Gradability in Natural Language
OXFORD STUDIES IN SEMANTICS AND PRAGMATICS

GENERAL EDITORS: Chris Barker, New York University, and Chris Kennedy, University of Chicago

PUBLISHED
1. Definite Descriptions
   Paul Elbourne
2. Logic in Grammar
   Polarity Free Choice, and Intervention
   Gennaro Chierchia
3. Weak Island Semantics
   Mártiá Abrusán
4. Reliability in Pragmatics
   Eric McCready
5. Numerically Qualified Expressions
   Christopher Cummins
6. Use-Conditional Meaning
   Studies in Multidimensional Semantics
   Daniel Gutzmann
7. Gradability in Natural Language
   Logical and Grammatical Foundations
   Heather Burnett
8. Subjectivity and Perspective in Truth-Theoretic Semantics
   Peter Lasersohn

IN PREPARATION

Meaning over Time
The Foundations of Systematic Semantic Change
Ashwini Deo
Measurement and Modality
Daniel Lassiter
Plural Reference
Friederike Moltmann
The Semantics of Evidentials
Sarah E. Murray
A History of Formal Semantics
Barbara Partee
Gradability in Natural Language

Logical and Grammatical Foundations

HEATHER BURNETT
To my parents
Contents

General preface x
Acknowledgments xi
List of abbreviations xiv

1 Introduction 1
   1.1. Organization of the book 6

2 Vagueness and linguistic analysis 9
   2.1. Introduction 9
   2.2. Our Classical Semantic Theory 11
      2.2.1. Classical First order logic 11
      2.2.2. Extensions in linguistics 15
   2.3. The phenomenon of vagueness 16
      2.3.1. Borderline cases 17
      2.3.2. Fuzzy boundaries 19
      2.3.3. The Sorites paradox 20
   2.4. Tolerant, Classical, Strict 22
      2.4.1. Definition 22
      2.4.2. Account of the puzzling properties 27
   2.5. Lasersohn’s (1999) Pragmatic Halos 29
      2.5.1. Definition 29
      2.5.2. Comparison with Tolerant, Classical, Strict 32
   2.6. Conclusion 33

3 Context sensitivity and vagueness patterns 34
   3.1. Introduction 34
   3.2. Adjectival context sensitivity patterns 37
   3.3. Universal vs existential context sensitivity 45
   3.4. Potential vagueness and adjectival vagueness patterns 48
      3.4.1. (A)symmetric vagueness 50
   3.5. Conclusion 54

4 The Delineation Tolerant, Classical, Strict framework 55
   4.1. Introduction 55
   4.2. Language and classical semantics 56
4.2.1. Classical semantics for relative adjectives
4.2.2. Classical semantics for absolute/non-scalar adjectives
4.2.3. The paradox of absolute scalar adjectives
4.3. Tolerant/strict semantics
4.3.1. The properties of indifference
4.4. Predictions of the DelTCS analysis
4.4.1. Context sensitivity results
4.4.2. Gradability results
4.4.3. Potential vagueness results
4.4.4. Other empirical consequences
4.5. Conclusion
4.6. Appendix: Longer proofs

5 Scale structure in Delineation semantics
5.1. Introduction
5.2. Scale structure patterns
5.2.1. Non-scalar adjectives
5.2.2. Summary of scale structure data
5.3. Scale structure in Delineation semantics
5.4. Conclusion

6 Beyond Delineation semantics
6.1. Introduction
6.2. Scale structure in Degree semantics
6.2.1. Kennedy’s (2007) Degree analysis
6.2.2. Comparison between Kennedy (2007) and DelTCS
6.3. Interpretive Economy and Bayesian pragmatics
6.3.1. Bayesian pragmatics
6.3.2. Adjectival interpretation
6.3.3. Summary
6.4. Degree Tolerant, Classical, Strict
6.4.1. Results
6.5. Conclusion

7 Beyond the adjectival domain
7.1. Introduction
7.2. Context sensitivity and vagueness patterns
7.2.1. Summary
7.3. Mereological Delineation Tolerant, Classical, Strict
7.3.1. Language
7.3.2. Classical semantics
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3.3. Tolerant/strict semantics</td>
<td>170</td>
</tr>
<tr>
<td>7.3.4. Summary</td>
<td>178</td>
</tr>
<tr>
<td>7.4. Definite plural DPs and maximality</td>
<td>178</td>
</tr>
<tr>
<td>7.4.1. Language and classical semantics</td>
<td>178</td>
</tr>
<tr>
<td>7.4.2. Tolerant/strict semantics</td>
<td>181</td>
</tr>
<tr>
<td>7.4.3. <em>Gather</em> predicates vs <em>numerous</em> predicates</td>
<td>187</td>
</tr>
<tr>
<td>7.4.4. Negation and homogeneity with definite plurals</td>
<td>193</td>
</tr>
<tr>
<td>7.5. Conclusion</td>
<td>194</td>
</tr>
<tr>
<td>7.6. Appendix: Longer proofs</td>
<td>195</td>
</tr>
<tr>
<td>8 Conclusion</td>
<td>198</td>
</tr>
</tbody>
</table>

References  
Index
General preface

Oxford Studies in Semantics and Pragmatics publishes original research on meaning in natural language within contemporary semantics and pragmatics. Authors present their work in the context of past and present lines of inquiry and in a manner accessible to both scholars whose core areas of expertise are in linguistic semantics and pragmatics, and to researchers in related and allied fields such as syntax, lexicology, philosophy, and cognitive science. The series emphasizes rigorous theoretical analysis grounded in detailed empirical investigation of particular languages.

This is a companion series to Oxford Surveys in Semantics and Pragmatics. The Surveys series provides critical overviews of the major approaches to core semantic and pragmatic phenomena, a discussion of their relative value, and an assessment of the degree of consensus that exists about any one of them. The Studies series equally seeks to put empirical complexity and theoretical debate into comprehensible perspective, but with a narrower focus and correspondingly greater depth. In both series, authors develop and defend the approach and line of argument that they find most convincing and productive.

In this contribution to the series, Heather Burnett investigates the semantics of gradability: the principles of meaning that provide the basis for expressions of comparison, intensification, and degree. Burnett’s central thesis is that gradability can be derived from the interaction of a set of more basic, and conceptually necessary, principles governing the interpretation of context-dependent expressions and the tolerance of categorization processes. This approach contributes to a long tradition in semantics and philosophy that seeks to link the semantics of gradability to a general theory of linguistic vagueness; where Burnett pushes beyond previous work is in developing a formal framework that can explain grammatical distinctions between classes of gradable predicates that have previously been thought to require reference to an additional abstract scalar representation system. This volume challenges this view and provides a fresh and sophisticated alternative to contemporary theories of the semantics of gradability.
Acknowledgments

This work would never have been possible without all the help and support that I have received from friends and colleagues during my time as a doctoral student in the Department of Linguistics at UCLA and as a Social Sciences and Humanities Research Council (SSHRC) postdoctoral fellow in the Département de linguistique et de traduction at l’Université de Montréal.

Although credit is due to many many people, I have to single out Paul Égré for the enormous contributions that he has made to both the content of this monograph and my personal and professional development. I thank him for so many things, including but not limited to: introducing me to the exciting world of non-classical logics, showing me how to fill in the CNRS application form, laughing at my jokes (even the ones making fun of philosophers, Normaliens and French people), showing me how to eat a hamburger French-style (i.e. with a knife and fork) and letting me sleep in his kid’s room when I was homeless in New York.¹

More generally, I thank the members of the Institut Jean Nicod at the Ecole normale supérieure in Paris (particularly members of the LIN-GUAE group: Emmanuel Chemla, Vincent Homer, Philippe Schlenker, Benjamin Spector, and Jérémy Zehr, as well as Claire Beyssade, Francis Corblin, Alda Mari, David Nicolas, and François Récanati) for welcoming me both as a student and as a visiting postdoctoral researcher. The time that I have spent at Jean Nicod has been incredibly rewarding (both academically and personally), and the influence of the ideas being developed in this lab can be clearly seen in the major themes explored in the book.

The first part of this book (chapters 2–5) is based on my 2012 dissertation *The Grammar of Tolerance: On vagueness, context-sensitivity and the origin of scale structure*, which was completed in the Linguistics department at UCLA. My supervisors, Ed Keenan and Dominique Sportiche, as well as Hilda Koopman, Jessica Rett, Yael Sharvit, and Ed Stabler have all made innumerable vital contributions to this project, and I will

¹ Credit for this last one is, of course, also due to Rachida, Amir, and, above all, to my roommate Isaac.
be forever grateful for their expertise, mentorship, and friendship in semantics and beyond.

The work that I began as a graduate student, I continued as a postdoc within the context of an SSHRC postdoctoral fellowship at l’Université de Montréal. In addition to being one of my closest collaborators and friends, Mireille Tremblay was the ideal postdoctoral supervisor, giving me guidance, but also an enormous amount of freedom to explore new interesting ideas and then to chase them across continents. This work never would have been possible without her.

During my postdoctoral work, I also benefitted greatly from shorter research stays in stimulating linguistics departments around the world. In the winter of 2013, I spent a very productive month working with Louise McNally and the rest of the GLiF group at the Universitat Pompeu Fabra. I would like to thank Louise, Gemma Barberà, Berit Gehrke, Scott Grimm, Laia Mayol, and Isidora Stojanovic for talking to me about adjectives and showing me how to grill and eat calçots. I also spent a very exciting couple of months at NYU in 2014, and I would like to especially thank my host, Chris Barker, as well as Dylan Bumford, Simon Charlow, Lucas Champollion, Orin Percus, Cara Shousterman, and Anna Szabolcsi for many helpful academic and non-academic discussions. Salvador Mascarenhas also gets a special mention for being my linguistics/philosophy drinking buddy in both Paris and New York . . . and for coming to rescue me at the 79th Police Precinct when I got caught up in tensions associated with the rapid gentrification of Brooklyn.

As I mentioned, many of the main ideas developed in the first part of the book come from my doctoral dissertation, and earlier versions of the proposals outlined here have been published in the journals Linguistics & Philosophy (37:1–39) and the Journal of Applied Non-Classical Logics (24:35–60). Likewise, Chapter 7 takes up and extends some of the proposals found in the chapter ‘Vague determiner phrases and distributive predication’ (which appears in Marija Slavkovik and Dan Lassiter (eds.) New Directions in Logic, Language, and Interaction. Springer: FoLLI Lecture Notes in Computer Science 7415. pp. 175–194.) Likewise, some of the discussion concerning logical theories of vagueness and linguistics found in chapter 2 is also taken up in the (joint) handbook chapter (with Peter Sutton) Vagueness and natural language semantics (under review). Furthermore, during the writing of my dissertation and this book, I have been generously supported by the Social Sciences and Humanities Research Council of Canada (Doctoral fellowship (#752-2007-2382) and
Postdoctoral fellowship (#756-2012-0045), the USA/France Partner University Fund grant (Theoretical/Experimental Linguistic Cognition Advanced Studies grant (UCLA/ENS, Paris)), and by the UCLA Department of Linguistics.

I have also been very fortunate to have had the chance to interact with a very large number of scholars in the areas of semantics, pragmatics, logic, and the philosophy of language. The long list of researchers that I have corresponded with about this material (and whose comments and suggestions have greatly improved the manuscript) includes, but is by no means limited to: Natasha Abner, Alan Bale, Mel Bervoets, Rajesh Bhatt, Denis Bonnay, John Burgess, Daniel Büring, Pablo Cobreros, Jenny Doetjes, Itamar Francez, Brendan Gillon, Thomas Graf, Meg Grant, Volker Halbach, Irene Heim, Greg Kobele, Manuel Kriš, Dan Lassiter, Eliot Michaelson, Friederike Moltmann, Rick Nouwen, Paul Pietroski, Walter Pedersen, Martin Prinzhorn, Dave Ripley, Galit Sassoon, Viola Schmitt, Roger Schwarzschild, Stephanie Solt, Alexis Wellwood, Yoad Winter, and Elia Zardini. Furthermore, I must single out the contributions of Robert van Rooij to this project. Some of the ideas found in this work were sparked by his class on vagueness at the 2011 ESSLLI in Ljubljana (co-taught with Pablo Cobreros, Dave Ripley, and Paul Égré) and his razor-sharp comments and critiques greatly improved both the form and the presentation of the DelTCS framework. I also particularly thank Chris Kennedy for his very detailed and insightful comments on versions of this manuscript, which greatly strengthened both the empirical and theoretical contributions of this work.

Finally, I thank Shiri Lev-Ari, Camilla Gibb, and Olivia Conway-Gibb for supporting me through finishing various versions of this work.

And, above all, I thank my parents and my sister Kate for all the love, support, and inspiration that they have given me over the course of my life.
List of abbreviations

AA  absolute scalar adjective
AA$^p$ partial absolute scalar adjective
AA$^T$ total absolute scalar adjective
AAA  absolute adjective axiom
BP  Be precise
CC  comparison class
CEM  classical extensional mereology
C-model  Classical model
CP  Contrast Preservation
CST  Classical Semantic Theory
DegP  Degree phrase
DegS  Degree semantics
DegTCS  Degree Tolerant, Classical, Strict
DelS  Delineation semantics
DelTCS  Delineation Tolerant, Classical, Strict
DD  Downward difference
DP  determiner phrase
FOL  First order logic
G  Granularity
I  Incomparability
IE  Interpretive Economy
LP  Logic of Paradox
MC  Mereological Convexity
MD  Minimal Difference
M-DelTCS  Mereological Delineation Tolerant, Classical, Strict
MI  Montagovian Individual
MSH  Mereological Structure Hypothesis
NR  No Reversal
NS  non-scalar adjective
P-vague potentially vague
PH  Pragmatic Halos
PA  Partial Axiom
RA  relative scalar adjective
S  Symmetry
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>scalar adjective</td>
</tr>
<tr>
<td>SC</td>
<td>Strict Convexity</td>
</tr>
<tr>
<td>SP</td>
<td>Shared Parts</td>
</tr>
<tr>
<td>SWO</td>
<td>strict weak order</td>
</tr>
<tr>
<td>TA</td>
<td>Total Axiom</td>
</tr>
<tr>
<td>TC</td>
<td>Tolerant Convexity</td>
</tr>
<tr>
<td>TCS</td>
<td>Tolerant, Classical, Strict</td>
</tr>
<tr>
<td>T-model</td>
<td>Tolerant model</td>
</tr>
<tr>
<td>UD</td>
<td>Upward Difference</td>
</tr>
<tr>
<td>VP</td>
<td>verb phrase</td>
</tr>
<tr>
<td>wff</td>
<td>well-formed formula</td>
</tr>
</tbody>
</table>
1

Introduction

This book presents a new theory of the relationship between vagueness, context sensitivity, and scale structure in natural language. In particular, this work is devoted to the description and analysis of the distribution of these phenomena within and outside the adjectival domain of English and other Indo-European languages.

A more precise and developed exposition of the phenomenon known as vagueness will be given in Chapter 2; however, we can illustrate some of the puzzles that it raises with the following example: Suppose we take someone who is 1.9 m tall, and suppose that we agree, because we are talking about average male heights, he is tall. Furthermore, suppose that we have a long line of people ordered based on height and that their heights differ by only 1 cm each. The 1.9 m tall man is at the front of the line, and there is someone who is only 1.5 m tall at the end. We can agree that the last person is not tall. Given this set-up, there must be some point in this line at which we move from a tall person to his not tall follower, who is 1 cm shorter than he is. But where is this point? Since adding or subtracting a single centimetre is such a small change, it seems absurd to think that changing someone’s height by this much could ever serve to affect whether or not we would call them tall. We call relations like ‘± 1 cm’ (in this context) tolerance relations or indifference relations, since they encode amounts of change that do not make a difference to categorization. When we can find a tolerance relation for an adjective, we call the adjective tolerant, i.e. we call tall a tolerant predicate because statements like (1) seem true.

(1) For all \(x, y\), if \(x\) is tall and \(x\) and \(y\)’s heights differ by at most 1 cm, then \(y\) is also tall.

Note, furthermore, that the negation of tall (not tall) is also tolerant: in a context such as the one described above, (2) also seems true.
(2) For all \( x, y \), if \( x \) is not tall and \( x \) and \( y \)’s heights differ by at most 1 cm, then \( y \) is also not tall.

Clearly the fact that both \textit{tall} and \textit{not tall} are tolerant creates a puzzle: why do we not conclude that both the 1.9 m man and the 1.5 m man are tall and not tall at the same time? Paradoxes of this type are known as \textit{Sorites} paradoxes,\textsuperscript{1} and they will be discussed in much greater detail throughout the book.

Another adjective that shows a similar pattern is \textit{straight}: In most situations, adding a 1/10 mm bend to a stick is such an irrelevant change that it will never be sufficient to make a stick that we call \textit{straight} not called \textit{straight}. Thus, if we were to line up a set of sticks that differ by 1/10 mm bend from the perfectly straight ones to the really bendy ones, then (3) seems true.

(3) For all \( x, y \), if \( x \) is straight and \( x \) and \( y \) differ by a single 1/10 mm bend, then \( y \) is also straight.

However, unlike \textit{tall}, whose negation is also tolerant, even though adding or subtracting a 1/10 mm bend is such a small change, the corresponding statement with \textit{not straight} is false: in particular (4) is falsified by the case where we move from \( x \) that has a 1/10 mm bend (so is not straight) to \( y \) that has absolutely no bends.

(4) False: For all \( x, y \), if \( x \) is not straight and \( x \) and \( y \) differ by a single 1/10 mm bend, then \( y \) is also not straight.

In summary, on the one hand, adjectives like \textit{tall} and \textit{straight} are both tolerant, but on the other, \textit{straight} displays an asymmetry that \textit{tall} does not.

The second phenomenon that will be treated in this work is context sensitivity. To be more specific, we will call a predicate \( P \) \textit{context sensitive} just in case, for some individual \( x \), we can find a context in which \( P \) applies to \( x \), and we can find another context in which \( P \) does not apply to \( x \), without changing the properties of \( x \) and \( y \). The adjectives \textit{tall} and \textit{straight} both have this property: someone who can be considered \textit{tall} when we are considering jockeys might not be considered \textit{tall} when we

\textsuperscript{1} The name of these puzzles comes from a puzzle attributed to Eubelides of Miletus known as “the Heap” (\textit{soros} being Greek for \textit{heap}):

Would you describe a single grain of wheat as a heap? No. Would you describe two grains of wheat as a heap? No . . . You must admit the presence of a heap sooner or later, so where do you draw the line? (from the Stanford Encyclopedia of Philosophy.)
are considering average men. Likewise, we saw above that an object with a very small bend can be sometimes considered to be straight; however, in a context in which very slight bends make a large difference to our purposes, the very same object would not be considered straight.2

This being said, tall and straight display a different pattern when it comes to being context sensitive. For example, as discussed in Kennedy (2007) and Syrett et al. (2010) (among others), adjectives like tall can shift their criteria of application across contexts in a way that adjectives like straight cannot. If I have two objects, one of which is (noticeably) taller than the other, but neither are particularly tall, I can still use the predicate tall to pick out the taller of the two.

(5) Pass me the tall one.
   OK: even if neither/both is/are tall.

However, using straight in such a linguistic construction is only possible if exactly one of the two is (very close) to perfectly straight.

(6) Pass me the straight one.
   # if neither/both is/are straight.

The third phenomenon treated in this work is scalarity. Again, tall and straight pattern alike on this dimension in that they can both appear in the comparative and many other degree constructions (7).

(7) a. This stick is taller/straighter than that one.
   b. This stick is very tall/straight.

However, once more, if we look at the full range of data concerning gradability and scale structure, tall and straight show a different pattern: for example, certain scalar modifiers like almost and completely are natural with straight, but not with tall.

(8) a. ??John is almost/completely tall.
   b. This stick is almost/completely straight.

The main goal of Chapters 2–5 is to develop an account of both the similarities and differences between various subclasses of adjectives with respect to each of these three phenomena (vagueness, context sensitivity, and scalarity). The principle subclasses that will be empirically distinguished are the following:

2 Consider, for example, the barrel of a rifle that must be perfectly straight for our shots to be accurate.
(9) **Relative scalar adjectives (RAs):**
e.g. tall, short, expensive, cheap, nice, friendly, intelligent, stupid, narrow, wide

(10) **Total absolute adjectives (AA^T)s:**
e.g. empty, full, clean, smooth, dry, straight, flat

(11) **Partial absolute adjectives (AA^P)s:**
e.g. dirty, bent, wet, curved, crooked, dangerous, awake

(12) **Non-scalar adjectives (NSs):**
e.g. atomic, geographical, polka-dotted, pregnant, illegal, dead, hexagonal

I propose that the patterns concerning the behavior of *tall* and *straight* described above and other patterns to be discussed in the work are all reflexes of a single underlying difference in the semantics of these lexical items involving (a certain kind of) context sensitivity. Moreover, I propose that the data concerning both vagueness and scale structure can be derived from the interaction between (lack of) context sensitivity and tolerance/indifference relations associated with general cognitive categorization processes. Building on insights into the connection between context sensitivity and scalarity from the work of Klein (1980) (among others) and insights into the connection between tolerance relations and the Sorites paradox from the work of Cobreros et al. (2012b) (among others), I propose a new logical framework called Delineation Tolerant, Classical, Strict (DelTCS) that captures the intimate and complex relationship between these three aspects of adjectival meaning.

Chapters 6 and 7 look at extensions of the framework developed in Chapters 2–5. An important class of the extensions we will look at concerns how the DelTCS framework can be applied outside the adjectival domain to develop an analysis of context sensitivity, vagueness, and gradability patterns associated with constituents of the determine phrase (DP) category. It has been long observed that there exist important parallels between certain kinds of adjectives and certain kinds of DPs when it comes to vagueness and scale structure. Although these parallels will be outlined in great detail in the book, we can observe a first cross-domain parallel using (among other constructions) degree modifiers that combine with constituents of different syntactic categories. For example, in many languages, a universal scalar modifier, such as French

---

3 As observed by Bolinger (1972) and Moltmann (1997), similar adjectival patterns are possible in some dialects of English (ex. *This room is all empty* ≈ *This room is completely empty*)
tou(te)s ‘all’, can combine with both adjectives like droit ‘straight’ and definite plural DPs like les filles ‘the girls’ to create a parallel maximizing interpretation.

(13) a. La rue est toute droite.
    The road is all straight
    ‘The road is completely straight.’

b. Toutes les filles sont arrivées.
    All the girls are arrived
    ‘All the girls arrived.’

Although examples such as (13) (and others to be discussed in Chapter 7) suggest that definite plural DPs and total adjectives have similar scale structure properties, we will also see that context sensitivity/vagueness/scale structure have slightly different manifestations with DPs than with adjectives. In particular, I will argue that we see a different typology of context sensitivity, vagueness, and scale structure patterns in the DP domain than in the adjectival domain. Chapters 6 and 7 are therefore devoted to capturing both the similarities (such as (13)) and the proposed differences between adjectival and DP constituents within a mereological extension (Simons, 1987; Hovda, 2008, among others) of the DelTCS system (called M-DelTCS).

