987. *Похищение автобуса.* Translated by Majja Kačaunova. Sofia: Svjat. [Шинов/Šinov]

Polish

Broszkiewicz, Jerzy. 1976. Imiona władzy. *Antologia Dramatu*, 279–337. Warsaw: Państwowy Instytut Wydawniczy. [Broszkiewicz]
Conan Doyle, Arthur. 1995. *Pies Baskerwille'ów*. Wróclaw: Wydawnictwo Dolnośląskie. [Conan Doyle, translation of the English *The Hound of the Baskervilles*]
Fitzgerald, F. Scott. 1994. *Wielki Gatsby*. Tłumaczył Jędrzej Polak. Poznań: SAWW, 1994. [Fitzgerald, translation of the English *The Great Gatsby*]
Galsworthy, John 1992. *Nowoczesna komedia*. Tłumaczył Jerzy T. Czaplicki. Warsaw: Da Capo. [Galsworthy, translation of the English *The Modern Comedy*].
Garwood, Julie. 1996. *Pożądanie*. Przełożyła Małgorzata Kicana. Warsaw: Wydawnictwo da Capo. [Garwood, translation of the English *Rebellious Desire*]
Herbert, Zbigniew. 1976. Drugi pokój. *Antologia Dramatu.* Warsaw: Państwowy Instytut Wydawniczy. [Herbert]
Korczak, Janusz. 1923. *Król Maciuś pierwszy.* Bielsko-Biała: Beskidzka Oficyna Wydawnicza. [Korczak]
Kruczkowski, Leon. 1973. Niemcy. *Dramaty*, 5–76. Warsaw: Państwowy Instytut Wydawniczy. [Kruczkowski 1]
Kruczkowski, Leon. 1973. Pierwszy dzień wolności. *Dramaty*, 77–162. Warsaw: Państwowy Instytut Wydawniczy. [Kruczkowski 2]
Mrożek, Sławomir. 1989. *Tango, Słoń, Wesele w Atomicach, Woda*. Warsaw: Młodzieżowa Agencja Wydawnicza. [Mrożek]
Olesky, Wiesław, and Oscar E. Swan. 1993. *W labiryncie.* Columbus, Ohio: Slavica Publishers. [Olesky and Swan]
Szaniawski, Jerzy. 1961. *Profesor Tutka*. Kraków: Wydawnictwo Literackie. [Szaniawski 1]
Szaniawski, Jerzy. 1976. Dwa teatry. *Antologia Dramatu*. Warsaw: Państwowy Instytut Wydawniczy, 5–58. [Szaniawski 2]

This page intentionally left blank

INDEX