I will propose that the scales associated with DPs are derived from statements about their context sensitivity and vagueness in the same basic way as with adjectives. This is what creates the observed cross-domain parallels. However, I will also propose that the different kinds of ontological relations that characterize the domains into which DPs and adjectives denote have important consequences for how the application of these constituents can vary across comparison classes and how they display the characteristic properties of vague language. In other words, by virtue of the fact that DP constituents are interpreted into domains that have mereological (i.e. part-structure) relations on them, their context sensitivity and vagueness is constrained in a way that the context sensitivity and vagueness of adjectival constituents is not. In turn, by virtue of the logical structure of the M-DelTCS framework, these differences in context sensitivity and vagueness will be translated into differences in scale structure. Based on these results, I conclude that

empty). However, adjectival all is not fully productive in English in the way that its counterparts in the Romance languages (or even in German) are. Indeed, Moltmann (1997) refers to English all as ‘deficient’ with respect to its cognates in other Indo-European languages.
the Delineation TCS system (and its mereological extension M-DelTCS) provides a broad and versatile framework for analyzing the connections between context sensitivity, vagueness, and gradability that we observe in natural language phenomena across different syntactic categories.

1.1 Organization of the book

Chapter 2 (Vagueness and linguistic analysis) serves as an introduction to one of the main empirical phenomena to be analyzed in the monograph and the formal tools that will be used in the analysis. As such, it has two main parts: in the first part (Sections 2.2–2.3), I present the empirical phenomenon known as vagueness in the linguistics and philosophical literatures, and I outline why this phenomenon appears so threatening to our classical semantic theories in logic and linguistics. In the second part of the chapter (Sections 2.4–2.5), I present the basic account of the puzzling properties of vague language that I will adopt in this work: Cobreros et al.’s (2012b) Tolerant, Classical, Strict (TCS) similarity-based non-classical logical framework. I then present a similar framework that has been very influential in linguistics: Lasersohn’s (1999) Pragmatic Halos framework. I give a comparison between the two approaches and argue that, while Cobreros et al.’s (2012b) analysis (as applied to the interpretation of English) is empirically superior, they share many of the same driving intuitions. I therefore suggest that one way of looking at TCS is as a more nuanced version of the halos approach. Readers that are already familiar with the puzzles associated with vagueness and their proposed solutions within TCS can easily skip this chapter without any consequences for their understanding of the rest of the book.

Chapter 3 (Context sensitivity and vagueness patterns) presents the main empirical patterns associated with adjectival context sensitivity and vagueness. In line with previous work on the topic, I argue that the different scale structure classes of adjectives in (9)–(12) vary with respect to comparison class-based context sensitivity. I argue that to properly understand this variation, it is useful to adopt two patterns of comparison class-based context sensitivity: (what I will call) universal context sensitivity and existential context sensitivity. Intuitively, predicates that are universally context sensitive show a greater range of meaning variation than predicates that are existentially context sensitive. In this chapter, I argue that three of the four scale-structure subclasses presented above can be distinguished based on their context
sensitivity: RAs are universally context sensitive, both partial and $\text{AA}^T$s are existentially context sensitive, and NSs are not context sensitive. This chapter also motivates an important empirical connection between vagueness (i.e. the appearance of the properties described in Chapter 2) and the scale structure classes in (9)–(12). In particular, I show that the distribution of the puzzling properties of vague language is tied to these lexical class distinctions, and, I propose, following authors such as Kennedy and McNally (2005) and Kennedy (2007), that the observed dependencies argue in favor of a closer relationship between the phenomena of vagueness and scale structure than is often assumed in the literature.

Chapter 4 (The Delineation Tolerant, Classical, Strict framework) presents the DeTCS non-classical logical system for modelling the relationship between context sensitivity, vagueness, and gradability in the adjectival domain. I give an analysis of the context sensitivity/vagueness patterns described in Chapter 3 within this framework, and I discuss the empirical predictions that my analysis makes for a wide range of semantic and pragmatic phenomena associated with adjectival predicates.

Chapter 5 (Scale structure in Delineation semantics) presents both new and previously discussed data associated with the scale structure of members of the four principle classes of adjectives that are studied in this work. Following much previous research, I argue that the adjectives in each of the classes shown at the beginning of this chapter are associated with scales that have different properties. In particular, as we will see, there are empirical arguments for proposing that absolute total adjectives are associated with scales that have maximal elements, absolute partial adjectives are associated with scales that have minimal elements, and relative adjectives are associated with scales that have neither minimal nor maximal elements. I show in this chapter that the association of an adjective with a scale with the correct properties is already predicted by the analysis presented in Chapter 4 set within the DeTCS architecture. In other words, I argue that, once we have an (independently necessary) analysis of context sensitivity and vagueness in the adjectival domain, we get an analysis of adjectival scale structure “for free.”

Chapter 6 (Beyond Delineation semantics) studies a first class of extensions of the framework developed in Chapters 2–5: theoretical/formal extensions. More precisely, this chapter explores to what
extent the enrichment of current theories of the semantics of gradable expressions, besides Delineation semantics (DelS), with the structure of a non-classical logic such as Tolerant, Classical, Strict can be useful to understanding the complex relationships between context sensitivity, vagueness, and scale structure that were discussed in Chapters 2–5. I focus particularly on the detailed comparison of the DelTCS framework and current analyses of the absolute/relative distinction set within Degree semantics (DegS), with particular attention to the differences and similarities between DelTCS and (a parallel formalization of) the account presented in Kennedy (2007). I argue that, while these two proposals do differ in important ways, significant, non-trivial (and not immediately obvious) similarities exist between them. Based on this observation, I propose that it is worthwhile to study how Kennedy’s DegS proposal could be extended within the structure of the TCS framework, and I give such an extension, called DegTCS. I argue that the DegTCS extension provided in this chapter allows not only for a more direct comparison between Delineation and Degree frameworks, but also allows for a new characterization of Kennedy’s (2007) important Interpretive Economy principle.

Chapter 7 (Beyond the adjectival domain) studies a second class of extensions of the framework: empirical extensions. More specifically, this chapter presents new data concerning the distribution of vagueness, context sensitivity, and scale structure properties outside the adjectival domain. I argue that three of the four principle “scale structure” classes of adjectives have analogues in the DP domain and that these subclasses of DPs can be distinguished in a similar (although not identical) manner to their counterparts in the adjectival domain through their vagueness, context sensitivity, and scale structure properties. I give a mereological extension of the DelTCS framework and show how statements concerning DP vagueness and context sensitivity within the new framework can be used to derive the mereologically based scales that are independently proposed to be associated with various kinds of DPs.

Finally, Chapter 8 (Conclusion) summarizes the main empirical and theoretical claims made by this work and draws some final conclusions on the nature of vagueness and scalarity in natural language.
2

Vagueness and linguistic analysis

2.1 Introduction

This chapter serves as an introduction to both one of the main empirical phenomena to be analyzed in this monograph and the formal tools that will be used in the analysis. As such, it has two main parts: in the first part (Sections 2.2–2.3), I present the empirical phenomenon known as vagueness in the linguistics and philosophical literatures, and I outline why this phenomenon appears so threatening to our classical semantic theories in logic and linguistics. Although the puzzles associated with vague language have received an enormous amount of attention in the field of philosophy, they have been much less studied from a grammatical perspective. Therefore, in the first part of the chapter, I describe the ways in which vague predicates challenge the currently dominant approaches to natural language semantics. Thus, I argue that the problem of accounting for vagueness is also a central problem for the field of formal linguistics.

In the second part of the chapter (Sections 2.4–2.5), I present the basic account of the puzzling properties of vague language that I will adopt in this work: Cobreros et al.’s (2012b) Tolerant, Classical, Strict (TCS) similarity-based non-classical logical framework. Unlike many other works on this topic, I will not begin by reviewing all the many and varied previous accounts of vague language, nor will I provide a comprehensive comparison between the TCS approach and its competitors. There are two reasons for this: firstly, excellent general introductions to the phenomenon of vagueness and the wide variety of approaches on the market already exist.1 Secondly, and more importantly, many of the debates in the philosophical literature that have given rise to the wide range of

---

1 See, for example, Keefe (2000), Chapter 2 of Smith (2008), the papers in Dietz and Moruzzi (2010) etc. See also van Rooij (2010), Cobreros et al. (2012b), and Cobreros et al. (2012a) for comparisons between TCS and other frameworks.
theories of vagueness are not particularly relevant for linguistics. For example, a major issue with respect to which philosophers tend to differ is to what extent a logical system that models vague language ought to preserve the features of classical First order logic (FOL). From a linguist’s point of view, comparisons with FOL are pertinent only in so much as, as we will see in Section 2.2, it appears that natural languages do share a certain set of properties with FOL. In other words, many concerns that motivate many philosophical theories of vagueness do not directly apply to the project of providing a semantics for fragments of natural language that contain vague expressions. However, as we will see, the consideration of some of the new data discussed in this book will have implications for what kind of theories of vagueness are appropriate for modelling the full range of patterns treated in this work. Thus, when this new data reveals ways in which existing accounts make different predictions, I will make remarks accordingly.

The chapter is organized as follows: I mentioned above that vague predicates appear to be problematic for our Classical Semantic Theories (CSTs); therefore, before discussing vagueness, I outline what I take to be the defining features of CSTs. Then, in Section 2.3, I briefly exemplify the phenomenon of vagueness with a couple of classic examples and discuss why these examples appear puzzling for our CSTs. Of course, a more in-depth empirical study of vague adjectives is given in the rest of the book. In Section 2.4, I present the TCS account of vague language, and, finally, in Section 2.5, I present (and formalize) a similar framework that has been very influential in linguistics: Lasersohn’s (1999) Pragmatic Halos (PH) framework. I give a comparison between the two approaches and argue that, while Cobreros et al.’s (2012b) analysis (as applied to the interpretation of English) is empirically superior, they share many of the

---

2 A concrete example: As we will see in Section 2.3, many speakers (the author included) judge sentences of the form \(A \text{ is both } P \text{ and not } P\) as non-contradictory when \(A\) is a borderline case of \(P\). However, given that such FOL translations of such sentences are contradictions, many philosophical theories maintain the contradictory nature of such statements. For example, Keefe (2000) says (p. 197) (and see also similar sentiments in Fine (1975) and van Deemter (1995)):

Many philosophers would soon discount the paraconsistent option (almost) regardless of how well it treats vagueness on the grounds of . . . the absurdity of \(p \land \lnot p\) both being true for many instances of \(p\).

Thus, we can already see that theories of vagueness, which by design, have no way of dealing with overt contradictions (either by allowing them, as in paraconsistent logics, or explaining them away in a non-paraconsistent approach), are already inadequate semantic/pragmatic theories for languages like English.
same driving intuitions. Thus, one way of looking at TCS is as a more nuanced version of the halos approach.

2.2 Our Classical Semantic Theory

Although the languages and the models that we will deal with in the rest of the book will be more complicated than those of classical FOL, it is useful to take a moment to review this system, while highlighting the aspects that will be challenged by the existence of vague constituents.

2.2.1 Classical First order logic

The language of FOL is defined as follows:

Definition 2.2.1 Vocabulary. The vocabulary of FOL consists of a series of individual constants $a_1, a_2, \ldots$, individual variables $x_1, x_2, \ldots$, unary predicate symbols $P, Q, R, \ldots$, quantifiers $\forall$ and $\exists$, and connectives $\land, \lor, \neg$ and $\rightarrow$, plus parentheses.

Definition 2.2.2 Syntax.

1. Variables and constants (and nothing else) are terms.
2. If $t$ is a term and $P$ is a predicate symbol, then $P(t)$ is a well-formed formula (wff).
3. If $\phi$ and $\psi$ are wffs, then $\neg\phi$, $\phi \land \psi$, $\phi \lor \psi$, $\phi \rightarrow \psi$, $\forall x \phi$, and $\exists x \phi$ are wffs.
4. Nothing else is a wff.

Now we define the semantics for FOL. We first define models that consist of a set of individuals $D$ and a function $m$.

Definition 2.2.3 Model. A model is a tuple $M = \langle D, m \rangle$ where $D$ is a non-empty domain of individuals and $m$ is a mapping on the non-logical vocabulary satisfying:

1. for a constant $a_i$, $m(a_i) \in D$
2. for a predicate $P$, $m(P) \subseteq D$.

The interpretation of variables is given by assignments.

In this chapter, for simplicity, I will limit the discussion to systems with unary predicate because the $n$-ary predicate case is simply a straightforward generalization of the unary predicate case.
Definition 2.2.4 Assignment. An assignment in a model $M$ is a function $g : \{x_n : n \in \mathbb{N}\} \rightarrow D$ (from the set of variables to the domain $D$).

A model together with an assignment is an interpretation.

Definition 2.2.5 Interpretation. An interpretation $\mathcal{I}$ is a pair $(M, g)$, where $M$ is a model and $g$ is an assignment.

We first associate an element from the domain $D$ with every interpretation $\mathcal{I}$ and every term $t$.

Definition 2.2.6 Interpretation of terms.

1. If $x_1$ is a variable, then $\mathcal{I}(x_1) = g(x_1)$
2. If $a_1$ is a constant, then $\mathcal{I}(a_1) = m(a_1)$

Finally, the satisfaction relation ($\models$) is defined as in Definition 2.2.7.\(^4\) In what follows, for an interpretation $\mathcal{I} = (M, g)$, a variable $x_1$, and $a_1$ a constant, let $g[a_1/x_1]$ be the assignment in $M$ that maps $x_1$ to $a_1$ and agrees with $g$ on all variables that are distinct from $x_1$. Also, let $\mathcal{I}[a_1/x_1] = (M, g[a_1/x_1])$.

Definition 2.2.7 Satisfaction ($\models$). For all interpretations $\mathcal{I} = (M, g),$

1. $\mathcal{I} \models P(t)$ iff $\mathcal{I}(t) \in m(P)$
2. $\mathcal{I} \models \neg \phi$ iff $\mathcal{I} \not\models \phi$
3. $\mathcal{I} \models \phi \land \psi$ iff $\mathcal{I} \models \phi$ and $\mathcal{I} \models \psi$
4. $\mathcal{I} \models \phi \lor \psi$ iff $\mathcal{I} \models \phi$ or $\mathcal{I} \models \psi$
5. $\mathcal{I} \models \phi \rightarrow \psi$ iff $\mathcal{I} \models \phi$, then $\mathcal{I} \models \psi$
6. $\mathcal{I} \models \forall x_1 \phi$ iff for every $a_1$ in $D$, $\mathcal{I}[a_1/x_1] \models \phi$
7. $\mathcal{I} \models \exists x_1 \phi$ iff there is some $a_1$ in $D$, $\mathcal{I}[a_1/x_1] \models \phi$

In the next sections, we will discuss a number of theorems and arguments of FOL. Thus, we define the consequence relation between sets of formulas as follows:

Definition 2.2.8 Consequence ($\models$). A set of formulas $\Psi$ is a consequence of a set of formulas $\Phi$ (written $\Phi \models \Psi$) iff every interpretation that is a model of $\Phi$ is also a model of $\Psi$.

1. Instead of $\{\psi\} \models \{\phi\}$, we will write $\psi \models \phi$.
2. A formula $\phi$ is valid (written $\models \phi$) iff $\emptyset \models \phi$.

\(^4\) As is common, I will use $a_2$ to refer to both the expression in the language and its interpretation ($\mathcal{I}(a_2)$), provided that it is clear from context which is meant.
2.2.1.1 Aspects of First order logic to note  The first aspect of FOL that will become important in the discussion of vague language is that the two-element Boolean algebra of truth values \{0,1\} (aka \{true,false\}) underlies Definition 2.2.7 above. Well-formed formulas of FOL are mapped to exactly one of these these values in the way described by the definition of satisfaction. There are only two truth values. Additionally, each interpretation of FOL is a (total) function: it is both total and single-valued from the language into \{0,1\}.

Furthermore, Definition 2.2.7 is recursive and truth-functional: which of the two truth values a wff is assigned is determined by the values assigned to its syntactic components. The components that are predicates are assigned a set of individuals. These sets have sharp boundaries. For a given predicate denotation, an individual’s degree of membership is either 0 or 1: in the set or out of the set. In this way, a unary predicate \(P\) naturally partitions the domain into the set of individuals included in \(P\) and its complement.

A final feature of FOL that is relevant for the puzzle of vagueness is the interpretation of negation. As shown in Definition 2.2.7, a formula of the form \(\neg P(a_1)\) is true just in case the corresponding formula \(P(a_1)\) is false. In other words, \(\neg P(a_1)\) is true just in case \(a_1\) is in the complement of \(P\) in \(D\). The partitioning nature of predicates and the definition of negation gives rise to certain validities in FOL (and related systems). For example, given Definition 2.2.7, it is impossible for an individual to be a member of both a predicate and its negation. This is known as the principle of bivalence (1).

(1) Bivalence:
For all \(\mathcal{I}\) and predicates \(P\),
\[
\mathcal{I}(\exists x_1 P(x_1) \land \neg P(x_1)) = 0
\]

In other words, there are no interpretations of FOL that can satisfy \(\exists(x_1 P(x_1) \land \neg P(x_1))\). This fact has an important effect on the semantic consequences that we can draw from such sentences. In particular, since no interpretations satisfy \(\exists(x_1 P(x_1) \land \neg P(x_1))\), by Definition 2.2.8, any formula is a consequence of this sentence. In general, (2) follows immediately from Definitions 2.2.4–2.2.7.

(2) Contradiction with explosion:
For all formulas \(\phi, \psi\),
\[
\{\phi, \neg \phi\} \models \psi
\]
Furthermore, by virtue of the definition of negation, every individual must be in either the extension of a predicate $P$ or its anti-extension ($D - m(P)$). This is the law of excluded middle (3).

(3) **Excluded middle:**
For all predicates $P$,
\[
\forall y \ (P(y) \lor \neg P(y))
\]
In other words, all interpretations satisfy $P(a_i) \lor \neg P(a_i)$, for all $a_i \in D$.

Finally, I take a moment to highlight some other facts that hold in FOL given the semantics that we outlined above. These will become relevant in the discussion of the Sorites paradox below and the adopted non-classical approach to solving it. Firstly, we can note that Modus Ponens is valid in FOL (4).

(4) **Modus Ponens:**
For all formulas $\phi, \psi$,
\[
\{ \phi \to \psi, \phi \} \models \psi
\]

Secondly, we note that the deduction (meta)theorem holds (5).

(5) **Deduction theorem:**
For all sets of formulas $\Gamma, \Delta$,
\[
\Gamma \models \Delta \iff \models \bigwedge \Gamma \to \bigvee \Delta
\]

Finally, we note that the consequence relation is transitive: (6) holds.

(6) **Transitivity:**
For all formulas $\phi, \psi, \chi$,
If $\phi \models \psi$ and $\psi \models \chi$, then $\phi \models \chi$

In summary, I have highlighted some basic features of classical FOL:

1. Every interpretation is total from expressions of the language into the set $\{0, 1\}$.
2. Predicates are assigned sets with (sharp) boundaries.
3. Negation partitions the domain, resulting in excluded middle, bivalence, and contradiction with explosion.
4. The consequence relation is transitive, the deduction theorem holds, and Modus Ponens is a valid rule of inference.

As we will see in Section 2.3, the semantic behavior of vague predicates will appear to be in conflict with the picture described in this section.
2.2.2 Extensions in linguistics

Clearly, the system just presented does not look very much like an interpreted grammar for English or any other possible natural language. And, we might wonder what bearing paradoxes for FOL might have on our theories of how meaning is constructed in human languages. However, within the Montagovian approach to the study of natural language semantics and pragmatics, the types of semantics that we give to grammars analyzing fragments of natural languages have much in common with the semantics of FOL already described, despite the many kinds of enrichments that linguistics have proposed. For example, many advances in linguistic semantics have been made by proposing that the domain of individuals $D$ is in fact sorted: it contains more than one kind of object. Some analyses of scalar adjectives propose that, instead of denoting properties of individuals, they denote binary relations between individuals and other kinds of objects in the domain: degrees on a scale. This is known as the degree theory of scalar predicates (Cresswell (1976), Bierwisch (1989) and very much subsequent work in the field).

(7) Degree analysis of $\text{tall}$:
   $\llbracket \text{tall} \rrbracket = \{ \langle x, d \rangle : x \text{ is tall to degree } d \}$

Linguists and philosophers have made similar proposals to enrich the ontology of possible referents to include, besides degrees, events, worlds, times, numbers, among other things. The other type of domain enrichment common in linguistics is to impose additional relations between individuals that are not present in classical models for FOL. These extensions are common in algebraic semantics (c.f. Link, 1983; Keenan and Faltz, 1985; Krifka, 1989, and much later work) and the analyses in Chapter 6 of this work will take advantage of some of them. However, we can observe that the extensions proposed by linguists within the Montagovian tradition all preserve the properties that I highlighted above as being challenged by the phenomenon of vagueness.

1. Every interpretation function is still total, with $\{0, 1\}$ being the only truth values.
2. Constituents are still assigned sets. These sets may have more structure or consist of different sorts of objects than in many interpretations of FOL; however, the set-theoretic boundaries of these relations are still sharp.
3. In the vast majority of linguistic theories, negation is treated either as a propositional truth-reversing operator (e.g. as in Chierchia and McConnell-Ginet, 2000; Heim and Kratzer, 1998, etc.) or as a more general complement operator (e.g. Keenan and Faltz, 1985; Winter, 2001). Thus, versions of excluded middle and bivalence are taken to hold in natural languages as well.

4. Entailment in natural language is generally taken to have the same properties as in FOL (Heim and Kratzer (1998), Chierchia and McConnell-Ginet (2000), and every other textbook in formal semantics). Namely, semantic consequence is taken to be transitive, i.e. we want inferences like (8) to hold between natural language sentences, and some sort of deduction theorem should also hold (9).

(8) If John came to the party early $\models_{\text{Eng}}$ John came to the party and, John came to the party $\models_{\text{Eng}}$ John was at the party at some time, then John came to the party early $\models_{\text{Eng}}$ John was at the party at some time.

(9) John came to the party early $\models_{\text{Eng}}$ John came to the party, iff $\models_{\text{Eng}}$ John came to the party early $\rightarrow$ John came to the party

In the rest of this chapter, I will discuss how vague predicates are problematic for an analysis within FOL, since this is how the puzzles of vagueness are standardly presented in the philosophical literature. It should be clear, however, that these problems apply not only to simple classical FOL systems, but to the vast majority of semantic theories for natural language expressions that are proposed in philosophy and linguistics.

2.3 The phenomenon of vagueness

In everyday language, the term vague has many uses. Not all of these uses refer to the particular linguistic phenomenon that will be studied in this book. For example, if you ask me,

(10) “Where do you study?”

And I answer,

(11) “In the United States and in France.”
Then, in a normal situation, you would probably accuse me of being vague because I have not included very much information in my answer. However, this kind of lack of specificity is not what is meant by the technical term *vague* in linguistics and philosophy. In the rest of this section, I present the three main characterizations of vague language in the sense relevant to this work and discuss how the properties of vague language appear to be problematic for our CSTs in logic and linguistics. These properties are the *borderline cases* property, the *fuzzy boundaries* property, and the *susceptibility to the Sorites paradox* property.

In what follows, the exemplification and discussion of vague language will be limited to “uncontroversial” cases of vagueness: so-called relative scalar adjectives like *tall* and *expensive*. In Chapter 4, I will argue that another class of adjectives (absolute adjectives like *empty* and *straight*) should be analyzed as vague; however, for the purpose of illustrating the phenomenon, in this chapter, I will stick to the classical examples.

### 2.3.1 Borderline cases

The first characterization of vague predicates found in the literature, going back to Peirce (1901), if not earlier, is the borderline cases property. That is, vague predicates are those that admit borderline cases: objects to which it is unclear whether or not the predicate applies. Consider the following example with the predicate *tall*: If we take the set of American males as the appropriate comparison class for *tallness*, we can easily identify the ones that are clearly tall: for example, anyone over 6 ft. Similarly, it is clear that anyone under 6 ft 9 in (the average) is not tall. But suppose that we look at John who is somewhere between 5 ft 9 in and 6 ft. Which one of the sentences in (12) is true?

\[(12) \quad \begin{align*}
    a. \text{John is } & \text{tall}. \\
    b. \text{John is not } & \text{tall}.
\end{align*}\]

For John, a borderline case of *tall*, it seems like the most appropriate answer is either “neither” or “both.” In fact, many recent experimental studies on contradictions with borderline cases have found that the “both” and/or “neither” answers seem to be favored by natural language speakers. For example, Alxatib and Pelletier (2011) find that many participants are inclined to permit what seem like overt contradictions of the form in (13) with borderline cases. Additionally, Ripley (2011) finds similar judgments for the predicate *near*. 

---

*OUP CORRECTED PROOF – FINAL, 7/11/2016, SPI*
(13)  a. Mary is neither tall nor not tall.
    b. Mary is both tall and not tall.

At first glance, we might hypothesize that what makes us doubt the principle of bivalence with borderline cases is that the context does not give us enough information to make an appropriate decision; for example, we are ignorant about John’s height. However, as observed by Peirce, adding the required information does not make any difference to resolving the question: finding out that John is precisely 5 ft 11 in does not seem to help us decide which sentence in (12) is true and which is false, or eliminate our desire to assent to contradictions for classical logical systems like (13).

Clearly, the existence of borderline cases poses a challenge for our classical semantic theories in both logic and linguistics. As mentioned in Section 2.2, these systems are all bivalent: there can be no individuals who are both members of a predicate and its negation. Furthermore, these systems all obey the law of excluded middle: there can be no individuals who are members neither of a predicate nor its negation. Thus, we have a puzzle.

The existence of borderline cases has been taken to be the defining property of vague language by a number of authors following Peirce (1901), including those advocating classical evaluationist (super/subevaluationist) frameworks like Fine (1975). However, many authors since Kamp (1975) have argued that the borderline cases property is too broad to properly characterize the constructions that we are interested in. As an illustration of the problem, consider the following predicate described by (Smith, 2008, p. 133):  

(14)  a. If $x$ is less than four feet in height, then ‘$x$ is short’ is true.
    b. If $x$ is more than six feet in height, then ‘$x$ is short’ is false.

(The end)

The predicate in (14) has borderline cases: all those individuals whose heights are between 4 ft and 6 ft. If John is 5 ft tall, then he is included in neither short’s extension nor its anti-extension. However, we can remark that, despite failing excluded middle, short is perfectly precise: there is a sharp division between its positive cases and its borderline cases.

---

5 Similar predicates are discussed in Fine (1975), Soames (1999), and Tappenden (1993). Note, in fact, that Fine, contrary to Kamp and Smith among others judges underspecified predicates like short to be paradigm cases of vagueness.
cases on the one hand, and another sharp division between its borderline cases and its negative cases on the other. Thus, although it applies to vague predicates, the borderline cases property does not seem to be what is at the heart of the phenomenon of vagueness.

We might note as well that many sentences in natural language appear to fail bivalence or excluded middle, like presupposition failures (15) or implicature failures (16), and we would not necessarily want to say that these expressions are vague because of it.

(15) The present king of France is bald.
(16) Dogs are in my yard right now.
   (In a context where there is a single dog in my yard)

Thus, in Section 2.3.2, I present a second characterization found in the literature that narrows the empirical domain of the study of vagueness: the fuzzy boundaries and the closely related tolerance property.

2.3.2 Fuzzy boundaries

A second characterization of vague predicates going back to Frege’s (1904) Grundgesetze is the fuzzy boundaries property. This is the observation that there are (or appear to be) no sharp boundaries between cases of a vague predicate $P$ and its negation. To take a concrete example: If we take a tall person and we start subtracting millimetres from their height, it seems impossible to pinpoint the precise instance where subtracting a millimetre suddenly moves us from the height of a tall person to the height of a not tall person. The same thing holds for expensive: if we take the price of an object that is clearly expensive (for that type of object) and we keep subtracting 1¢ from its cost, at some point, we will arrive at a price that is not expensive, but precisely specifying this point does not seem possible.

The fuzzy boundaries property is problematic for our classical semantic theories because we assign set-theoretic structures to predicates and their negations, and these sets have sharp boundaries. In principle, if we line up all the individuals in the domain according to height, we ought to be able to find an adjacent pair in the tall-series consisting of a tall person and a not tall person. However, it does not appear that this is possible.

Of course, one way to get around this problem would be to just stipulate where the boundary is, say, at another contextually given value for tall; however, if we were to do this, we would be left with the
impression that the point at which we decided which of the borderline cases to include and which to exclude was arbitrary. The inability to draw sharp, non-arbitrary boundaries is often taken to be the essence of vagueness (e.g. by Fara, 2000), and it is intimately related to another characterization of vague language: vague predicates are those that are tolerant. We will call a predicate tolerant with respect to a scale or a dimension $\Theta$ if there is some degree of change in respect of $\Theta$ that is insufficient ever to affect the justice with which the predicate is applied to a particular case. This novel definition of vagueness was first proposed by Wright (1975) as a way to give a more general explanation to the fuzzy boundaries feature; however, versions of this idea have, more recently, been further developed and taken to be at the core of what it means to be a vague expression (e.g. Eklund, 2005; Smith, 2008; van Rooij, 2010; Cobreros et al., 2012b). This property is more nuanced than the fuzzy boundaries property in that it makes reference to a dimension and to an incremental structure associated with this dimension, and it puts an additional constraint on what can be defined as a vague predicate: the distance between the points on the associated dimension must be sufficiently small such that changing from one point to an adjacent one does not affect whether we would apply the predicate. Immediately, we can see that tall is tolerant. There is an increment, say 1 mm, such that if someone is tall, then subtracting 1 mm does not suddenly make them not tall. Similarly, adding 1 mm to a person who is not tall will never make them tall. Since height is continuous, we will always be able to find some increment that will make tall tolerant. So, if we are considering very small things for whom 1 mm makes a significant difference in size, we can just pick 0.5 mm or whatever.

In summary, the second characterization of vague predicates in the literature is the fuzzy boundaries characterization, or its more specific tolerance characterization: vague predicates are those whose application is insensitive to extremely small changes, and thus, they appear to lack sharp boundaries.

### 2.3.3 The Sorites paradox

One of the reasons that vagueness has received so much attention in philosophy is that vague predicates seem to give rise to arguments that result in contradiction in FOL. The first discussion of the Sorites paradox (literally, the paradox of the “heaper”) is generally attributed to
the Megarian philosopher Eubulides of Miletus and, informally, it can be laid out as below:6

Would you describe a single grain of wheat as a heap? No. Would you describe two grains of wheat as a heap? No. . . . You must admit the presence of a heap sooner or later, so where do you draw the line?

Formally, the paradox can set up in a number of ways in FOL. A common one found in the literature is (17), where \( \sim P \) is a ‘little by little’ or ‘indifference’ relation.7

\[
(17) \quad \textbf{The Sorites paradox} \\
\quad \text{a. Clear case: } P(a_i) \\
\quad \text{b. Clear non-case: } \neg P(a_k) \\
\quad \text{c. Sorites series: } \forall i \in [1, n] (a_i \sim P a_{i+1}) \\
\quad \text{d. Tolerance: } \forall x \forall y ((P(x) \land x \sim P y) \rightarrow P(y)) \\
\quad \text{e. Conclusion: } P(a_k) \land \neg P(a_k)
\]

Thus, in FOL and other classical systems, as soon as we have a clear case of \( P \), a clear non-case of \( P \), and a Sorites series, through universal instantiation and repeated applications of Modus Ponens8 we can conclude that everything is \( P \) and that everything is not \( P \). We can see that tall (for a North American male) gives rise to such an argument. We can find someone who measures 6 ft to satisfy (17-a), and we can find someone who measures 5 ft 6 in to satisfy (17-b). In the previous subsection, we concluded that tall is tolerant, so it satisfies (17-d), and, finally, we can easily construct a Sorites series based on height to fulfill (17-c). Therefore, we would expect to be able to conclude that this 5 ft 6 in tall person (a non-borderline case) is both tall and not tall. I stress again that the Sorites is not only a paradox for FOL. As discussed in Section 2.1, the semantic theories that linguists employ all validate bivalence, excluded middle, and Modus Ponens. Thus, the puzzles that vague predicates raise are widespread in (at least) the nominal and adjectival domains, and shake the very core of the logical approach to natural language semantics.

6 From the Stanford Encyclopedia of Philosophy.

7 Note that, technically speaking, the Sorites argument is not stateable in the system that I set out above because the language does not contain binary predicates like \( \sim P \). Thus, the Sorites must be formulated in a slightly enriched language.

8 Note that universal instantiation is not even necessary for the paradox: we can replace the quantified statements in (17) by individual conditionals and the result is the same; it is the validity of Modus Ponens that is important for the Sorites.
Finally, we can oppose predicates like *tall*, *expensive*, and *heap* that display these three properties with precise or non-vague predicates that do not display every member of this characterizing cluster. Some examples of precise adjectival predicates are shown in (18).

(18)  

a. Mary is *Canadian* (in the ‘citizenship’ sense).
b. This algebra is *atomic*.
c. This number is *prime*.

We will see further examples of precise predicates and constituents throughout the course of this book, and one of the main goals of this work is to investigate the grammatical factors that determine whether or not a predicate can be vague.

2.4 Tolerant, Classical, Strict

In this section, I outline the logical framework for the analysis of the puzzling properties of vague language that will form the backbone of the analyses of vagueness and scale structure in this work: Cobreros et al.’s (2012b) TCS framework. In what follows, I provide a succinct definition of the system for expository purposes; however, a version of this system extended with the structure of Klein’s delineation system presented in Chapter 1 will be given in Chapter 4.

2.4.1 Definition

This system was originally developed as a way to preserve the intuition that vague predicates are *tolerant* (i.e. satisfy $\forall x \forall y (P(x) \& x \sim_p y \rightarrow P(y))$, where $\sim_p$ is an indifference relation for a predicate $P$), without running into the Sorites paradox. Cobreros et al. (2012b) adopt a non-classical logical framework with three notions of satisfaction: classical truth, tolerant truth, and its dual, strict truth. Formulas are tolerantly/strictly satisfied based on classical truth and predicate-relative, possibly non-transitive *indifference relations*. For a given predicate $P$, an indifference relation, $\sim_p$, relates those individuals that are viewed as sufficiently similar with respect to $P$. For example, for the predicate *tall*, $\sim_{tall}$ would be something like the relation “not looking to have distinct heights.” In this framework, we say that *John is tall* is tolerantly true just in case John has a very similar height to someone who is classically tall (i.e. has a height greater than or equal to the contextually given ‘tallness’ threshold). The framework is defined as follows:
Definition 2.4.1 Language. The language of TCS is that of first order predicate logic with neither identity nor function symbols. Additionally, for every predicate $P$, there is a binary predicate: $I_P$.

- If $t_1, t_2$ are terms, then $t_1 I_P t_2$ is a wff.

For the semantics, we define three notions of satisfaction: one that corresponds to satisfaction in classical FOL ($c$-satisfaction), and two that are novel ($t$-satisfaction and its dual $s$-satisfaction).

Definition 2.4.2 C-model. A model is a tuple $M = \langle D, m \rangle$ where $D$ is a non-empty domain of individuals and $m$ is a mapping on the non-logical vocabulary satisfying:

1. For a constant $a_1$, $m(a_1) \in D$.
2. For a predicate $P$, $m(P) \subseteq D$.

Definition 2.4.3 T-model. A $T$-model is a tuple $\langle D, m, \sim \rangle$, where $\langle D, m \rangle$ is a C-model and $\sim$ is a function that takes any predicate $P$ to a binary relation $\sim_P$ on $D$. For any $P$, $\sim_P$ is reflexive and symmetric (but possibly not transitive).

A non-empty set with a reflexive, symmetric relation on it is often called a tolerance space (e.g. Pogonowski, 1981). Thus, for any $P$, the structure $\langle D, \sim_P \rangle$ is a tolerance space.

The interpretation of variables is given by assignments.

Definition 2.4.4 Assignment. An assignment in a model $M$ is a function $g : \{x_n : n \in \mathbb{N}\} \rightarrow D$ (from the set of variables to the domain $D$).

A model together with an assignment is an interpretation.

Definition 2.4.5 Interpretation. An interpretation $\mathcal{I}$ is a pair $\langle M, g \rangle$, where $M$ is a model and $g$ is an assignment.

We associate an element from the domain $D$ with every interpretation $\mathcal{I}$ and every term $t$.

Definition 2.4.6 Interpretation of terms.

1. If $x_i$ is a variable, then $\mathcal{I}(x_i) = g(x_i)$.
2. If $a_i$ is a constant, then $\mathcal{I}(a_i) = m(a_i)$. 

The classical satisfaction relation in TCS is simply the satisfaction relation of FOL (extended to Ios as shown in Definition 2.4.7). In what follows, for an interpretation $\mathcal{I} = \langle M, g \rangle$, a variable $x_i$, and $a_i$ a constant, let $g[a_i/x_i]$ be the assignment in $M$ which maps $x_i$ to $a_i$ and agrees with $g$ on all variables that are distinct from $x_i$. Also, let $\mathcal{I}[a_i/x_i] = \langle M, g[a_i/x_i] \rangle$.

**Definition 2.4.7 Classical satisfaction ($\models^c$).** Let $M$ be a $T$-model such that $M = \langle D, m, \sim, \cdot \rangle$, and let $\mathcal{I}$ be an interpretation. For all predicates $P$ and terms $t_1, t_2$:

1. $\mathcal{I} \models^c P(t_1) \text{ iff } (t_1) \in m(P)$
2. $\mathcal{I} \models^c t_1 t P t_2 \text{ iff } \mathcal{I}(t_1) \sim_P \mathcal{I}(t_2)$
3. $\mathcal{I} \models^c \neg \phi \text{ iff } \mathcal{I} \not\models^c \phi$
4. $\mathcal{I} \models^c \phi \land \psi \text{ iff } \mathcal{I} \models^c \phi \text{ and } \mathcal{I} \models^c \psi$
5. $\mathcal{I} \models^c \phi \lor \psi \text{ iff } \mathcal{I} \models^c \phi \text{ or } \mathcal{I} \models^c \psi$
6. $\mathcal{I} \models^c \phi \rightarrow \psi \text{ iff } \mathcal{I} \models^c \phi, \text{ then } \mathcal{I} \models^c \psi$
7. $\mathcal{I} \models^c \forall x \phi \text{ iff for every } a_i \text{ in } D, \mathcal{I}[a_i/x_i] \models^c \phi$
8. $\mathcal{I} \models^c \exists x \phi \text{ iff there is some } a_i \text{ in } D, \mathcal{I}[a_i/x_i] \models^c \phi$

**Definition 2.4.8 Tolerant/strict satisfaction ($\models^t/s$).** Let $\mathcal{I}$ be an interpretation. For all predicates $P$ and terms $t_1, t_2$:

1. $\mathcal{I} \models^t P(a_1) \text{ iff } \exists x_1 \sim_P a_1 : \mathcal{I} \models^c P(a_1)$
2. $\mathcal{I} \models^t t_1 t P t_2 \text{ iff } \mathcal{I}(t_1) \sim_P \mathcal{I}(t_2)$
3. $\mathcal{I} \models^t \neg \phi \text{ iff } \mathcal{I} \not\models^c \phi$
4. $\mathcal{I} \models^t \phi \land \psi \text{ iff } \mathcal{I} \models^t \phi \text{ and } \mathcal{I} \models^t \psi$
5. $\mathcal{I} \models^t \phi \lor \psi \text{ iff } \mathcal{I} \models^t \phi \text{ or } \mathcal{I} \models^t \psi$
6. $\mathcal{I} \models^t \phi \rightarrow \psi \text{ iff } \mathcal{I} \not\models^t \phi \text{ or } \mathcal{I} \models^t \psi$
7. $\mathcal{I} \models^t \forall x \phi \text{ iff for every } a_i \text{ in } D, \mathcal{I}[a_i/x_i] \models^t \phi$
8. $\mathcal{I} \models^t \exists x \phi \text{ iff there is some } a_i \text{ in } D, \mathcal{I}[a_i/x_i] \models^t \phi$
9. $\mathcal{I} \models^t P(a_1) \text{ iff } \forall x_1 \sim_P a_1 : \mathcal{I} \models^c P(a_1)$
10. $\mathcal{I} \models^t t_1 t P t_2 \text{ iff } \mathcal{I}(t_1) \sim_P \mathcal{I}(t_2)$
11. $\mathcal{I} \models^t \neg \phi \text{ iff } \mathcal{I} \not\models^t \phi$
12. $\mathcal{I} \models^t \phi \land \psi \text{ iff } \mathcal{I} \models^t \phi \text{ and } \mathcal{I} \models^t \psi$
13. $\mathcal{I} \models^t \phi \lor \psi \text{ iff } \mathcal{I} \models^t \phi \text{ or } \mathcal{I} \models^t \psi$
14. $\mathcal{I} \models^t \phi \rightarrow \psi \text{ iff } \mathcal{I} \not\models^t \phi \text{ or } \mathcal{I} \models^t \psi$
15. $\mathcal{I} \models^t \forall x \phi \text{ iff for every } a_i \text{ in } D, \mathcal{I}[a_i/x_i] \models^t \phi$
16. $\mathcal{I} \models^t \exists x \phi \text{ iff there is some } a_i \text{ in } D, \mathcal{I}[a_i/x_i] \models^t \phi$
Note that the predicates that refer to indifference relations are interpreted “crisply” (in the words of Cobreros et al., 2012b): their interpretation is the same on all kinds of satisfaction.

2.4.1.1 Consequence relations The framework has three notions of satisfaction, and from these notions we can derive nine consequence relations (defined in a similar manner to the consequence relation of FOL in Definition 2.2.8). As discussed in Cobreros et al. (2012b), these relations are in the following lattice order (based on inclusion, shown in Figure 2.1), where $\models^{mn}$ stands for reasoning from $m$ interpreted premises to $n$ interpreted conclusions. Note (as shown in Cobreros et al., 2012b) that $\models^{cc}$ is equivalent to consequence in classical FOL (i.e. reasoning from classical premises to classical premises). Furthermore, $\models^{tt}$ is equivalent to consequence in Priest’s 1979 *Logic of Paradox* (LP), and $\models^{ss}$ is equivalent to strong Kleene logic (K3) (Kleene, 1950).

How appropriate are these systems as basic semantic theories for natural languages? We saw in Sections 2.2 and 2.3 that, to model a language like English, we want a system that validates certain arguments and invalidates others. First of all, we wanted the principle of tolerance (19) to be valid, since it seems that vague predicates permit this kind of reasoning.

(19) **Tolerance:**
\[
\forall x \forall y (P(x) \land x \sim_P y \rightarrow P(y))
\]
Secondly, we do not want the Sorites argument (17) to be valid because it is paradoxical. Thirdly, I suggested in Section 2.2 that Modus Ponens (4) and the deduction theorem (5) ought to be valid for English. Furthermore, since, as discussed in Section 2.3, contradictions do seem to be possible with borderline cases, we want Explosion (2) to be invalid. Finally, whether (6) should be valid is somewhat unclear. Sometimes, as discussed in Section 2.2, it seems like transitive inferences go through, but, on the other hand, transitivity is part of what gets us into trouble with the Sorites paradox. Thus, we want a system that (in)validates the following arguments:

1. Tolerance = valid (√)
2. Sorites = invalid (×)
3. Modus Ponens = valid (√)
4. Deduction theorem = valid (√)
5. Explosion = invalid (×)
6. Transitivity = ?

I first consider non-mixed consequence, as shown in Table 2.1.

From Table 2.1, we can observe that the only non-mixed consequence relation that validates tolerance without also validating the Sorites is ⊨_{tt} (LP). However, systems with ⊨_{tt} have neither Modus Ponens nor the deduction theorem; therefore, they are not so useful for modelling natural language. We already argued that the consequence relation in classical FOL ⊨_{cc} was inadequate, and we can also note that ⊨_{ss} (K3) does not validate the tolerance principle and validates explosion, which is undesirable.

<table>
<thead>
<tr>
<th>Argument</th>
<th>⊨_{cc} (FOL)</th>
<th>⊨_{tt} (LP)</th>
<th>⊨_{ss} (K3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tolerance</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Sorites</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Modus Ponens</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Deduction theorem</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Explosion</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Transitivity</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

\* Note however that excluded middle is invalid in K3.
Therefore, I now consider mixed consequence relations. Since we are interested in consequence relations that validate the tolerance principle, and tolerance is never classically nor strictly valid, I will only consider $\models^c$ and $\models^{ct}$.\textsuperscript{10}

As shown in Table 2.2, $\models^c$ and $\models^{ct}$ both validate tolerance but avoid the Sorites paradox (we will return to this point in Section 2.5). The only difference between them is that only $\models^{ct}$ validates the deduction theorem. I therefore conclude (with Cobreros et al., 2012b) that $\models^{ct}$ is the system that is the most appropriate for modelling reasoning associated with vague predicates in natural language. With this system in mind, I turn to how TCS explains the cluster of properties that characterize vague predicates.

2.4.2 Account of the puzzling properties

TCS (with $\models^{st}$) explains the puzzling properties of vague language in the following way. Firstly, although classical negation partitions the domain (like it does in FOL), the definition of tolerant negation actually allows for $P(a_1)$ and $\neg P(a_1)$ to be tolerantly true for some individual $a_1$. Individuals like $a_1$ are the borderline cases. The reason that we have difficulty deciding whether a borderline individual is part of a predicate’s extension or anti-extension is that such an individual is actually part of both sets. In other words, at the level of tolerant truth, TCS is paraconsistent: contradictions involving borderline cases do not result in explosion (like they do in classical logic: see Table 2.2). Secondly, TCS preserves the intuition behind the fuzzy boundaries/tolerance property because the principle of tolerance is, in fact, valid at the level of tolerant truth. Note that it is neither classically valid nor strictly valid.

\textsuperscript{10} Because single premise validity for $\models^{mm}$ is entirely dependent on $n$. 

\textbf{Table 2.2} Mixed consequence relations validating tolerance

<table>
<thead>
<tr>
<th>Argument</th>
<th>$\models^{ct}$</th>
<th>$\models^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tolerance</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sorites</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Modus Ponens</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Deduction theorem</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Explosion</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Transitivity</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>
How this system avoids the Sorites paradox is a bit more complicated. Firstly, following Cobreros et al. (2012b, p. 27), we can distinguish two syntactic versions of the argument. The first version proceeds directly from indifference relations:

(20) **Sorites version 1:**

a. \( P(a_1) \)

b. \( \forall i \in [1, n](a_i I a_{i+1}) \)

c. \( P(a_k) \)

This version of the Sorites is \( st \)-invalid. However, what is interesting is that TCS (with \( \models st \)) validates each step along the way, which seems appropriate.

(21) **Step-wise tolerance**

a. \( P(a_1) \)

b. \( a_1 I a_2 \)

c. \( P(a_2) \)

The reason that (20) is invalid, despite the validity of (21) for all individuals adjacent on the scale, is that \( \models st \) is not transitive (cf. Table 2.2).

There is, however, a second version of the Sorites that is more similar to the formulation presented in (17) and is \( st \)-valid:

(22) **Sorites version 2:**

a. \( P(a_1) \)

b. \( \forall i \in [1, n](a_i I a_{i+1}) \)

c. \( \forall x \forall y((P(x) \land x I y) \rightarrow P(y)) \)

d. \( P(a_k) \)

However, we still avoid paradox. Although (22) is valid, it is not sound. Recall that, with \( \models st \), we are reasoning from strict premises to tolerant conclusions. As I mentioned, the principle of tolerance is neither \( c \)-valid nor \( s \)-valid; thus, (22-c) will never be strictly true.

In summary, TCS is a paraconsistent indifference relation-based logical framework that preserves the intuition that vague predicates are tolerant, but avoids the Sorites paradox. It will form the basis of the analyses presented in this monograph. In Section 2.5, I will briefly outline a very similar framework that has received a lot of attention in the field of linguistics: Lasersohn’s (1999) PH framework.
2.5 Lasersohn’s (1999) Pragmatic Halos

Although it is sometimes presented as such (cf. Sauerland and Stateva, 2007), Lasersohn’s PH framework was not designed as a theory of vagueness, at least not in that it addresses the challenges to classical semantics posed by the properties that we discussed earlier in this chapter. The empirical domain of Lasersohn’s proposal is the phenomenon that he calls \textit{pragmatic slack}. Two examples of slack are the sentences in (23) and (24) used when, for example, the theater has a couple of seats filled or a few irrelevant townspeople are awake.

(23) The theatre is empty.
(24) The townspeople are asleep.

Although sentences like (23) often allow for exceptions, as Lasersohn observes, this slack is context dependent. For example, while (24) is acceptable in the general context of describing a town at night, for a similar sentence such as (25) to be said in the context of a sleep study, it is absolutely crucial that all of the members of the group denoted by the subject DP be asleep.

(25) The subjects are asleep.

In Chapters 3 and 4, I will argue that many of the examples discussed in Lasersohn’s paper should, in fact, be treated as instances of vagueness, and, indeed, it is interesting to note how similar \textit{Pragmatic Halos} (PH) is in spirit to TCS, despite not being devised as a framework for modelling vague language. The framework is laid out (in my notation) below.

2.5.1 Definition

In order to make the comparison with TCS optimally perspicuous, I will present ‘halo’ semantics for the language of FOL (defined in Section 2.1). Of course, Lasersohn (1999) provides analyses of the semantics and pragmatics of many expressions of English that have no counterparts in FOL (definite plurals, slack regulators/hedges like \textit{exactly} etc.), but this simple language is sufficient to understand how PH works.

Like TCS, PH starts with a classical model for interpreting the language, and extends it with more structure.

\footnote{Lasersohn does not give comprehensive definitions of his system; however, its architecture is easy to reconstruct given his remarks on pages 48–50.}
Definition 2.5.1 Model. A model is a tuple $M = \langle D, m \rangle$ where $D$ is a non-empty domain of individuals and $m$ is a mapping on the non-logical vocabulary\(^{12}\) satisfying:

1. For a constant $a_1$, $m(a_1) \in D$.
2. For a predicate $P$, $m(P) \subseteq D$.

Models are extended by a function that applies to elements of the language and returns relational structures associated with them. Lasersohn (1999, p. 526) proposes that, in addition to its regular denotation, for each expression in the language, the pragmatic context associates this denotation with a set of objects of the same logical type as the denotation itself. Each object in this set is understood to differ from the denotation only in some respect that is pragmatically ignorable from the context.

This set of objects is an expression's Pragmatic Halo, and Lasersohn proposes that it is partially ordered. Thus, models are extended to halo models as follows:

Definition 2.5.2 Halo model. A halo model is a tuple $M = \langle D, m, h \rangle$, where $\langle D, m \rangle$ is a model (as defined in Definition 2.2.3) and $h$ is a function from the non-logical vocabulary satisfying:

1. For an individual constant $a_1$, $h(a_1) = \langle X, \leq_{h(a_1)} \rangle$, where $X \subseteq D$ and $\leq_{h(a_1)}$ is a reflexive, transitive, and anti-symmetric relation.
2. For a predicate $P_1$, $h(P_1) = \langle X, \leq_{h(P_1)} \rangle$, where $X \subseteq P(D)$ and $\leq_{h(P_1)}$ is a reflexive, transitive, and anti-symmetric relation.


1. For an individual constant $a_1$, $m(a_1) \in \text{dom}(h(a_1))$ and there is no $a_2 \in \text{dom}(h(a_1))$ such that $m(a_1) \neq a_2$ and $a_2 \leq_{h(a_1)} m(a_1)$.
2. For a predicate $P_1$, $m(P_1) \in \text{dom}(h(P_1))$ and there is no $P_2 \in \text{dom}(h(P_1))$ such that $m(P_1) \neq P_2$ and $P_2 \leq_{h(P_1)} m(P_1)$.

In other words, the basic denotation of a non-logical term is the center of its halo. The denotation of variables are given on assignment.

Definition 2.5.3 Assignment. An assignment in a model $M$ is a function $g : \{x_n : n \in \mathbb{N}\} \rightarrow D$ (from the set of variables to the domain $D$).

We now extend the function $h$ to define halos for variables.

\(^{12}\) Technically, Lasersohn suggests that every expression in the language is assigned a halo; however, for the purpose of exposition, I will simply define halos for the non-logical vocabulary.
Definition 2.5.4 Variable halos. For a variable \( x_1 \), \( h(x_1) = (X, \leq_{h(x_1)}) \), where \( X \subseteq D \) and \( \leq_{h(x_1)} \) is a reflexive, transitive, and anti-symmetric relation.

- Furthermore, \( g(x_1) \in \text{dom}(h(x_1)) \) and there is no \( a_2 \in \text{dom}(h(x_1)) \) such that \( g(x_1) \neq a_2 \) and \( a_2 \leq_{h(x_1)} g(x_1) \).

A halo model together with an assignment is an interpretation.

Definition 2.5.5 Interpretation. An interpretation \( \mathcal{I} \) is a pair \( \langle M, g \rangle \), where \( M \) is a halo model and \( g \) is an assignment.

Definition 2.5.6 Interpretation of terms

1. If \( x_1 \) is a variable, then \( I(x_1) = g(x_1) \).
2. If \( a_1 \) is a constant, then \( I(a_1) = m(a_1) \).

“Real” truth/satisfaction is defined in the standard way.\(^ {13} \)

Definition 2.5.7 Satisfaction (\( \models \)). Let \( M \) be a halo model such that \( M = (D, m, h) \), and let \( \mathcal{I} \) be an interpretation. For all predicates \( P_1 \) and terms \( t_1, t_2 \):

1. \( I \models P_1(t_1) \) iff \( I(t_1) \in m(P) \)
2. \( I \models \neg \phi \) iff \( I \not\models \phi \)
3. \( I \models \phi \land \psi \) iff \( I \models \phi \) and \( I \models \psi \)
4. \( I \models \phi \lor \psi \) iff \( I \models \phi \) or \( I \models \psi \)
5. \( I \models \phi \rightarrow \psi \) iff \( I \models \phi \), then \( I \models \psi \)
6. \( I \models \forall x_1 \phi \) iff for every \( a_1 \) in \( D \), \( I[a_1/x_1] \models \phi \)
7. \( I \models \exists x_1 \phi \) iff there is some \( a_1 \) in \( D \), \( I[a_1/x_1] \models \phi \)

We now define a second notion of truth based on the halos. This kind of satisfaction is called “close enough to truth” by Lasersohn.

Definition 2.5.8 Close enough to truth (\( \models^e \)). Let \( M \) be a halo model such that \( M = (D, m, h) \), and let \( \mathcal{I} \) be an interpretation. For all predicates \( P_1 \) and terms \( t_1, t_2 \):

1. \( I \models^e P_1(t_1) \) iff there is some \( a_1 \in \text{dom}(h(t_1)) \) : \( a_1 \in P_2 \), for some \( P_2 \in \text{dom}(h(P_2)) \)
2. \( I \models^e \neg \phi \) iff \( I \not\models^e \phi \)
3. \( I \models^e \phi \land \psi \) iff \( I \models^e \phi \) and \( I \models^e \psi \)

\(^ {13} \) Recall that, for an interpretation \( \mathcal{I} = (M, g) \), a variable \( x_1 \), and \( a_1 \) a constant, \( g[a_1/x_1] \) is the assignment in \( M \) which maps \( x_1 \) to \( a_1 \) and agrees with \( g \) on all variables that are distinct from \( x_1 \). Furthermore, \( I[a_1/x_1] = (M, g[a_1/x_1]) \).
4. \( I \models^e \phi \lor \psi \iff I \models^e \phi \) or \( I \models^e \psi \)
5. \( I \models^e \phi \rightarrow \psi \iff I \models^e \phi \), then \( I \models^e \psi \)
6. \( I \models^e \forall x \phi \iff \) for every \( a \) in \( D \), \( I[a/x] \models^e \phi \)
7. \( I \models^e \exists x \phi \iff \) there is some \( a \) in \( D \), \( I[a/x] \models^e \phi \)

In section 2.5.2, I discuss the relationship between the PH described above and TCS.

2.5.2 Comparison with Tolerant, Classical, Strict

Two similarities between TCS (in the version presented in Cobreros et al. (2012b) and here) and PH are immediately apparent. Firstly, both frameworks start from assigning lexical items a semantic denotation that is consistent with our classical semantic theory. Secondly, we can draw a parallel between TCS’s tolerant truth and PH’s close enough to true. Both of these kinds of satisfaction are built on the classical semantic denotations of lexical items and take into consideration relations in the model that are meant to model contributions from the context: ‘indifference’ (in the case of TCS) and ‘pragmatically ignorable’ difference/irrelevance (in the case of PH). Furthermore, in both frameworks, tolerant denotations/halos are defined through existential quantification over elements related by indifference relations/halo partial order relations. Thus, the conditions for being ‘tolerantly true’ or ‘close enough to true’ are weaker than being classically true. This is Cobreros et al.’s (2012b) Lemma 1, and we can prove the same result for the PH system: we show that if \( \phi \) is true, then it is also close enough to true.

**Theorem 2.5.9** \( I \models \phi \Rightarrow I \models^e \phi \)\(^{14}\)

However, just because some formula is close enough to true, it does not mean that it is necessarily true, as shown in Theorem 2.5.10.

**Theorem 2.5.10** \( \text{It is not the case that for all } I, \phi, I \models^e \phi \Rightarrow I \models \phi \).\(^{15}\)

In other words, both TCS and PH share a common core intuition that at least one aspect of vagueness/pragmatic slack involves loosening the conditions of application of an expression with a precise semantic

---

\(^{14}\) Proof. By induction on \( \models \). Non-trivial case: Suppose \( I \models P_1(a_1) \). By Definition 2.5.2, \( m(a_1) \in dom(h(a_1)) \) and \( m(P_1) \in dom(h(P_1)) \). So by Definition 2.5.8, \( I \models^e P_1(a_1) \).

\(^{15}\) Proof. Let \( D = \{a_1, a_2\} \). Let \( m(P_1) = \{a_1\} \) and let \( P_2 = D \). Let \( h(P_1) = \{P_1, P_2\} \), \( \{P_1 \leq h(P_1) \} \) (reflexivity for \( \leq h(P_1) \)). Suppose \( h(a_2) = \{a_2\} \). Let \( g \) be an assignment. So, by the definition of \( \models^e \), \( I \models^e P_1(a_2) \), but \( I \models P_1(a_2) \), because \( a_2 \notin m(P_1) \).
denotation to include other objects that are considered to differ in only ‘pragmatically ignorable’ ways. Indeed, as shown by the coincidences between Cobreros et al.’s (2012b) Lemma 1 and Theorems 2.5.9 and 2.5.10, these two logical systems share a common structural core and, therefore, I suggest that there is a sense in which TCS extends the PH framework with a non-classical interpretation of negation and duality between tolerant and strict denotations.

However, as defined above, TCS and PH do have some differences. A primary difference concerns the structure of the tolerant denotations/halos. In PH, halos are given in the model and the ordering on their members is also given in the model (so, by context in Lasersohn, 1999). In TCS, indifference relations are given in the model, and tolerant interpretations are constructed from these relations in conjunction with classical interpretations. Furthermore, as we will see in Chapter 4, it will be possible to derive orderings between individuals based on predicates from looking at their tolerant interpretations across contexts within the Delineation semantics extension of TCS that will be the focus of this book. In other words, I also suggest that, from the perspective of the derivation of non-classical denotations and orderings associated with expressions, there is a sense in which TCS is a refinement of PH.

2.6 Conclusion

In conclusion, I have argued that the challenges that vague predicates raise for modelization within FOL are also challenges for the kinds of theories that we adopt in the field of linguistic semantics. Thus, a comprehensive analysis of vague constituents in languages like English is of great importance to the logical approach to meaning in natural language. I presented the framework for modelling the properties of vague language that I will extend throughout the rest of the book (Cobreros et al.’s 2012b TCS) and show how it bears certain important structural relationships to another similar influential framework: Lasersohn’s (1999) Pragmatic Halos.
Context sensitivity and vagueness patterns

3.1 Introduction

This chapter presents the main empirical patterns associated with two of the principle phenomena that are treated in this work: context sensitivity and vagueness in the adjectival domain. Broadly speaking, we will call a predicate \textit{context sensitive} just in case its criteria of application can be different in different contexts. More specifically, when we consider adjectival predicates, a major source of contextual variation that we observe concerns variation across \textit{comparison classes} (CCs). CCs are contextually given sets of individuals that influence (in a way to be discussed in Section 3.2) the assignment of the semantic denotations of adjectives. In line with previous work on the topic, I argue that the different classes of adjectives mentioned in the introduction (and repeated below) vary with respect to CC-based context sensitivity.

(1) \textbf{Relative adjectives (RAs)}:
\begin{itemize}
  \item e.g. tall, short, expensive, cheap, nice, friendly, intelligent, stupid, narrow, wide
\end{itemize}

(2) \textbf{Total absolute adjectives (AA}\textsuperscript{T}s)\textbf{):}
\begin{itemize}
  \item e.g. empty, full, clean, smooth, dry, straight, flat
\end{itemize}

(3) \textbf{Partial absolute adjectives (AA}\textsuperscript{P}s)\textbf{):}
\begin{itemize}
  \item e.g. dirty, bent, wet, curved, crooked, dangerous, awake
\end{itemize}

(4) \textbf{Non-scalar adjectives (NSs)}:
\begin{itemize}
  \item e.g. atomic, geographical, polka-dotted, pregnant, illegal, dead, hexagonal
\end{itemize}

Furthermore, following previous work, I argue that to properly understand this variation, it is useful to adopt two patterns of CC-based context sensitivity: (what I will call) \textit{universal} context sensitivity and
existential context sensitivity. These patterns will be exemplified in great detail in Section 3.2; however, intuitively, predicates that are universally context sensitive will show a greater range of meaning variation than predicates that are existentially context sensitive. In this chapter, I argue that the four scale-structure subclasses presented above display the following context sensitivity patterns (shown in Table 3.1): RAs show both context sensitivity patterns, both AA$^P$s and AA$^T$s are not universally context sensitive, but are existentially context sensitive, and NSs are neither universally nor existentially context sensitive.

This chapter also motivates an important empirical connection between vagueness (i.e. the appearance of the properties described in Chapter 2) and the same “scale structure” classes of adjectives shown in (1)–(4). In particular, I show that the distribution of the puzzling properties of vague language is tied to these lexical class distinctions, and I propose, following authors such as Kennedy and McNally (2005) and Kennedy (2007), that the observed dependencies argue in favor of a closer relationship between the phenomena of vagueness and scale structure than is often assumed in the literature.

As mentioned in Chapter 2, the RAs are uncontroversial examples of vague constituents, and, indeed, that is why the discussion in Chapter 2 was limited to them. However, as we will see later in this chapter (Section 3.4), in some (or indeed most) contexts, the adjectives in (2) and (3) also seem to display the symptoms of vagueness. An open debate has emerged in the linguistic and philosophical literatures as to whether the appearance of borderline cases, fuzzy boundaries/tolerance, and Sorites susceptibility with AAs should be analyzed in a parallel manner to their appearance with RAs. The dominant view in philosophy, both historically and recently (cf. Fine, 1975; Lewis, 1979; Keefe, 2000; Fara, 2000; Smith, 2008; Égré and Klinedinst, 2011, among many others), is that the aspects of the meaning of both *tall* and *straight* that trouble our classical semantic theories should be given a unified analysis. However, this view has been challenged on empirical grounds by a number of authors (cf. Pinkal, 1995; Kennedy and McNally, 2005; Kennedy, 2007; Sauerland and Stateva, 2007; Moryzcki, 2011; and Husband, 2011, among

---

**Table 3.1** Context sensitivity patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>RA</th>
<th>AA$^T$</th>
<th>AA$^P$</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal context sensitivity</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Existential context sensitivity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>
I propose that it is possible to reconcile these two views and arrive at a more accurate description of the phenomenon of vagueness and its distribution across contexts by employing a context-relative notion that I call potential vagueness, defined (informally) in (5).

(5) **Potential vagueness (informal):**

An adjective \( P \) is *potentially vague* (P-vague) iff there is some context \( c \) such that \( P \) gives rise to the Sorites paradox in \( c \).

I argue that RAs and AAs do display variability in how they instantiate (potential) vagueness; however, it is *complement*-based, rather than context-based, as is often claimed in the literature. In particular, I show that RAs have both P-vague positive forms (\( P \)) and negative forms (\( \neg P \)), while, for AAs, only one of the two are P-vague. I show that whether an AA has a P-vague positive or negative form is straightforwardly predictable from which well-established scale-structure AA subclass it belongs to: the total class or the partial class (cf. Cruse, 1980; Yoon, 1996; Rotstein and Winter, 2004, among others). More precisely, I show that AA\( ^T \)'s (e.g. empty, straight, clean) have P-vague positive forms and non-P-vague negative forms; whereas, AA\( ^P \)'s (e.g. wet, dirty, bent) have P-vague negative forms and non-P-vague positive forms (Table 3.2).

A formal analysis of the context sensitivity and vagueness properties of the different classes of adjectival predicates will be given in Chapter 4 within a Delineation semantics extension of the Tolerant, Classical, Strict framework that was introduced in Chapter 2.

This chapter is laid out as follows: The first part is devoted to the study of context sensitivity patterns in the adjectival domain. In Section 3.2, I present the context sensitivity patterns exhibited by the classes of adjectives in Table 3.1. Then, in Section 3.3, I present a discussion of the universal/existential context sensitivity distinction outside the adjectival domain. In particular, I suggest that this distinction broadly corresponds to the difference between the phenomena known

<table>
<thead>
<tr>
<th>Pattern</th>
<th>RA</th>
<th>AA(^T)</th>
<th>AA(^P)</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-vague ~( P )</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>P-vague ( P )</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>
as indexicality/saturation and imprecision/loose talk/modulation in the literature. The second part of the chapter is devoted to the study of the vagueness patterns associated with the adjectives in (1)–(4). In Section 3.4, I introduce and exemplify the notion of potential vagueness and show how it is realized across subclasses of adjectives. Finally, Section 3.5 gives a basic summary of the main empirical proposals argued for in this chapter in anticipation of the Delineation Tolerant, Classical, Strict analysis which will be given in Chapter 4.

3.2 Adjectival context sensitivity patterns

This section presents the data concerning the distribution of context sensitivity patterns in the adjectival domain. It has been observed since at least Sapir (1944) that the syntactic category of bare adjective phrases can be divided into two principled classes: scalar (or gradable) vs non-scalar (non-gradable). The principle test for scalarity of an adjective P is the possibility of P to appear in the explicit comparative construction. Thus, we find a first distinction between scalar adjectives like tall, expensive, straight, and empty (6) on the one hand and non-scalar atomic, pregnant, geographical, and hexagonal on the other (7).1

(6) a. John is taller than Phil.
   b. This watch is more expensive than that watch.
   c. This road is straighter than that one.
   d. My cup is emptier than your cup.

(7) a. ?This algebra is more atomic than that one.
   b. ?Mary is more pregnant than Sue.
   c. ?This map is more geographical than that one.
   d. ?This shape is more hexagonal than that one.

Furthermore, scalar adjectives are just those that can appear with other kinds of degree modifiers like very, so, and this.

(8) a. John is very/so/this tall.
   b. This watch is very/so/this expensive.
   c. This road is very/so/this straight.
   d. My cup is very/so/this empty.
   e. This towel is very/so/this dry.

1 Note that the non-scalars can be very easily transformed into scalar adjectives (i.e. Mary is more pregnant than Sue: she’s farther along). Gradable uses of NSs will be discussed later in this section.
(9)  a. ?This algebra is very/so/this atomic.
    b. ?Mary is very/so/this pregnant.
    c. ?This map is very/so/this geographical.
    d. ?This shape is very/so/this hexagonal.

Since Unger (1975), it is common to propose the further division of the class of scalar adjectives into two subclasses: what are often called the relative class and the absolute class. In particular, (following others) I show that, in languages like English, adjectives such as tall and expensive pattern differently from ones like straight and empty with respect to a variety of tests associated with how their denotations can vary across contexts. The tests that I present in this section are only a very small subset of the context sensitivity-based diagnostics for the RA/AA distinction described in the literature, and the reader is referred to works such as Unger (1975), Kyburg and Morreau (2000), Kennedy and McNally (2005), Kennedy (2007), Récanati (2010), Foppolo and Panzeri (2011), van Rooij (2011c), McNally (2011), and Toledo and Sassoon (2011) for more information.

The first way in which we can see the difference in context sensitivity between RAs and AA\(^T\)s and AA\(^P\)s is through the definite description test. As observed by, for example, Kyburg and Morreau (2000), Kennedy (2007), Syrett et al. (2010), Foppolo and Panzeri (2011), adjectives like tall and empty differ in whether they can “shift” their thresholds (i.e. criteria of application) to distinguish between two individuals in a two-element CC when they appear in a definite description. For example, as shown in Figure 3.1, suppose there are two containers (A and B), and neither of them are particularly tall; however, A is (noticeably) taller than B. In this situation, if someone asks me (10), then it is very clear that I should pass A. Now suppose that container A has less liquid than container B, but neither container is particularly close to being completely empty. In this situation, unlike what we saw with tall, (11) is infelicitous.

(10)  Pass me the tall one.

(11)  Pass me the empty one.

In other words, unlike RAs, AAs cannot change their criteria of application to distinguish between objects that lie in the middle of their associated scale. Using this test, we can now make the argument that adjectives like full, straight, and dry are absolute, since (12-a) is infelicitous if
neither object is (close to) completely full/straight/dry. Likewise, we can make the argument that *dirty*, *wet*, and *bent* are also absolute, since (12-b) is infelicitous when comparing two objects that are at the middle of the dirtiness/wetness/curvature scale (i.e. both of them are dirty/wet/bent).

(12) **Absolute adjectives**
   a. Pass me the full/straight/dry one.
   b. Pass me the dirty/wet/bent one.

Furthermore, we can make the argument that *long*, *expensive*, and *nice* are relative, since (13) is felicitous when comparing two objects when both or neither are particularly long/expensive/nice.

(13) Pass me the long/expensive/nice one.

A correlation of the observation that AAs are not context sensitive in the same way as RAs is the observation that only members of the latter class permit modification by a prepositional phrase headed by *for* that puts some restriction on a contextually given CC (Siegel, 1979).

(14) a. John is tall/short for a 3-year-old.
    b. This delicious baguette is expensive for Los Angeles/Paris.

AAs are more resistant to being modified by an expression that makes reference to a CC (cf. McNally, 2011; Toledo and Sassoon, 2011; and Bylinina, 2011, for discussion).

(15) a. #This towel is wet/dry for a used towel.
    b. #This glass is full/empty for a plastic glass.
40 CONTEXT SENSITIVITY AND VAGUENESS PATTERNS

(16) Compared to the glass on the table, this glass is full. McNally (2011, p. 159)

Of course, saying that AAs are not at all context sensitive is clearly false. As discussed by very many authors, such as Austin (1962), Unger (1975), Lewis (1979), Pinkal (1995), Kennedy and McNally (2005), Kennedy (2007), and Récanati (2010), although they may not be able to shift their semantic denotation to distinguish between any individuals on their scales, it is easy to see that their criteria of application can change depending on at least some contexts. For example, if we consider a particular large theater with two spectators in it, the same theater might be considered empty in the context of evaluating attendance at a play (17-a), however, it might not be considered so in the context of ensuring that no one is left inside during a fumigation or demolition process (17-b).

(17) a. Only two people came to opening night; the theater was empty.
   b. Two people didn’t evacuate; the theater wasn’t empty when they started fumigating.

Likewise, a road that has some twists in it might be considered straight in a context in which we are trying to avoid getting car sick, but it may no longer be considered so in a context in which we are surveying the land. And it is very easy to think of similar cases that show the context sensitivity of adjectives like flat, dry, clean, etc.

While these examples show us that absolute scalar predicates are, after all, context sensitive, it is important to observe that, in line with the data discussed in Section 3.1, the context sensitivity of predicates like empty is more restricted than that of predicates like tall. Crucially, the examples in (17) involve shifting the application of an AA from only objects at the endpoint of the scale to those that lie very close to the endpoint (like theatres with two people or roads with few bends). In other words, all of these examples involve what are known in the literature as “rough” (cf. Austin, 1962), “loose” (Unger, 1975; Sperber and Wilson, 1985) “modulated” (Récanati, 2004, 2010), or “imprecise” (Pinkal, 1995; Kennedy and McNally, 2005, among others) uses. Furthermore, as observed by McNally (2011), Toledo and Sassoon (2011), and Bylinina (2011), once we are in a context in which a loose use of an AA is possible, modification by for phrases becomes much more acceptable. For instance, the sentence in (18) is most felicitous in a context in which we are describing a
restaurant with a couple of people in it. It is bizarre if the restaurant is completely empty.

(18) This restaurant is empty for a Friday night.

Data such as (18) and those discussed in the works of McNally and others suggest not only that CCs are relevant for some aspects of the meaning of AAs, but also that they interact with the phenomenon of imprecision/loose talk.

Finally, through looking at the distribution of for phrases with AAs, we can see the existence of an interaction between scalarity and loose talk. As observed by McNally (2011), adding an explicit scalar modifier that moves the threshold of application of the adjective away from the endpoint greatly facilitates the presence of a for phrase: while (19-a) is awkward unless the context makes it very clear that full is being used imprecisely, (19-b) with a scalar modifier that forces an imprecise use is fine.

(19) a. For a Friday, the dentist’s schedule is full.
    b. For a Friday, the dentist’s schedule is very full.

In other words, for phrases are compatible with loose uses of AAs, and scalar modifiers enforce this use. McNally’s observations about data such as (19-b) gives us a first connection between imprecision or “loose talk” associated with AAs and the distribution of scalar modifiers. This connection will be made more explicit throughout the book particularly in Chapter 5, which presents a context sensitivity and vagueness-based theory of scale structure.

Based on the empirical observations made in this section, I argue that we can identify two types of context sensitivity in the adjectival domain: The first pattern, which I will descriptively call universal context sensitivity, corresponds to the ability to shift one’s threshold in any non-trivial CC. As we saw above, RAs have this property. Thus, tall can shift its criteria of application in the CC in Figure 3.2, but empty cannot.

---

2 Of course, the objects being compared must be perceived as distinct with respect to a dimension: if the height of two objects is so close that they seem to have roughly the same height, then tall will not distinguish between them. Note also that there is a difference between definite descriptions with the positive form of the adjective and the comparative form. Using the tall one seems to require that the object picked out exceed its comparison classmates by a greater degree than using the taller one. See Kennedy (2007), Kennedy (2011), Syrett et al. (2010), and van Rooij (2011c) for discussion. This point will be addressed in Chapter 4.
The second pattern (which I will call **existential context sensitivity**) corresponds to the ability to shift one’s threshold in some CC (but not necessarily the minimal CCs). AAs have this property, and I suggested that this kind of context sensitivity is related to the pragmatic phenomenon of “loose talk” or “imprecision”. If we consider again the predicate *empty*, now with reference to the two CC in Figure 3.3: while the leftmost container in this figure has a negligible amount of liquid in it and may not be considered empty in the upper CC, it may be considered empty in the lower CC.

---

3 Recall that I argued that CC are involved in at least some uses of AAs, as shown in the discussion concerning examples (18), (19-a), and (19-b).
I now consider the context sensitivity properties of non-scalar adjectives. It is easy to see that, like AAs, NS adjectives are not universally context sensitive. Firstly, as shown in (20), they uniformly fail the definite description test: they are only licit in contexts in which exactly one object is atomic/prime/hexagonal.

(20) a. Point to the atomic one.
(But neither/both is/are atomic.)
b. Point to the prime one.
(But neither/both is/are prime.)
c. Point to the hexagonal one.
(But neither/both is/are hexagonal.)

Secondly, as shown in (21), they are much more awkward with for phrases than are RAs.

(21) a. ?This algebra is atomic, for a boolean algebra.
b. ?This shape is hexagonal, for a shape in a geometry textbook.

What about existential context sensitivity? A famous example in the literature of a context sensitive use of hexagonal (originally due to Austin (1962) and discussed in the context of vagueness and imprecision in Lewis (1979)) is the one in (22).

(22) France is hexagonal.

If we are comparing France to shapes in geometry textbooks, it will not be considered hexagonal (its coastline has very many more “sides” than six!); however, when we are comparing it to other countries, all of whom also have bumpy coastlines, it may be considered hexagonal. Thus, we have found a case where the criteria for application of the predicate hexagonal vary depending on CC, and we can conclude that hexagonal, in what Austin calls its “rough” use, is context sensitive. In this way, at first glance, it may look as if we do find existentially context sensitive non-scalar adjectives.

But I believe that this conclusion would be premature. In fact, “rough” hexagonal is perfectly natural in the comparative construction.

(23) France is more hexagonal than Canada.

So, as soon as we are licensed by the context to apply hexagonal to France, we are licensed to compare things in terms of how close they
are to being in the extension of the non-scalar use of the predicate. (23) shows that, while the rough use of *hexagonal* is context sensitive, it is also scalar. In other words, it has been turned into an AA. Correspondingly, we can notice that, when we use a *for phrase* with a “non-scalar” adjective, a scalar modifier is not only possible, but, in fact, ameliorates the example (24).

(24) a. ?France is hexagonal, for a country.

b. France is very hexagonal, for a country.

I therefore propose that non-scalar adjectives are neither universally nor existentially context sensitive.

The observation that non-context sensitive non-scalar predicates have properly existential context sensitivity scalar counterparts is a general one. For example, if we consider a context sensitive use of *pregnant* as in (25-a) (someone can be considered pregnant if they meet the medical criteria in one context, but not be considered pregnant if they do not display the characteristic properties of pregnancy in another context), then we see that the comparative is licensed (25-b).

(25) a. Mary is technically pregnant, but she’s not showing and doesn’t go on and on about how wonderful pregnancy is (like Jane does), so she’s not really pregnant.

b. Jane is more pregnant than Mary.

We can replicate these examples with other non-scalars: *illegal, Canadian, and dead*.

(26) a. Smoking marijuana in Montréal is prohibited by law, but the police do not ever arrest anyone for it, like they do for breaking and entering. So smoking pot is not really illegal.

b. Breaking and entering is more illegal than smoking pot.

(27) a. Although both Heather and Dominique have Canadian citizenship, Dominique only lived in Canada for 8 years, has a European citizenship and accent. So, in most situations, you would not call Dominique Canadian.

b. Heather is more Canadian than Dominique.

4 Note, of course, that even ‘gradable’ *hexagonal* fails the definite description test: to be considered loosely *hexagonal*, an object has to be considered to be at least somewhat close to having six sides.
Table 3.3 Adjectival context sensitivity patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>RA</th>
<th>AA^I</th>
<th>AA^P</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal context sensitivity</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Existential context sensitivity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>

(28) a. Both zombies are dead; however, unlike zombie B, zombie A is highly mobile and chasing after us to eat our brains. So zombie A is not really dead.

b. Zombie B is more dead than zombie A.

Data associated with gradable uses of non-scalar adjectives show us that context sensitivity and scalarity go hand in hand. In particular, based on observations made in this section, I propose a new empirical generalization regarding gradability in the adjectival domain:

(29) **Scalarity generalization:**

An adjective is scalar iff it is existentially context sensitive.

This intimate connection between context sensitivity and scalarity that (I argue) we see in natural language will be the driving force behind the new theory of the origin of scale structure distinctions that will be developed throughout the course of this work.

In summary, in the first part of this chapter, I argued that we find the context sensitivity patterns in the adjectival domain shown in Table 3.3.

Furthermore, I proposed that there exists an important dependency between context sensitivity and gradability in natural language (cf. (29)).

In the next section, I look more closely at the universal/existential context sensitivity distinction with a view to bringing the observations that I made about adjectives in line with a more general theory of context sensitivity and the semantics/pragmatics interface.

### 3.3 Universal vs existential context sensitivity

As I have already mentioned, the patterns that distinguish universal and existential context sensitivity are very similar to patterns that characterize a distinction that is commonly made in the fields of linguistic semantics and pragmatics between context-based meaning variation due to indexicality (cf. Morris, 1938; Bar-Hillel, 1954; Montague, 1968; Kaplan, 1989 among very many others) and context-based meaning variation...
due to imprecision/loose talk (cf. Austin, 1962; Unger, 1975; Lewis, 1979; Sperber and Wilson, 1985; Lasersohn, 1999; Récanati, 2004; Syrett et al., 2010). Broadly speaking, indexical expressions are those whose literal (i.e., semantic) meanings are “gappy”; that is, the context is required to make some contribution before the expression can be assigned any kind of referent whatsoever. For example, if we consider an indexical pronoun like I, without knowing what the context of utterance is, it is impossible to have any kind of idea who is designated by this expression.

(30) I am in Paris.

Furthermore, we can observe that the referent of I can change depending on context: if I utter the sentence in (30), then I refers to Heather Burnett, and the sentence is true. However, if the speaker in the context of utterance of (30) is my mother, then I designates my mother, and (since she is currently in Ottawa) the sentence is false.

Clear cases of meaning variation due to imprecision/loose talk show a different pattern. Many linguistic expressions of various syntactic and semantic categories are not indexical; that is, their semantic denotation is not dependent on the extra-linguistic context. An example of such an expression would be 17.00 CET, May 28, 2013. Without using this expression in context, it is possible to assign some referent to it: the point of time that consists precisely of 17.00 on May 28. However, we can nevertheless observe (following Lasersohn (1999) among others) that the actual period of time that may be designated by this expression in context can vary. For example, in a context in which a very precise computer is keeping track of the time (31-a), 17.00 CET on May 28, 2013 might designate only exactly 17.00. However, if someone uses (31-b) while recounting their weekend, 17.00 CET on May 28, 2013 will most likely designate a wider temporal interval consisting of multiple minutes around 17.00.

(31) a. The deadline for the submission of abstracts is 17.00 CET on May 28, 2013.
   b. The concert started at 17.00 CET on May 28, 2013.

There is a natural similarity between the extreme context-dependence of indexical pronouns and the extreme context-dependence of RAs, on the one hand, and the more moderate context-dependence of temporal expressions and AAs on the other. Without knowing what the appropriate contextually given CC for tallness is (basketball players?
jockeys?), it is impossible to have any idea about who the tall individuals are; however, despite the possible contextual variation that was studied above (Section 3.2), I know that I can always apply the predicate empty to a container that contains zero objects. Likewise, I know that a stick that does not have a single bend in it can be always referred to as straight, even regardless of the level of granularity that the context imposes. In other words, like the expressions of time in (31), AAs have a context-independent prototypical core extension that may be broadened if the context requires it; whereas the extension of an RA has no such context-independent meaning.²

In fact, (a subset of) the patterns discussed in Section 3.2 has already been analyzed as involving the indexicality (i.e. variable semantic denotation) vs imprecision (i.e. variable pragmatic denotation) distinction. For example, Syrett et al. (2010, p. 30) describe their analysis of the experimental results of the definite description test as follows,

(32) If our interpretation of the facts is correct, then we have experimental evidence for a distinction between two types of interpretative variability. One type, exhibited by relative GAs [gradable adjectives], is fundamentally semantic in nature and is based on the conventional meaning of particular expressions (or combinations thereof). A second type, exhibited by imprecise uses of maximum standard absolute GAs, is fundamentally pragmatic and involves computation of a set of alternative denotations and a judgment about which of them count as tolerable deviations from the actual, precise meaning of the expression.

Similarly, Récanati (2010, pp. 66–70) analyzes the difference in context sensitivity between RAs and AAs as being the result of differences in the indexicality of the literal meanings of the different classes: he proposes that members of the former class have hidden indexical arguments that must be saturated in order for their semantic denotation to be determined; while members of the latter class have a non-indexical literal meaning, but are subject to a context-dependent pragmatic modulation process that takes the literal meaning as input.

² Indeed, a more cognitively oriented way of analyzing the difference between RAs and AAs that reflects their context sensitivity properties might be to say that AAs have certain distinguished members (i.e. prototypes, cf. Rosch (1973), Lakoff (1987), among others) that must always be included in their semantic or pragmatic extension; whereas, RAs have no such members. I will briefly revisit this idea in Chapter 4.
The account of the RA/AA distinction that I will develop in this work will be firmly in the tradition espoused by the aforementioned authors (see also Sapir (1944), Unger (1975), and Lewis (1979) for proposals in this vein). In particular, I propose:

(33) **The RA/AA distinction:**
   a. RAs are indexical expressions; their semantic denotation is assigned relative to a contextually given CC. As such, as the value of the CC indexical varies, so too will their semantic denotation.
   b. AAs and NSs are not indexical expressions. As such, their semantic denotation does not vary across CCs.

In the second half of this chapter, I examine the patterns associated with the properties of vague predicates discussed in Chapter 2 that characterize the classes of adjectives discussed above.

### 3.4 Potential vagueness and adjectival vagueness patterns

As discussed in Chapter 2, the uncontroversial examples of vague predicates are RAs like *tall*, *long*, and *expensive*. These lexical items allow borderline cases, fuzzy boundaries and, provided certain basic conditions on the domain are met, they give rise to the Sorites. But what about AAs like *empty*, *straight*, and *clean*. Do AAs also have borderline cases and fuzzy boundaries? Are they tolerant and do they give rise to the Sorites?

On the one hand, it has been observed (by Pinkal (1995), Kennedy (2007), and others) that, in some contexts, the symptoms of vagueness with AAs disappear. As a first example, we might consider Kennedy’s (2007) discussion of the absolute predicate *straight*. He observes that, in some very special cases where our purposes require the object to be perfectly straight, it is possible to say something like (34).

(34) The rod for the antenna needs to be *straight*, but this one has a 1 mm bend in the middle, so unfortunately it won’t work.

(Kennedy, 2007, p. 25)

In this situation, *straight* has no borderline cases: even a 1 mm bend is sufficient to move an object from *straight* to *not straight*. Similarly, the boundary between *straight* and *not straight* is sharp and located between
the perfectly straight objects and those with any small bend. We can see
the same pattern with empty. Suppose, instead of evaluating the success
of a play, we are describing the process of fumigating a theatre. In this
case, since having even a single person inside would result in a death, the
cut-off point between empty theaters and non-empty theaters would be
sharply at “one or more spectators.”

On the other hand, it is easily observed that, in at least some contexts,
straight and empty display certain properties that are eerily similar to
the properties displayed by tall and expensive. Likewise, suppose we are
going on a car trip and, since I get car sick very easily, we only want
to drive on straight roads. Clearly, it is not important in this situation
that the road that we take have absolutely no bends at all (we would
never go anywhere!). In this context, then, adding or subtracting a single
bend to a road would not make a difference in whether we would call
it straight; that is, in this context (unlike in the context described by
Kennedy above), ±1 mm is an indifference relation for straight. As such,
in this (tolerant) context, I think that we would assent to the tolerance
statement in (35).

(35) For all \(x, y\), if \(x\) is straight and \(x\) and \(y\) differ by a single \(1\) mm bend,
then \(y\) is also straight.

Likewise, consider a context in which we are talking about theaters
and whether or not a particular play was well attended. In this kind
of situation, we often apply the predicate empty to theaters that are not
completely empty (i.e. those with a couple of people in them), and, in
this context, it is tolerant; if we are willing to call a theater with a couple
of people in it empty, then at what number of spectators does it become
not empty?

In sum, at least in some contexts, AAs also display the characteristic
properties of vague language, and I propose that the similarities between
the fuzziness of tall and the fuzziness of loose uses of empty strongly
suggest that we are dealing with a single phenomenon at work in both
cases. Furthermore, we have seen both contexts in which AAs display
the characteristic properties of vagueness and contexts in which they
do not; thus, when we ask whether AAs are vague, I conclude that the
appropriate answer to this question is “sometimes (but not always).”
In other words, I argue that being vague (by which I mean “exhibiting
the cluster of properties discussed in Chapter 2”) is a stage-level prop-
erty, i.e. one that is subject to contextual variation. This picture is at
odds with the traditional use of the term *vague* (beginning with Peirce (1901)) which takes it to be an individual-level, context-independent property. Thus, I propose that, in order to account for the empirical patterns described above, and in the literature on vagueness and the absolute/relative distinction, we should employ a more nuanced notion, one that makes the contribution of the context fully explicit. I suggest that this notion is *potential vagueness*, as defined in (36).

(36) **Potential vagueness** (formal):
An adjective $P$ is $P$-vague iff there is some context $c$ such that $P$ gives rise to a Soritical argument in $c$.

In what follows, I argue that the *potential vagueness* property is yet another way through which RAs and AAs can be distinguished.

### 3.4.1 (A)symmetric vagueness

We saw in Chapter 2 that *tall* was P-vague: we found some context in which *tall* was tolerant, for example, if we are evaluating the height of North American males and we take the relation ‘± 1 mm’ to be an appropriate indifference relation. Moreover, we can make the same observation about *not tall*: in the context of evaluating the heights for North American males, at what point does adding 1 mm to the height of a ‘not tall’ man change them into a ‘tall’ man? In the contexts in which ‘± 1 mm’ counts as an irrelevant change, then *not tall* will also be tolerant; that is, in this context, we will generally assent to both the statements in (37).

(37) **Potential vagueness of tall and not tall**:
   a. **Tall**:
      For all $x, y$, if $x$ is tall and $x$ and $y$’s heights differ by 1 mm, then $y$ is tall.
   b. **Not tall**:
      For all $x, y$, if $x$ is not tall and $x$ and $y$’s heights differ by 1 mm, then $y$ is not tall.

I will refer to the property of having both a P-vague positive and negative form as being *symmetrically vague*.

**Definition 3.4.1 Symmetric vagueness.** A predicate $P$ is symmetrically vague iff $P$ is $P$-vague and ‘not $P$’ is $P$-vague.

I argue that AAs, on the other hand, display a different pattern. Consider, firstly, AA $^T$’s like *straight* and *empty*. I argued in Section 3.3 that these predicates were P-vague, and we can think of contexts (such as going on
a car trip or attending a play) in which we would assent to the principle of tolerance for these predicates. But we can observe that the negations of AA’s behave differently. In particular, even in the same contexts as described above, the principle of tolerance is not valid for not straight and not empty (38).\(^6\)

(38) **Intolerant not straight and not empty:**

a. **False:** For all \(x, y\), if \(x\) is not straight and \(x\) and \(y\)'s shapes differ by a single 1 mm bend, then \(y\) is not straight.

b. **False:** For all \(x, y\), if \(x\) is not empty and \(x\) and \(y\)'s contents differ by a single item, then \(y\) is not empty.

Crucially, the statements in (38) are falsified by the cases where we move from individuals who are at the endpoint of the relevant scale to those who lie at the second to last degree: in particular, when we move from a road with a single millimetre bend in it (so one which is not straight) to a perfectly straight road, i.e. one which cannot be called not straight, even in this context. Similarly with empty: (38-b) is falsified by the case where \(x\) has a single spectator (so is not empty) and \(y\) contains no spectators (so is not not empty). Although I have been discussing these data in the context of tolerance and Soritical arguments (since one of the aims of this chapter is to make a connection between the reasoning patterns associated with AAs and similar patterns traditionally analyzed as instances of vague language), the fundamental asymmetry outlined in this section, which will interest us in the rest of the book, is actually quite simple and is shown in (39).

(39) **Asymmetric similarity judgment:**

Although, depending on the context, we might consider an object that is not perfectly straight/empty to be straight/empty, we will never consider an object that is perfectly straight/empty to be not straight/empty.

The observation that we will never consider an object that is perfectly straight not straight is very similar to (and indeed, as I argue in Chapter 6, may even be considered an instantiation of) ideas by Pinkal (1995) and Kennedy (2007), who propose that AAs have natural precisifications and are not vague in the same way as RAs. This being said, the other observation (i.e. that precision is non-symmetric) actually makes the

\(^6\) Note that we are analyzing the expression *if...then* as a material implication.
empirical proposal in (39) weaker than the Pinkal/Kennedy proposal. A full comparison of the theory developed in this work and the alternative analysis of Kennedy (2007) is given in Chapter 6.

Thus, with (39) in mind, I will refer to the property of differing in vagueness with one’s negation as being *asymmetrically vague*:

**Definition 3.4.2 Asymmetric vagueness.** A predicate $P$ is asymmetrically vague iff one of $\{P, \text{not } P\}$ is not $P$-vague.

We can now be more precise about how total predicates are $P$-vague as in (40), and the reader is encouraged to verify that this correlation does hold of the entire list of $AA^T$s at the beginning of this chapter.

(40) **AA$^T$s generalization:**

$Q$ is an $AA^T$ iff $Q$ is $P$-vague and $\text{not } Q$ is not $P$-vague.

What about partial $AA^P$s? We can immediately see a difference between adjectives like *wet*, *dirty* and *empty*, *straight*: the negations of partial adjectives are $P$-vague. For example, suppose I am getting out of the shower and I need a towel to dry myself with. In this situation, I need to pick a towel that is not wet; however, it is not going to make a huge difference to my purposes if there happen to be a few stray drops of water on the towel that I pick. Thus, in this situation, we can pick ± one drop of water as an indifference relation for *wet*, because how could adding or substracting a single drop of water to a towel make a difference to whether or not I can dry myself with it?

Contrary to what we saw with *straight* above, however, this time it is *not wet* that is tolerant (41-a). And we can make the same observation with *not dirty*: this negated predicate will satisfy the principle of tolerance in cases where one speck of dirt is perceived as irrelevant (41-b).

(41) **Tolerance of not wet and not dirty:**

a. For all $x, y$, if $x$ is not wet, and $x$ and $y$ differ by one drop of water, then $y$ is not wet.

b. For all $x, y$, if $x$ is not dirty, and $x$ and $y$ differ by one speck of dirt, then $y$ is not dirty.

However, with $AA^P$s, it is the positive form of the adjective that is not $P$-vague: even if a single drop/speck is perceived as irrelevant, *wet* and *dirty* do not satisfy tolerance. In particular, objects that are completely dry and completely clean cannot ever be described as *wet* or *dirty* respectively.
(42) Intolerance of wet and dirty:
   a. False: For all \( x, y \), if \( x \) is wet, and \( x \) and \( y \) differ by one drop of water, then \( y \) is wet.
   b. False: For all \( x, y \), if \( x \) is dirty, and \( x \) and \( y \) differ by one speck of dirt, then \( y \) is dirty.

Again, I highlight the existence of the asymmetric similarity judgment in (43).

(43) Asymmetric similarity judgment:
   Although, depending on context, we might consider an object that has one drop of water on it to be not wet, we will never consider a bone-dry object to be wet.

Thus, \( AA_P \)’s are also asymmetrically vague and conform to the generalization in (44).

(44) \( AA_P \) generalization:
   \( Q \) is an \( AA_P \) iff \( Q \) is not \( P \)-vague and not \( Q \) is.

Finally, consider NSs like atomic or hexagonal: as discussed in Égré and Klinedinst (2011) and van Rooij (2011c) (among others), these constituents are typical examples of precise predicates, and we can verify that both their positive forms and negative forms are intolerant in all contexts.\(^7\)

(45) Intolerance of prime and hexagonal:
   a. False: For all numbers \( x, y \), if \( x \) is prime and \( x \) and \( y \) differ by one, then \( y \) is prime.
   b. False: For all shapes \( x, y \), if \( x \) is hexagonal and \( x \) and \( y \) differ by one side, then \( y \) is hexagonal.

(46) Intolerance of not prime and not hexagonal:
   a. False: For all numbers \( x, y \), if \( x \) is not prime and \( x \) and \( y \) differ by one, then \( y \) is not prime.
   b. False: For all shapes \( x, y \), if \( x \) is not hexagonal and \( x \) and \( y \) differ by one side, then \( y \) is not hexagonal.

\(^7\) Note that we are only talking about non-gradable uses of NSs. We can observe that “loose” uses of non-scalars display the characterizing properties of vague language: how many sides does an object need to have before it cannot be considered loosely hexagonal?
Thus, by looking at how adjectival predicates behave in (contextually appropriate) Soritical arguments, we can replicate the traditional non-scalar/relative/partial/total scale structure typology as in Table 3.4.

### 3.5 Conclusion

This chapter gave a description of the distribution of two semantic/pragmatic phenomena in the adjectival domain: context sensitivity and (potential) vagueness. I argued that the basic scale structure typology that we find in the adjectival domain can be (almost) completely generated with reference to context sensitivity patterns and completely generated with reference to potential vagueness. Furthermore, I argued that the potential vagueness notion, which I proposed, allows us to pursue a unified analysis of the Sorites susceptibility of both RAs and AAs, something that is both empirically and theoretically justified. A summary of the main context sensitivity and vagueness patterns described in this work is shown in Table 3.5.

In Chapter 4, I develop an analysis of these patterns within Delination Tolerant, Classical, Strict, which is a new logical framework that I create for modelling the relationships between context sensitivity, vagueness, and scale structure in natural language.

---

**Table 3.4** (Potential) vagueness patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>RA</th>
<th>( AA^T )</th>
<th>( AA^P )</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-vague ( \neg P )</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>P-vague ( P )</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

---

**Table 3.5** Correspondences between context sensitivity and potential vagueness

<table>
<thead>
<tr>
<th>Pattern</th>
<th>RA</th>
<th>( AA^T )</th>
<th>( AA^P )</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Context Sensitivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal context sensitivity</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Existential context sensitivity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Potential Vagueness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-vague ( \neg P )</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>P-vague ( P )</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>
The Delineation Tolerant, Classical, Strict framework

4.1 Introduction

This chapter presents the Delineation Tolerant, Classical, Strict (DelTCS) framework: a new logical architecture for modelling the relationships between context sensitivity, vagueness, gradability, and scale structure in natural language. In simplest terms, DelTCS is a Tolerant, Classical, Strict extension of a (simplified) version of the system proposed in Klein (1980), which belongs to a class of logical systems known (after Lewis (1970)) as Delineation semantics (DelS) systems. Within the DelTCS framework, I set an analysis of the semantic and pragmatic properties of relative (RA), absolute (AA), and non-scalar (NS) adjectives, and I show how my analysis accounts for the context sensitivity and potential vagueness patterns described in Chapter 3. Furthermore, I show that, by virtue of the logical structure of the framework, the analysis of the RA/AA/NS distinction makes correct predictions concerning the gradability properties of the respective classes of adjectives. Other predictions of the proposal, including scale structure, modification, and inferential patterns associated with the different kinds of adjectives, will be discussed in Chapter 5.

The chapter is therefore laid out as follows: In Section 4.2, I define the language of DelTCS and give a DelS analysis of the classical semantic denotations of RAs, AAs, and NSs. Then, in Section 4.3, I extend the basic Delineation analysis using the structure of the TCS non-classical logic and make proposals concerning the tolerant and strict interpretations of these predicates. With the full proposal in place, in Section 4.4, I show that my analysis set within DelTCS correctly predicts the context sensitivity, vagueness, and gradability properties that have been outlined so far. Finally, Section 4.5 summarizes the main theoretical proposals made in the chapter and their key empirical consequences.
4.2 Language and classical semantics

The analyses that I give in this chapter deal with modelling the semantics and pragmatics of the following range of syntactic constructions: copular sentences with the positive form of relative adjectives (1-a), the positive form of total/partial absolute adjectives (1-b), comparatives with relative adjectives (1-c), comparatives with total/partial absolute adjectives (1-d), and, finally, positive NSs (1-e).

(1) a. John is tall.
   John is not tall.
b. John is clean/wet.
   John is not clean/wet.
c. John is taller than Peter.
   John is not taller than Peter.
d. John is cleaner/wetter than Peter.
   John is not cleaner/wetter than Peter.
e. John is dead.
   John is not dead.

Therefore, the language of DelTCS has the following vocabulary:

1. A series of individual constants: \(a_1, a_2, a_3 \ldots\)
2. Four series of unary predicate symbols:
   • Relative scalar adjectives (RA): \(P, P_1, P_2, P_3 \ldots\)
   • Total absolute scalar adjectives (AA\(^T\)): \(Q, Q_1, Q_2, Q_3 \ldots\)
   • Partial absolute scalar adjectives (AA\(^P\)): \(R, R_1, R_2, R_3 \ldots\)
   • Non-scalar adjectives (NS): \(S, S_1, S_2, S_3 \ldots\)
3. For every unary predicate symbol \(P\), there is a binary predicate \(> P\).
4. A one-place connective \(\neg\).

I will often refer to the entire class of scalar adjectives as \(SA\) (\(RA \cup AA = SA\)). Also, if the relative/absolute distinction is irrelevant for a particular definition, I will use members of the \(P\) series to notate members of \(SA\) and members of the \(Q\) series to notate members of \(AA \cup NS\). I trust that this will not cause confusion. The syntax of DelTCS is as follows:

1. Constants (and nothing else) are terms.
2. If \(t\) is a term and \(P\) is a predicate symbol, then \(P(t)\) is a well-formed formula (wff).
3. If $t_1$ and $t_2$ are terms and $P$ is a predicate symbol, then $t_1 >_P t_2$ is a wff.
4. If $\phi$ is a wff, then $\neg \phi$ is a wff.
5. Nothing else is a wff.

4.2.1 Classical semantics for relative adjectives

The heart of the proposal lies in the semantics for this very simple language. In particular, I adopt a simplified version of Klein’s (1980) framework, which, as mentioned in the introduction, is a DelS system.

DelS is a framework for analyzing the semantics of gradable expressions that takes the observation that they are context sensitive to be their key feature. Delineation approaches to the semantics of positive and comparative constructions were first proposed by McConnell-Ginet (1973), Kamp (1975), Lewis (1979), and Klein (1980) and this general framework has been further developed by van Benthem (1982), Keenan and Faltz (1985), Larson (1988), Klein (1991), van Rooij (2011b), Doetjes (2010), van Rooij (2011a), Doetjes et al. (2011), Burnett (2014a), among others. In what follows, I will present a very basic version of the theory because it will be sufficient to account for the data discussed in this book. However, presumably, more enriched theories, such as those proposed by the authors cited above, will be necessary to account for the wide range of scale-based constructions in natural language.¹

In this framework, scalar adjectives denote sets of individuals and, furthermore, they are evaluated with respect to comparison classes (CCs), i.e. subsets of the domain. The basic idea is that the extension of a gradable predicate can change depending on the set of individuals that it is being compared with. For example, consider the predicate tall and the graphic (based on Klein (1980, p. 18)) in Figure 4.1.

![Figure 4.1 Two-element (minimal) comparison class X](image)

¹ See Kennedy (1997) for challenges raised by certain kinds of empirical phenomena (cross-polar anomaly, incommensurability, etc.) for the Delineation approach.
If we apply the predicate *tall* to the elements in the minimal two-element CC $X = \{u, v\}$, then the extension of *tall* in $X$, written $\llbracket \text{tall} \rrbracket_X$, would be $\{u\}$, and $\llbracket \text{not tall} \rrbracket_X = \{v\}$ (Figure 4.2).

Now consider the larger CC $X'$ (also based on Klein (1980, p. 18), Figure 4.3).

If we apply the predicate *tall* in $X'$, despite the fact that their actual sizes have not changed, it is conceivable that both $u$ and $v$ could now be in $\llbracket \text{not tall} \rrbracket_{X'}$ (cf. Figure 4.4).

These examples illustrate how the semantic denotation of an RA can be relativized to and vary depending on CCs. More formally, we define our models and satisfaction in them as follows:

**Definition 4.2.1 Classical model.** A c-model is a tuple $M = \langle D, \llbracket \cdot \rrbracket \rangle$ where $D$ is a non-empty domain of individuals, and $\llbracket \cdot \rrbracket$ is a function from pairs consisting of a member of the non-logical vocabulary and a CC (a subset of the domain) satisfying:

![Figure 4.2 Application of *tall* in comparison class $X$](image1)

![Figure 4.3 Four-element comparison class $X'$](image2)

![Figure 4.4 Application of *tall* in $X'$](image3)
1. For each individual constant $a_1$, $[a_1] \in D$.

2. For each $X \subseteq D$ and for each predicate $P$, $[P]_X \subseteq X$.

Observe that, unlike in first order logic where predicates are assigned any subset of the domain (cf. Chapter 2), in the Delineation analysis presented here, predicates are assigned different properties in different CCs (i.e. subsets of the domain). Moreover, for simplicity, we assume that the interpretation of a predicate is defined on all members of the powerset of $D$. I will discuss imposing more restrictions on both the definition of $[\cdot]$ and the CCs on which $[\cdot]$ is defined later in this section. For a given utterance of a sentence containing the positive form of a scalar adjective, the relevant CC against which the statement is evaluated is given purely by context. Thus, in our analysis, truth in a model is always given with respect to a distinguished CC. Note that if the subject of the sentence is not included in the distinguished CC, the truth value of the sentence is undefined (as suggested by our judgments concerning sentences like (2)).

(2) # Mary is tall for a boy in this class.

**Definition 4.2.2 Classical semantics of the positive form.** For all models $M$, all CCs $X \subseteq D$, all predicates $P$, and individuals $a_1 \in X$,

\[
[P(a_1)]_{X,M} = \begin{cases} 
1 & \text{if } [a_1]_M \in [P]_{X,M} \\
0 & \text{if } [a_1]_M \in X - [P]_{X,M} \\
i & \text{otherwise}
\end{cases}
\]

We will take the basic semantics of negation to be classical, as shown below:

**Definition 4.2.3 Semantics of negation.** For all models $M$, $X \subseteq D$ and wffs $\phi$,

\[
[\neg \phi]_{X,M} = \begin{cases} 
1 & \text{if } [\phi]_{X,M} = 0 \\
0 & \text{if } [\phi]_{X,M} = 1 \\
i & \text{otherwise}
\end{cases}
\]

2. Note that this is an idealization: there are counterexamples to this pattern, such as (i) whose treatment goes beyond the scope of this work.

(i) Mia wants an expensive hat for a three-year-old.

See Schwartz (2010) and Solt (2011) for more discussion of these cases.

3 In the bulk of the book, for readability considerations, I will often omit the model notation, writing only $[\cdot]_X$ for $[\cdot]_{X,M}$.
The definitions given above constitute a very simple analysis of the semantics of positive and negative gradable adjectives. In fact, it is too simple. As discussed in Klein (1980) and van Benthem (1982), if we put no restrictions on how the denotations of scalar predicates can be applied across CCs, then we will allow some counterintuitive results. For example, suppose that we apply tall in the CC X as in Figure 4.2. So, in some CC, we have categorized u as tall and v as not tall. Suppose, furthermore, that, when we move to the larger CC X', instead of applying tall as in Figure 4.4, we apply it such that now v is tall and u is not tall (Figure 4.5).

At the moment, nothing in our definitions prohibits applying tall as in Figure 4.5. But clearly scalar predicates in natural language do not work like this. So we have a problem.

The standard solution to this problem involves imposing some constraints on how predicates like tall can be applied in different CCs. The proposal that the interpretation of scalar predicates is constrained by certain intuitive axioms is an integral part of the Delineation framework. Unlike other approaches (like Degree semantics (DegS), cf. Kennedy (1997) and Chapter 6 of this work for an overview of this framework) that put constraints on degree structures in the ontology, in the Kleinian framework, these constraints are put directly on how scalar predicates can be interpreted at different CCs. In other words, we can observe that the application of relative scalar predicates like tall in natural language is guided by certain monotonicity principles, and so we build these principles into the interpretation function.

There exist a few proposed constraint-sets in the literature (e.g. Klein (1980), van Benthem (1982), (van Rooij 2011a, 2011c)) and, indeed, analyses set within the DeS architecture can differ in which CC-based axioms they assume. Probably the best-known constraint-set is that of van Benthem (1982, 1990). Van Benthem proposes three axioms governing the categorization of individuals across CCs. They are the following (presented in my notation):
For all models $M$, all $a_1, a_2 \in D$ and $X \subseteq D$ such that $a_1 \in \llbracket P \rrbracket_{X,M}$ and $a_2 \notin \llbracket P \rrbracket_{X,M}$.

**Axiom 4.2.4 No Reversal (NR):** There is no $X' \subseteq D$ such that $a_2 \in \llbracket P \rrbracket_{X',M}$ and $a_1 \notin \llbracket P \rrbracket_{X',M}$.

**Axiom 4.2.5 Upward Difference (UD):** For all $X'$, if $X \subseteq X'$, then there is some $a_3, a_4 : a_3 \in \llbracket P \rrbracket_{X',M}$ and $a_4 \notin \llbracket P \rrbracket_{X',M}$.

**Axiom 4.2.6 Downward Difference (DD):** For all $X'$, if $X' \subseteq X$ and $a_1, a_2 \in X'$, then there is some $a_3, a_4 : a_3 \in \llbracket P \rrbracket_{X',M}$ and $a_4 \notin \llbracket P \rrbracket_{X',M}$.

NR states that if there is some CC in which $a_1$ is classified as $P$ and $a_2$ is classified as $not P$, there is no other CC in which they switch; in other words, NR rules out the weird application of tall in Figure 4.5 that we just discussed. UD states that if, in one CC, there is a $P/not P$ contrast, then a $P/not P$ contrast is preserved in every larger CC. In other words, if there is some reason that, in one CC, we made a distinction between some individuals with respect to $P$, adding extra individuals to CCs cannot erase all distinctions (although they might shift). Finally, DD says that if, in one CC, there is a $P/not P$ contrast involving $a_1$ and $a_2$, then there remains a contrast in every smaller CC that contains both $a_1$ and $a_2$.

In what follows, I will adopt van Benthem’s analysis outlined above as an analysis of RAs as a class and propose that the classical semantic interpretations of RAs across CCs are subject to (only) these three constraints. Of course, the proposal that the interpretation of all RAs is constrained only by these three very weak axioms is undoubtedly overly simplistic. It is well known that RAs can be divided into multiple subclasses based, for example, on their implicatures in various syntactic constructions. For example, the question in (3) with tall is neutral; however, the same question with short suggests that John is short. So we can make a distinction between tall and short with respect to this evaluativity property (in the words of Rett (2008)) in degree questions.

(3) a. How tall is John?
   Implies nothing.

   b. How short is John?
   Implies that John is short.

We can make a further distinction between short and ‘extreme’ RAs like brilliant: as shown in (4), a comparative with tall or short is not evaluative, but a comparative with brilliant is.
(4) a. John is taller/shorter than Mary.  
Implies nothing.  
b. John is more brilliant than Mary.  
Implies John is brilliant.

Although a full analysis of these contrasts is out of the scope of this work, the most promising way to analyze them in a DelS framework would be to propose that the interpretation of short is subject to some constraint(s) that tall is not subject to, and brilliant is subject to some constraints that neither tall nor short is subject to. This being said, since the goal of this work is to give an account of the interaction between context sensitivity, vagueness, and scale structure, and all RAs appear to behave the same way in the tests associated with these three phenomena, I will not make any finer distinctions within the RA class. Thus, for the purposes of this project, I assume that van Benthem’s NR, UD, and DD give us a characterization of RAs, at least where their context sensitivity and scale structure properties are concerned.

4.2.1.1 From context sensitivity to scalarity  Although the CC-based variation restricted by van Benthem’s axioms gives us a nice analysis of the extreme context sensitivity of predicates such as RAs, in fact, it gives us much more. A major feature of the Delineation approach is that the scalarity/gradability of an adjective is derived from its context sensitivity. The scales associated with particular adjectival predicates (as well as the denotation of the comparative) are defined as follows:

**Definition 4.2.7 Semantics for the comparative.** For all models $M$, all $X \subseteq D$, all individuals $a_1, a_2$, and predicates $P$,

$$\|a_1 \succ_P a_2\|_{X,M} = \begin{cases} 1 & \text{if there is some } X' \subseteq D : \|P(a_1)\|_{X',M} = 1 \text{ and } \|P(a_2)\|_{X',M} = 0 \\ 0 & \text{otherwise} \end{cases}$$

Informally, in this framework, *John is taller than Mary* is true just in case there is some CC with respect to which John counts as tall and Mary counts as not tall. Thus, in the example in the Figures 4.2–4.4, we can establish the ordering $u \succ_tall v$ since $u$ is tall in $X$ and $v$ is not tall in $X$ (cf. Figure 4.2). Furthermore, we can establish the orderings $t \succ_tall u$, $s \succ_tall u$, $t \succ_tall v$, and $s \succ_tall v$ from the CC $X'$ (Figure 4.4). More generally, van Benthem shows that these axioms give rise to strict weak orders: irreflexive, transitive, and almost connected relations.
**Definition 4.2.8 Strict weak order.** A relation $>$ is a strict weak order just in case $>$ is irreflexive, transitive, and almost connected.\(^4\)

As discussed in Klein (1980), van Benthem (1990), and van Rooij (2011b), strict weak orders (also known as ordinal scales in measurement theory) intuitively correspond to the types of relations expressed by many kinds of comparative constructions. For example, one cannot be taller than oneself; therefore, $>_\text{tall}$ should be irreflexive. Also, if John is taller than Mary, and Mary is taller than Peter, then we know that John is also taller than Peter. So $>_\text{tall}$ should be transitive. Finally, suppose John is taller than Mary. Now consider Peter. Either Peter is taller than Mary or he is shorter than John. Therefore, $>_\text{tall}$ should be almost connected. Thus, Theorem 4.2.9 is an important result in the semantic analysis of comparatives, and it shows that scales associated with gradable predicates can be constructed from the context sensitivity of the positive form and certain axioms governing the application of the predicate across different contexts.

**Theorem 4.2.9 Strict weak order.** For all $P \in RA$, $>_P$ is a strict weak order.

**Proof.** van Benthem (1982), van Benthem (1990, p. 116).

It is important to note that strict weak orders are weaker than linear orders (i.e. the kinds of orders assumed in DegS), but it is easy to construct a linear order from them in the following way (à la Krantz et al. (1971), van Rooij (2011b), Bale (2011), i.a.): first we define an equivalence relation $\approx$ based on $>$. 

**Definition 4.2.10 Equivalent ($\approx_P$).** For all $a_1, a_2 \in D$ and predicates $P$, $a_1 \approx_P a_2$ iff $a_1 \not>_P a_2$ and $a_2 \not>_P a_1$.

In other words, John and Mary are equivalent with respect to the predicate tall just in case John is not taller than Mary, and Mary is not

---

\(^4\) **Definition 4.2.11 Irreflexivity.** A relation $>$ is irreflexive iff there is no $x \in D$ such that $x > x$.

**Definition 4.2.12 Transitivity.** A relation $>$ is transitive iff for all $x, y, z \in D$, if $x > y$ and $y > z$, then $x > z$.

**Definition 4.2.13 Almost connectedness.** A relation $>$ is almost connected iff for all $x, y \in D$, if $x > y$, then for all $z \in D$, either $x > z$ or $z > y$. 
taller than John; that is, John and Mary 'have the same height'\(^5\) just in case there is no CC in which the predicate *tall* distinguishes them.

Now we can order the equivalence classes of individuals (i.e. \([a_1]_{\approx_P}\) is the set of individuals that are related to \(a_1\) by the \(\approx_P\) relation) in the following way:

**Definition 4.2.14** *Degree ordering* (\(\gg\)). For all predicates \(P\) and individuals \(a_1, a_2\):

\[ [a_1]_{\approx_P} \gg_P [a_2]_{\approx_P} \text{ iff for all } a_3 \in [a_1]_{\approx_P} \text{ and all } a_4 \in [a_2]_{\approx_P}, a_3 >_P a_4. \]

This derived ordering is a strict linear order.\(^6\) Thus, if necessary, we can collapse the \(>_P\) relations into linear orders to get a more degree-like structure (e.g. one to which measure functions can apply) in which each degree is an equivalence class of individuals. For further comparisons between the approach developed here and the DegS approach, see Chapter 6.

**4.2.1.2 Positive vs comparative forms** With Axioms 4.2.1–4.2.3, we can observe that the comparison relation has the following property: if \(a_1 >_P a_2\), then, if we look at the minimal two-element CC (i.e. if we compare \(a_i\) directly with \(a_2\), \(a_i\) will be \(P\) and \(a_2\) will be not \(P\). I call this property *two-element reducibility*.

**Theorem 4.2.15** *Two-element reducibility.* For all \(a_1, a_2 \in D\) and relative predicates \(P\) (i.e. predicates that obey NR, UD, DD), \(a_1 >_P a_2\) iff \(\|P(a_1)\|_{[a_1, a_2]} = 1\) and \(\|P(a_1)\|_{[a_1, a_2]} = 0.\(^7\)

---

\(^5\) Note that this equivalence relation should not be taken as an analysis of the equative construction (e.g. *John is as tall as Mary*), since this construction has its own particularities (cf. Rett (2008)) that go beyond the scope of this work.

\(^6\) Proof. Irreflexivity. Immediately by the irreflexivity of \(>_P\). Transitivity. Immediately by the transitivity of \(>_P\). Totality. Let \(a_1, a_2 \in D\) such that \([a_1]_{\approx_P} \neq [a_2]_{\approx_P}\). Suppose that \([a_2]_{\approx_P} \gg_P [a_1]_{\approx_P}\) to show that \([a_1]_{\approx_P} >_P [a_2]_{\approx_P}\). Suppose, for a contradiction, that there is some \(a_3 \in [a_1]_{\approx_P}\) such that \(a_3 \neq_P a_4\), for some \(a_4 \in [a_2]_{\approx_P}\). Since, by assumption, \([a_1]_{\approx_P}\) and \([a_2]_{\approx_P}\) are distinct, \(a_3 \neq_P a_4\). So \(a_3 >_P a_4\), since \([a_3]_{\approx_P} = [a_1]_{\approx_P}\) and \([a_4]_{\approx_P} = [a_2]_{\approx_P}\), by the fact that \(\approx_P\) is an equivalence relation (cf. Krantz et al. (1971)), \([a_3]_{\approx_P} >_P [a_1]_{\approx_P}\). \(\square\)

Where *totality* is defined as follows:

**Definition 4.2.16** *Total. A relation > is total if for all \(x, y \in D\), either \(x > y\) or \(y > x\).*

\(^7\) Proof. \(\Rightarrow\) Suppose \(a_1 >_P a_2\) to show \(\|P(a_1)\|_{[a_1, a_2]} = 1\) and \(\|P(a_2)\|_{[a_1, a_2]} = 0.\) Since \(a_1 >_P a_2\), there is some \(X \subseteq D\) such that \(\|P(a_1)\|_{X} = 1\) and \(\|P(a_2)\|_{X} = 0.\) Clearly \([a_1, a_2] \subseteq X\). So, by DD, there is some \(a_3, a_4 \in [a_1, a_2]\) such that \(\|P(a_3)\|_{[a_1, a_2]} = 1\) and \(\|P(a_4)\|_{[a_1, a_2]} = 0.\) By NR, \(\|P(a_1)\|_{[a_1, a_2]} = 1\) and \(\|P(a_2)\|_{[a_1, a_2]} = 0.\) \(\Leftarrow\) Immediately from the definition of \(>_P.\) \(\square\)
There are reasons to think that this prediction is too strong. As observed by Kennedy (2011) and van Rooij (2011a), there are some cases in which we would like to apply the comparative form of an adjective, but would not necessarily apply the positive form. Consider the following example (based on Kennedy (2011)): if we compare the size of the planets Uranus and Venus (schematized in Figure 4.6, based on Kennedy’s figure 1), it seems appropriate to say both the sentences in (5).

(5)  a. Uranus is the *bigger* one.
    b. Uranus is the *big* one.

This pattern is predicted by the analysis that I gave earlier in this section: it predicts that Uranus is bigger than Venus just in case, if we compare Uranus directly with Venus, then we will call Venus *big* and Uranus *not big*. However, we see a different pattern when we compare Uranus with another, larger, planet: Neptune (cf. Figure 4.7, based on Kennedy’s figure 2).

In this case, although we would assent to (6-a), we would generally deny (6-b).

(6)  a. Uranus is the *bigger* one.
    b. Uranus is the *big* one.

Similarly, in the situation exemplified in Figure 4.7, although (7-a) is appropriate, (7-b) is inappropriate.

(7)  a. Uranus is *bigger* than Neptune.
    b. Uranus is *big* compared to Neptune.
Thus, cases like the one just discussed are problematic for the simple Delineation analysis using van Benthem's axioms. Fortunately, van Rooij (2011a) gives an analysis of precisely this contrast within a DelS framework. I will not go through the details of van Rooij's analysis here, since it would take us too far afield, but the reader is referred to the paper for the technical aspects of the theory. In a nutshell, he gives two ways of solving this puzzle: one solution involves modifying the set of constraints that relative predicates obey, which, as I mentioned, is a natural move within DelS. The second solution involves adding a series of constraints on what can count as a "pragmatically appropriate" CC. In other words, while I have allowed the interpretation of relative predicates to be defined for all subsets of the domain, van Rooij proposes that it should only be defined for CCs that meet certain conditions (discussed in his paper). Then he shows that the comparative relations, defined earlier in this section, give rise to semi-orders from which the required strict–weak orders can be derived. With this modification of van Benthem's analysis, the equivalence in Theorem 4.2.15 ceases to hold and we can account for contrasts like those in (7-a) and (7-b).

I will not incorporate van Rooij's (2011a) analysis into the current system, particularly, because it implies a slightly different analysis of the properties of vague language than the one that I adopt here. However, see Burnett (2014b, 2015) for a DelS analysis that captures Kennedy's observations in the adjectival and determiner phrase domain. In other words, I simply highlight in this section that there exists an account of contrasts between the positive and comparative forms in the literature that is consistent with the general approach developed in this monograph.
4.2.2 Classical semantics for absolute/non-scalar adjectives

I suggested in Chapter 3 (Section 3.3) that both AAs and NSs have semantic denotations that are assigned independently of a contextually given CC. That is, in order to know which rooms are empty or which sticks are straight, we don’t need to compare them to a certain group of other individuals; we just need to look at their properties. Similarly, for NSs: to know whether a shape is hexagonal, we do not need to compare it to other shapes, we simply need to count its sides. To incorporate this idea into the Delineation approach, I propose (following a suggestion from van Rooij (2011c)) that, in a semantic framework based on CCs, what it means to be non-context sensitive is to have your denotation be invariant across classes. Thus, for an AA or NS Q and a CC X, it suffices to look at what the extension of Q is in the maximal CC, the domain D, in order to know what \( \llbracket Q \rrbracket_X \) is. I therefore propose that an additional axiom governs the semantic interpretation of the members of the absolute class that does not apply to the relative class: the absolute adjective axiom (AAA).

\[\text{Axiom 4.2.16 Absolute adjective axiom. For all absolute and non-scalar predicates } Q, \text{ all interpretations } \llbracket \cdot \rrbracket_M, \text{ all } X \subseteq D \text{ and } a_i \in X,\]

1. If \( \llbracket Q_1(a_i) \rrbracket_{X,M} = 1 \), then \( \llbracket Q_1(a_i) \rrbracket_{D,M} = 1 \).
2. If \( \llbracket Q_1(a_i) \rrbracket_{D,M} = 1 \), and \( \llbracket Q_1(a_i) \rrbracket_{X,M} \neq i \), then \( \llbracket Q_1(a_i) \rrbracket_{X,M} = 1 \).

In other words, the semantic denotation of an AA or NS is set with respect to the total domain, and then, by the AAA, the interpretation of Q in D is replicated in each smaller CC.\(^8\) As an illustration, consider the absolute predicate \textit{empty} and the CC \{A, B\}. In this example, only container A is truly empty, so when we apply the predicate (as in Figure 4.8), only A is in its semantic denotation.

If we now move to a larger CC like \{A, B, C\}, when we reapply the predicate, the AAA tells us that, despite the fact that the CC has changed,

\(^8\) In this approach, I take an important difference between RAs and AAs to be the constraints that these predicates must obey across CCs. Another option would be to have the two classes of predicates have different kinds of individuals in their CCs. This is done by Toledo and Sassoon (2011), where RAs involve CCs composed of different individuals and AAs involve CCs composed of different intensional counterparts of the same individual. This being said, Toledo and Sassoon’s account is not strictly speaking a Delineation account, since it also adopts a DegS analysis of comparatives. However, it would be interesting in future work to see to what extent having different kinds of CCs could be useful in solving the puzzles associated with AAs within a pure DelS system.
A still has to be in the extension of empty and B still has to be in its anti-extension (cf. Figure 4.9).

Note that if empty were an RA, a possible interpretation of this predicate in \{A, B, C\} would be \{A, B\}. But such an interpretation (in which \(B \not\in \text{empty}_{\{A,B\}}\) and \(B \in \text{empty}_{\{A,B,C\}}\)) is ruled out by the AAA.

Since, in the Delineation framework, scalarity is derived from context sensitivity, the proposal presented in this section already makes some predictions concerning the scales that are associated with AAs and NSs. In particular, the AAA is very powerful. In fact, even without any of van Benthem’s axioms, we can prove that the scales associated with AAs are strict weak orders.

**Theorem 4.2.17**  If \(Q\) is an AA or NS (i.e. the interpretation of \(Q\) satisfies the AAA), \(>_{Q}\) is a strict weak order.\(^9\)

The scales that the semantic denotations of absolute constituents give rise to are very small, essentially trivial. In particular, the relations denoted by the absolute and non-scalar comparative (\(>_{Q}\)) do not allow

\(^9\) Proof. **Irreflexivity.** An individual \(a_1\) cannot be both in \(Q\) and not in \(Q\).

**Transitivity.** Trivially. **Almost connected.** Let \(a_1, a_2, a_3 \in D\) and suppose \(a_1 >_{Q} a_2\). Since, by the AAA, all classical denotations are subsets of \(Q\), we have two cases: 1) if \(a_3 \in Q\), then \(a_3 >_{Q} a_2\), and 2) if \(a_3 \not\in Q\), then \(a_1 >_{Q} a_3\).
for the predicate to distinguish three distinct individuals. This result is stated as Theorem 4.2.18.

**Theorem 4.2.18**  If $Q$ is an absolute or non-scalar predicate (i.e. $Q$'s interpretation obeys the AAA), then there is no model $M$ such that, for distinct $a_1, a_2, a_3 \in D$, $a_1 >_Q a_2 >_Q a_3$.\(^{10}\)

More simply, if we look at the sets of individuals that are equivalent with respect to $Q$ (i.e. related by the $\approx_Q$ relation), we see that AAs and NSs allow for only two equivalence classes. Alternatively, we could say that the degree scales ($\gg_Q$) associated with AAs and NSs have at most two degrees.

**Theorem 4.2.19**  If the interpretation of a predicate $Q$ satisfies the AAA, then there is no model $M$ such that for distinct $a_1, a_2, a_3 \in D$, $[a_1]_{\approx_Q} \not= [a_2]_{\approx_Q} \not= [a_3]_{\approx_Q}$.\(^{11}\)

In other words, if we were to look at the equivalence classes based on $>_Q$, all the elements that are completely empty or perfectly hexagonal are in one equivalence class, and all the elements that are not completely empty or perfectly hexagonal are treated as equivalent. Recall that this is very different from what we see with RAs: these predicates can distinguish between more than two individuals, and, correspondingly, the degree-type scales that are associated with them ($\gg_p$) can have more than two degrees.

### 4.2.3 The paradox of absolute scalar adjectives

Is an analysis that associates trivial scales with AAs and NSs a descriptively adequate one? On the one hand, it would seem so for true non-scalar predicates. For example, if someone tells me (8), my reaction to

\(^{10}\) Proof. Let $Q$ satisfy the AAA. Suppose for a contradiction that there is some model $M = \langle D, [\cdot] \rangle$ such that $a_1, a_2, a_3$ are distinct members of $D$, and $a_1 >_Q a_2 >_Q a_3$. Then, by Definition 4.2.7, there is some $X \subseteq D$ such that $a_1 \in [Q]_X$ and $a_2 \not\in [Q]_X$. Therefore, by the AAA, $a_2 \not\in [Q]_D$. Furthermore, since $a_3 >_Q a_1$, there is some $X' \subseteq CC$ such that $a_2 \in [Q]_{X'}$ and $a_1 \not\in [Q]_{X'}$. Since $a_2 \in [Q]_{X'}$, by the AAA, $a_2 \not\in [Q]_{D'}$. So there is no model $M$ such that, for distinct $a_1, a_2, a_3 \in D$, $a_1 >_Q a_2 >_Q a_3$. \(\Box\)

\(^{11}\) Proof. Let $Q$ satisfy the AAA and suppose for a contradiction that there is some model in which, for $a_1, a_2, a_3 \in D$, $[a_1]_{\approx_Q} \not= [a_2]_{\approx_Q} \not= [a_3]_{\approx_Q}$. Since $[a_1]_{\approx_Q} \not= [a_2]_{\approx_Q}$, $a_1 \not\approx_Q a_2$. Without loss of generality, suppose $a_1 >_p a_2$. So there is some $X \subseteq D$ such that $[Q(a_1)]_X = 1$ and $[Q(a_2)]_X = 0$. Since $Q$ satisfies the AAA, $[Q(a_1)]_D = 1$ and $[Q(a_2)]_D = 0$. Since $[a_2]_{\approx_Q} \not= [a_3]_{\approx_Q}$, $a_2 \not\approx_Q a_3$. Suppose without loss of generality that $a_2 >_Q a_3$. Then, by Definition 4.2.7, there is some $X' \subseteq D$ such that $[Q(a_3)]_{X'} = 1$ and $[Q(a_2)]_{X'} = 0$. So, by the AAA, $[Q(a_2)]_D = 1$. \(\Box\) So there is no model $M$ such that for $a_1, a_2, a_3 \in D$, $[a_1]_{\approx_Q} \not= [a_2]_{\approx_Q} \not= [a_3]_{\approx_Q}$. \(\Box\)
this, after I have recovered from the strangeness of the statement, is to say, “Why yes; yes it certainly is.”

(8) 5 is **more prime** than 6.

But, on the other hand, it is clear that trivial scales are inappropriate for AAs: as discussed in Section 3.2, it is perfectly acceptable and very natural to use absolute comparatives like those in (9).

(9) a. Room A is emptier than room B, which is emptier than room C.
   b. This road is straighter than that road, which is straighter than this third road.
   c. This towel is wetter than that one, which is wetter than this other one.
   d. Ottawa is cleaner than Montréal, which is cleaner than Paris.
   e. The table is flatter than the desk, which is flatter than the sidewalk.

So something needs to be modified or added to our analysis. Note that, in the way that the framework is set up at the moment, it is not clear exactly what can be modified to account for examples like (9).

To analyze certain differences between RAs and AAs, we proposed that AAs had a classical semantic denotation that was not context sensitive. Within the Delineation framework, gradability is derived from CC-based context sensitivity (and constraints on this contextual variation). Thus, as observed by van Rooij (2011c) and McNally (2011), AAs pose the following puzzle for the CC-based approach:

(10) **The paradox of absolute adjectives:**
   a. If gradability is derived from CC-based context sensitivity, **and**
   b. AAs are not context sensitive, **then**
   c. How can they be gradable?

Although I have framed the paradox of the gradability of AAs as a problem for the Delineation approach, the puzzle extends beyond this particular framework and is, in fact, a longstanding problem in the semantics and pragmatics of gradable constituents. For example, the contradictory nature of AAs is summarized by Récanati (2010, 117), who gives an analysis of the predicate *empty* within the DegS framework, in the following way:

(11) As a matter of fact, we know perfectly well which property the adjective *empty* expresses. It is the property (for a container)
of not containing anything, of being devoid of contents. This is how we define empty. Note that this is an absolute property, a property which a container has or does not have. Either it contains something, or it does not contain anything. So the property which the adjective expresses and which determines its extension is not a property that admits of degrees. How, then, can we explain the gradability of the adjective?

And Récanati is not the only one to make this observation. The paradox was also discussed in the 1970s by Peter Unger and David Lewis. In Scorekeeping in a Language Game (Lewis, 1979), Lewis (p. 245) shows, for the predicate flat, how reasoning about its gradability seems to devolve into absurdity:

(12) Peter Unger has argued that hardly anything is flat. Take something that you claim is flat; he will find something else and get you to agree that it is even flatter. You think the pavement is flat—but how can you deny that your desk is flatter? But flat is an absolute term: it is inconsistent to say that something is flatter than something that is flat. Having agreed that your desk is flatter than the pavement, you must concede that the pavement is not flat after all.

Finally, similar observations about the seemingly paradoxical use of comparative morphology with absolute terms go back even to Sapir (1944), who proposes (p. 115) the following analysis of comparatives formed with the AA perfect:

(13) Observe that the ‘less perfect’ of B is really as illogical as ‘more perfect’ would be. It may be considered an ellipsis for the logical ‘less than perfect’ or ‘less nearly perfect’ based on a secondary extension of the range of meaning of the term ‘perfect’. The superlative implication of ‘perfect’, which should make of it a unique and ungradable term, tends to be lost sight of for the simple reason that it belongs to the class of essentially gradable terms (e.g. good). Such terms as ‘less perfect’ are psychological blends of unique terms of the type ‘perfect’ and graded terms of the type ‘less good’. The polar term is stretched a little, as it were, so as to take in at least the uppermost (or nethermost) segment of the gradable gamut of reality.
In the rest of the chapter, I will give a new solution to the paradox of the gradability of AAs within a TCS extension of the Delineation framework. Crucially, I will argue that the solution to this longstanding puzzle lies in the appropriate analysis of the existential context sensitivity property that holds of these predicates that was argued for in Chapter 3, and its relation to the phenomenon of potential vagueness. As such, the form of my solution will bear many similarities to Sapir’s, and it can even be viewed as a more complete and formalized implementation of his intuitions. In particular, in Section 4.3, I will show that, through giving an appropriate tolerant and strict semantics for AAs, we can arrive at an understanding of how it is possible to “stretch” the meaning of an absolute term to (in the words of Sapir) take in the gradable gamut of reality.

4.3 Tolerant/strict semantics

The basic idea behind the TCS extension of the Delineation framework that I propose is to build tolerant and strict structures on top of classical semantic structures in the way done in Cobreros et al. (2012b), and use the properties of this logic to model the properties of vague language that were discussed in Chapter 3. We therefore extend our models to Tolerant models (T-models) by adding the function $\sim$ in the way shown in Definition 4.3.1.

**Definition 4.3.1 T-model.** A T-model is a tuple $M = \langle D, m, \sim \rangle$, where $\langle D, m \rangle$ is a model and $\sim$ is a function from predicate/CC pairs such that:

- For all $P$ and all $X \subseteq D$, $\sim X^P$ is a binary relation on $X$.

There are two ways in which this definition differs from the original definitions in Cobreros et al. (2012b) (presented in Chapter 2): Firstly, note that now, instead of mapping a predicate to an indifference relation as in classical TCS, $\sim$ maps a predicate and a CC to an indifference relation on the members of the class. Thus, indifference relations are also relativized to CC. Secondly, in Chapter 2, we immediately put constraints of reflexivity and symmetry on the $\sim_P$s. In what follows, we will put similar constraints on the definition of $\sim$; however, due to the fact that they now relate elements of CCs, these constraints will be more complicated. I will present each proposed constraint in detail; however, we first define tolerant and strict denotations (relativized to CCs) as in
Definition 4.3.2. As a terminological note: in what follows, as a way of talking about the multiple values that we will assign to predicates and comparatives, I will refer to the (classical) denotations that we assigned to (non-)scalar predicates as their semantic denotations and refer jointly to the secondary tolerant and strict denotations as pragmatic denotations.\footnote{In this framework, tolerant and strict denotations are constructed from classical (i.e. basic semantic) denotations in conjunction with context sensitive indifference (\(\sim\) relations). As I will argue in Section 4.3.1, these indifference relations should be taken to model general cognitive judgments of very close similarity or approximation, which are most likely not exclusive to natural language. Indeed, there is a fair amount of research that suggests that Sorites-type paradoxes can be constructed based not only on linguistic data, but also on perceptual data (see, for example, Raffman (2000) and Égré (2009) for some examples). Thus, I consider it to be in, line with standard terminology to refer to the tolerant and strict denotations as pragmatic objects. Note, however, that these pragmatic denotations and the scales that are derived from them will have grammatical reflexes (cf. Chapter 5). Given this fact, some readers may prefer to refer to all three denotations assigned adjectival predicates as features of their semantics. Therefore, I caution the reader not to read too much into this particular choice of labels.}

Definition 4.3.2 Tolerant/strict CC denotations. For all predicates \(P\) and \(X \subseteq D\),

1. \(\|P\|_t^X = \{x : \exists d \sim^X x : d \in \|P\|_X\}\).
2. \(\|P\|_s^X = \{x : \forall d \sim^X x : d \in \|P\|_X\}\).

Finally, the tolerant/strict semantics for the positive form of an adjective with respect to a CC is given as in Definition 4.3.3.

Definition 4.3.3 Positive form. For all \(T\)-models \(M\), all \(X \subseteq D\), all predicates \(P\), and all \(a_1 \in D\),

1. \(\|P(a_1)\|_{X,M}^t = \begin{cases} 1 & \text{if } \|a_1\|_M \in \|P\|_X^t, M \\ 0 & \text{if } \|a_1\|_M \in X - \|P\|_X^t, M \\ i & \text{otherwise} \end{cases} \)
2. \(\|P(a_1)\|_{X,M}^s = \begin{cases} 1 & \text{if } \|a_1\|_M \in \|P\|_X^s, M \\ 0 & \text{if } \|a_1\|_M \in X - \|P\|_X^s, M \\ i & \text{otherwise} \end{cases} \)

The tolerant and strict interpretations of negative sentences in DelTCS are given in Definition 4.3.4.
Definition 4.3.4  Tolerant/strict semantics for negation.

1. For all models $M$, $X \subseteq D$ and wffs $\phi$,

$$\llbracket \neg \phi \rrbracket^t_{X,M} = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket^s_{X,M} = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket^s_{X,M} = 1 \\ i & \text{otherwise} \end{cases}$$

2. For all models $M$, $X \subseteq D$ and wffs $\phi$,

$$\llbracket \neg \phi \rrbracket^s_{X,M} = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket^s_{X,M} = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket^s_{X,M} = 1 \\ i & \text{otherwise} \end{cases}$$

Finally, we can define the tolerant and strict interpretation of the comparative as follows:

Definition 4.3.5  Comparative form.  For all models $M$, $X \subseteq D$, predicates $P$, and $a_1, a_2 \in D$,

1. $\llbracket a_1 >_P a_2 \rrbracket^t_X = 1$ iff there is some $X' \subseteq D$ such that $\llbracket P(a_1) \rrbracket^t_{X',M} = 1$ and $\llbracket P(a_2) \rrbracket^t_{X',M} = 0$.
2. $\llbracket a_1 >_P a_2 \rrbracket^s_X = 1$ iff there is some $X' \subseteq D$ such that $\llbracket P(a_1) \rrbracket^s_{X',M} = 1$ and $\llbracket P(a_2) \rrbracket^s_{X',M} = 0$.

The analysis in Definitions 4.3.3, 4.3.4, and 4.3.5 is a simple implementation of the Klein-ian approach to the semantics of scalar adjectives within a TCS account of the semantics/pragmatics of vague predicates. Thus, in a particular situation with a particular CC $X$, a predicate $P$ of either semantic class can have borderline cases (objects that in both $\llbracket P \rrbracket^t_X$ and $\llbracket \neg P \rrbracket^t_X$), and members of $X$ can be related by $\sim_P$ in a way that forms a Soritical series; therefore, we can construct a Sorites paradox.

4.3.1 The properties of indifference

As it stands, we have not placed any constraints on the definition of $\sim$. However, like the case discussed in Section 4.2 with the interpretation of relative predicates, if we do not say anything about how indifference relations can be established across CCs, the $\sim_P$s will not look at all
like the cognitive indifference relations that they are supposed to be modelling. In what follows, I will propose a series of constraints that the \( \sim \) function must satisfy across CCs. These constraints are meant to be, at the same time, intuitive in nature and inspired by previous proposals in the linguistics and psychological literature about the properties of indifference and the more general notion of similarity. This being said, it is important to note that the precise question of indifference or even similarity within and across adjectival CCs is not something that, to my knowledge, is explicitly examined in the psychological literature. Therefore, the proposals outlined in this section are not meant to constitute a comprehensive psychological analysis of this phenomenon. Nevertheless, I believe that the constraints proposed here will be sufficient to encode a useful (albeit simplistic) notion of approximation/close similarity into the logical system developed in this work, one that will be appropriate for modelling the context sensitive uses of potentially vague (P-vague) predicates.

The first property that is generally proposed to characterize similarity/approximation relations is reflexivity (cf. Luce (1956), Pogonowski (1981), Cobreros et al. (2012b), among many others). Intuitively, every individual is indifferent from itself. Thus, we adopt the constraint in (14) that enforces reflexivity across CCs.

\[
\text{(14) Reflexivity (R): For all predicates } P, \text{ all models } M, \text{ all } X \subseteq D, \text{ for all } a_1 \in X, a_1 \sim_{P_X} a_1. 
\]

In addition to being reflexive, indifference and similarity relations are generally proposed to be symmetric (e.g. the original formulation of TCS in Cobreros et al. (2012b)). At first glance, this seems reasonable: if an individual \( a \) is considered indifferent from an individual \( b \), then surely \( b \) must also be considered indifferent from \( a \). However, there is a fair amount of literature in both philosophy and psychology that argues that, in certain cases, judgments of similarity are directional (e.g. Tversky, 1977; Tversky and Gati, 1978; Rosch, 1978; Ortony et al., 1985; Lakoff, 1987; and Égré and Bonnay, 2010). The cases for which it has been proposed that symmetry fails in judgments of similarity and indifference particularly involve relations between individuals that differ in terms of “prototypicality” (cf. Tversky, 1977; Rosch, 1978; Ortony et al., 1985; Lakoff, 1987). The generalization concerning asymmetric judgments of similarity can be stated (in the words of (Ortony et al. 1985, p. 570)) as follows:
(15) **Prototypicality generalization:**
Atypical members of categories tend to be judged as more similar to typical members than the other way around.

Example (15) is a robust generalization that has been observed in studies of judgments of similarity with respect to colours, geographical concepts, letters, sounds, and shapes (cf. Tversky (1977) and Lakoff (1987) for literature reviews). For example, Tversky (1977) shows that, when asked to judge similarity between pairs of countries, participants overwhelmingly judge the less prominent country to be more similar to the more prominent country than vice versa. More specifically, out of 69 participants, 66 preferred the sentence (16-a) over (16-b), and similar results were obtained for pairs of sentences like (17) and (18).

(16) a. North Korea is similar to Red China.
   b. Red China is similar to North Korea.

(17) a. Mexico is similar to the USA.
   b. The USA is similar to Mexico.

(18) a. Luxembourg is similar to Belgium.
   b. Belgium is similar to Luxembourg.

This discussion of asymmetric similarity judgments is important for the present purposes because, in the previous chapter (Section 3.4), I argued that we saw a similar asymmetry in judgments of indifference with AAs. In particular, we saw that, with AA$^T$'s like *empty*, members of the adjective's semantic denotation (i.e. those individuals that always count as *empty*) are never indifferent to members outside the semantic denotation. However, we also saw that, depending on context, individuals that are not completely empty can be considered indifferent from the completely empty ones. Thus, we have a similar case to examples like (16-a) and (16-b) in (19).

(19) a. This container with no liquid in it $\sim_{\text{empty}}$ this container with a small amount of liquid in it.
   b. This container with a small amount of liquid in it $\not\sim_{\text{empty}}$ this container with no liquid in it.

We also saw that AA$^P$'s display the opposite pattern:

(20) a. This towel with no water on it $\not\sim_{\text{wet}}$ this towel with a small amount of water on it.
b. This towel with a small amount of water on it $\sim_{\text{wet}}$ this towel with no water on it.

Therefore, I believe that it is a reasonable hypothesis that the patterns with AAs discussed in Chapter 3 are instances of a more general phenomenon in which prototypical members of a predicate’s denotation have a different status than less prototypical members.\textsuperscript{13} Formally, I propose that these asymmetries are encoded into the indifference relations associated with $\text{AA}^T$s and $\text{AA}^P$s by means of the following two pragmatic axioms.\textsuperscript{14} Since the $\sim_P$ relation is now not necessarily symmetric, $a \sim b$ can now be read as $\langle a, b \rangle \in \sim_X \text{Q}_t$, ‘b can count as a’ or ‘b approximates a.’ I highlight here that the $\sim_P$ relations are not meant to be picking out metaphysical relations like ‘± one drop of water’ (which, of course, are symmetric), but rather epistemic relations that express approximation with respect to categorization using $P$. This being said, to be consistent with the terminology used in the TCS framework, I will still sometimes refer to the $\sim$ relations as indifference relations, even though they are not symmetric.

(21) Possible (non-)symmetry in $\sim$:

1. **Symmetry (S):** For a relative predicate $P_1$, a model $M$, and $a_1, a_2 \in D$, if $a_1 \sim_{X_1}^P a_2$, then $a_2 \sim_{X_1}^P a_1$.

2. **Total axiom (TA):** For a total predicate $Q_1$, a model $M$, and $a_1, a_2 \in D$, if $\llbracket Q_1(a_1) \rrbracket_{M,D} = 1$ and $\llbracket Q_1(a_2) \rrbracket_{M,D} = 0$, then $a_2 \not\sim_{X_1}^Q a_1$, for all $X \subseteq D$.

3. **Partial axiom (PA):** For a partial predicate $R_1$, a model $M$ and $a_1, a_2 \in D$, if $\llbracket R_1(a_1) \rrbracket_{M,D} = 1$ and $\llbracket R_1(a_2) \rrbracket_{M,D} = 0$, then $a_1 \not\sim_{X_1}^R a_2$, for all $X \subseteq D$.

Furthermore, as discussed in Chapter 3, NSs have both precise positive and negative forms. To account for this, I propose that indifference relations associated with NSs are subject to a pragmatic constraint that prohibit indifference relations from being established across the

\textsuperscript{13} The view of the difference between RAs and AAs that I am suggesting here is similar in spirit (although very different in its execution) as a proposal by McNally (2011) based on Hahn and Chater (1998) in which the semantic denotations of RAs are determined based on context sensitive similarity relations and the denotations of AAs are determined based on lexical rules.

\textsuperscript{14} A similar strategy was adopted in the analysis of the total/partial pair clear/unclear by Égré and Bonnay (2010), but these authors do not consider any other adjectival predicates or the relation between non-symmetric indifference relations and the total/partial distinction.
boundaries of their semantic denotations. This constraint, which I call
*Be precise*, is the conjunction of the TAs and PAs, and it is stated as
in (22).

(22) **Be precise (BP):**
For a non-scalar predicate \(S\), a model \(M\), and \(a_1, a_2 \in D\),
a. If \(\|S(a_1)\|_{M,D} = 1\) and \(\|S(a_2)\|_{M,D} = 0\), then \(a_2 \not\sim^X S a_1\), for all
\(X \subseteq D\).
b. If \(\|S(a_1)\|_{M,D} = 1\) and \(\|S(a_2)\|_{M,D} = 0\), then \(a_1 \not\sim^X S a_2\), for all
\(X \subseteq D\).

The previous axioms (in Section 4.3.1) made a distinction between the
four subclasses of gradable predicates that were identified; however, the
rest of the pragmatic axioms that I will propose will apply to all adjectival
predicates in the same way. The first general axiom that I propose is
called *tolerant convexity*:\(^{15}\)

(23) **Tolerant Convexity (TC):** For all predicates \(P\), all models \(M\), all
\(X \subseteq D\), and all \(a_1, a_2 \in X\),
• If \(a_1 \sim^X P a_2\) and there is some \(a_3 \in X\) such that \(a_1 \geq^t_P a_3 \geq^t_P a_2\),
then \(a_1 \sim^X P a_3\).

TC says that, if person A is indistinguishable from person B, and there's a
person C lying in between persons A and B on the relevant tolerant scale,
then A and C (the greater two of \(\{A, B, C\}\)) are also indistinguishable.
For example, suppose I have two containers: one that has absolutely
no liquid in it (i.e. is in the semantic denotation of *empty*), one with
a very small amount of liquid, and then one which is a \(1/3\)-full of
liquid. Although it might be conceivable that in some (extremely) large
CC, I may consider the \(1/3\)-full container to be indifferent from
the completely empty container, I will never be able to do so while main-
taining a distinction between the completely empty and almost empty
container.

\(^{15}\) The definition of \(\geq_P\) that is featured in the next two constraints is as follows:

**Definition 4.3.6 Greater than or equal. (≥)** For a model \(M\), a predicate \(P\), \(a_1, a_2 \in D\):
1. \(a_1 \geq_P p a_2 \iff \|a_1 > p a_2\|_{X,M} = 1 \text{ or } a_1 \equiv p a_2\).
2. \(a_1 \geq^t_P p a_2 \iff \|a_1 >^t_P a_2\|_{X,M} = 1 \text{ or } a_1 \equiv^t_P a_2\).
3. \(a_1 \geq^s_P p a_2 \iff \|a_1 >^s_P a_2\|_{X,M} = 1 \text{ or } a_1 \equiv^s_P a_2\).
As a short illustration of how this works, consider the following example: if I compare container B, which has a small amount of liquid, with the completely empty container A (Figure 4.10), the two would not be considered indifferent with respect to emptiness (B has some liquid).

However, if I add container C (which contains very much liquid) into the CC (Figure 4.11), then perhaps A and B start looking much more similar, when it comes to emptiness.

However, given TC, it would never be the case that, in some CC, A and C could be considered to be indifferent with respect to emptiness, but not B (Figure 4.12). Thus, I propose that it should be possible to order individuals with respect to how close to being completely empty they are by looking at in which CCs they are considered indifferent to completely empty objects.

As shown in the appendix, TC performs a very similar function to van Benthem’s NR. I propose a second axiom that is, in some sense, the dual of TC: strict convexity:

\[ (24) \text{ Strict convexity: For all predicates } P_1, \text{ all models } M, \text{ all } X \subseteq D, \text{ and all } a_1, a_2 \in X, \]

- If \( a_1 \sim_{X}^{P_1} a_2 \) and there is some \( a_3 \in X \) such that \( a_1 \not\geq_{P_1}^{X} a_3 \not\geq_{P_1}^{X} a_2 \),

  then \( a_3 \sim_{X}^{P_1} a_2 \).
Strict convexity says that, if person A is indistinguishable from person B, and there’s a person C lying in between persons A and B on the relevant strict scale, then B and C (the lesser two of \{A, B, C\}) are also indistinguishable. For example, strict convexity rules out situations such as in Figure 4.13 in which a very wet towel (towel 3) approximates a bone-dry towel (towel 1), but a slightly wet towel (towel 2) does not also approximate towel 1.

The next axiom deals with how indifference relations can change across CCs. At the moment, \(\sim\)'s can be established and destroyed in different CCs in a more or less arbitrary way, provided that the previous three constraints are respected. But presumably we might want some more restrictions on the distribution of the \(\sim\)'s. I therefore propose the following, which specifies how distinctions can be made between individuals at different sizes of CCs:

\begin{align*}
\text{(25) Granularity (G): } & \text{For all predicates } P_i, \text{ all models } M, \text{ all } X \subseteq D, \\
& \text{and all } a_1, a_2 \in X, \text{ if } a_1 \preceq^{X}_{P_i} a_2, \text{ then for all } X' \subseteq D : X \subseteq X', \\
& \text{if } a_1 \preceq^{X'}_{P_i} a_2.
\end{align*}

Granularity says that if person A and person B are indistinguishable in CC X, then they are indistinguishable in all supersets of X. This is
meant to reflect the fact that the larger the domain (i.e. the larger the CC), the more things can cluster together. In other words, the larger the CC, the more it is possible to collapse fine distinctions that were made in smaller CCs, and once you collapse such a “fine-grained” distinction, you cannot make it again at a more “coarse-grained” level.

These first three axioms make very similar proposals to some other existing analyses of what I have been calling existential context sensitivity in the literature. For example, in response to a passage in Unger (1975) who evokes the paradox of AAs outlined in Section 4.2.3 (i.e. how can an object be both straighter/flatter than something and still not be straight/flat?), Lewis says (p. 353):

(26) The right response to Unger, I suggest, is that he is changing the score on you. When he says that the desk is flatter than the pavement, what he says is acceptable only under raised standards of precision. Under the original standards the bumps on the pavement were too small to be relevant either to the question whether the pavement is flat or to the question whether the pavement is flatter than the desk. Since what he says requires raised standards, the standards accommodatingly rise. Then it is no longer true enough that the pavement is flat. That does not alter the fact that it was true enough in its original context [Lewis’ italics-HB].

Thus, according to Lewis, existential context sensitivity should be analyzed as changing standards of precision. This idea is taken up more formally in van Rooij (2011c), who sets his account within a Delineation approach, and Hobbes (1985) also makes proposals in this spirit.

While the previous three axioms talk about how indifference is preserved, the final two axioms deal with the preservation of differences across CCs.

(27) For all predicates $P_i$, models $M$, all $X \subseteq D$,

a. **Contrast preservation (CP):** For all $X' \subseteq D$, and $a_1, a_2 \in X$, if $X \subset X'$ and $a_1 \not\sim^{P_i} a_2$ and $a_1 \sim^{X'}_{P_i} a_2$, then $\exists a_3 \in X' - X : a_1 \not\sim^{P_i}_{X'} a_3$.

---

16 For a full comparison of van Rooij’s DelS analysis of AAs and the one developed here, see Burnett (2014a).
b. **Minimal difference (MD):** For all \( a_1, a_2 \in D \), if \( a_1 \succ_P a_2 \) \( M, g, X = 1 \), then \( a_1 \not\sim_{P(X)} a_2 \).

**MD** says that, if, at the finest level of granularity, you would make a distinction between two individuals with respect to the semantic denotation of a predicate, then they are not indistinguishable at that level of granularity. **MD** is similar in spirit to van Benthem’s **DD** because it allows us to preserve contrasts down to the smallest CCs. **CP** says that, if person A and person B are distinguishable in one CC, \( X \), and then there’s another CC, \( X' \), in which they are indistinguishable, then there is some person C in \( X' - X \) that is distinguishable from person A. For example, suppose I have two containers that, when we restrict our attention to them, we make a distinction between them in terms of *emptiness* (perhaps one is completely empty and one is almost empty). Then, suppose that in a larger CC, these two containers are now treated as indifferent. According to **CP**, this can only occur because of the introduction of a new container into the CC (perhaps a container with a very large amount of liquid) which is viewed as distinct from the other containers. This axiom is similar in spirit to van Benthem’s **UD** in that it ensures that, if there is a contrast/distinction in one CC, the existence of a contrast is maintained in all the larger CCs.

In summary, my analysis of the pragmatic constraints associated with scalar and non-scalar adjectives is given in Table 4.1.

In Section 4.4, I discuss the predictions that this analysis makes for the context sensitivity patterns and potential vagueness patterns exhibited by RAs, AAs, and NSs.

**Table 4.1** Pragmatic axioms for (non-)scalar adjectives

<table>
<thead>
<tr>
<th>Axiom</th>
<th>RA</th>
<th>AA(^T)</th>
<th>AA(^P)</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflexivity (R)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tolerant convexity (TC)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Strict convexity (SC)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Granularity (G)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Minimal difference (MD)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Contrast preservation (CP)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Symmetry (S)</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Total axiom (TA)</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Partial axiom (PA)</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Be precise (BP)</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>
4.4 Predictions of the DelTCS analysis

In Chapter 3, I argued RAs, AAs, and NSs could be empirically distinguished through looking at both their context sensitivity and potential vagueness properties. Namely, I argued in favor of the patterns shown in Table 4.2.

In this section, I show how the analysis outlined in Section 4.3 set within DelTCS derives these patterns, as well as certain other empirical properties associated with the different classes of adjectives.

4.4.1 Context sensitivity results

We saw in Chapter 3 that the context sensitivity of AAs is more limited than that of RAs, and I argued that NSs were not context sensitive (at least on their precise use). In particular, we diagnosed the existence of context sensitivity differences through the use of the Definite Description Test (i.e. Pass me the tall/empty one); however, I suggested that the results of this test correspond to a more general distinction between universal and (properly) existential context sensitivity: RAs could potentially change their criteria of application in all kinds of CCs; however, despite still displaying some contextual variation, some instances of possible variation were ruled out. Concretely speaking, the situation that the Definite Description Test probes is the following: there is some large-ish CC $X$ in which there are two elements $a_1$ and $a_2$, and neither of them (even tolerantly) satisfies a predicate $P$. However, there is a smaller subset of $X$, $X'$, in which $a_1$ (at least tolerantly) satisfies $P$. As we saw in Chapter 3, RAs allow this situation, and models in

<table>
<thead>
<tr>
<th>Pattern</th>
<th>RA</th>
<th>$AA^T$</th>
<th>$AA^P$</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal context sensitivity</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Existential context sensitivity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Potential vagueness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-vague $\neg P$</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>P-vague $P$</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>
which the classical/tolerant denotations of predicates vary in this way are acceptable models for RAs.\(^\text{17}\)

However, such a pattern of contextual variation is not permitted with AA\(^T\)s, and we can show that, in the proposed framework, there are no models in which the tolerant denotations of predicates of the AA\(^T\) class change in this way. This is stated as Theorem 4.4.1. The analysis presented above therefore predicts that we should be able to say “Pass me the tall one” to distinguish between objects that are not even tolerantly tall in a large CC, despite being tall in a smaller one; however, the only time in which we can say “Pass me the empty one” is when one object is tolerantly empty in a small CC and all of its supersets.

**Theorem 4.4.1** If Q is an AA\(^T\), there are no models M such that there is some CC X \(\subseteq D\) and \(a_1, a_2 \in X\) such that \(a_1 \notin \llbracket Q \rrbracket^X\) and \(a_2 \notin \llbracket Q \rrbracket^X\) in M, and, for another CC \(X' \subset X\), \(a_1 \in \llbracket Q \rrbracket^X\) and \(a_2 \notin \llbracket Q \rrbracket^X\) in M.\(^\text{18}\)

Clearly, this property also holds dually for AA\(^P\)s and their context sensitive strict denotations: there are no models in which two objects are strictly \(Q\) in a large CC, yet one is not even tolerantly \(Q\) in a smaller CC.

Although AA\(^T\)s have context sensitive tolerant denotations and AA\(^P\)s have context sensitive strict denotations (which explains their existential context sensitivity), it turns out that the constraints on \(\sim\) that we proposed above (Section 4.3.1) have consequences for the properties of the strict denotations of total predicates and the tolerant denotations of partial predicates. In particular, the partial and total axioms have

---

\(^{17}\) For example, consider the model \(M=\langle\{a, b, c\}, \llbracket \cdot \rrbracket, \sim\rangle\), where \(\llbracket \cdot \rrbracket\) is defined (restricted to \(P\)) as follows:

1. \(\llbracket P \rrbracket_1 = \{\}; \llbracket P \rrbracket_{\{a\}} = \{a\}; \llbracket P \rrbracket_{\{b\}} = \{b\}; \llbracket P \rrbracket_{\{c\}} = \{c\}.\)
2. \(\llbracket P \rrbracket_{\{a,b\}} = \{a\}; \llbracket P \rrbracket_{\{b,c\}} = \{b\}; \llbracket P \rrbracket_{\{a,b,c\}} = \{a\}.\)

And suppose that the only indifference relations are the reflexive relations associated with the pertinent members of the CCs. Clearly, \(\llbracket \cdot \rrbracket\) satisfies NR (there are no “switching” pairs like in Figure 4.5). UD (the existence of a contrast is preserved across supersets), and DD (the existence of a contrast is preserved across subsets), so it is a possible model, and there is some individual, namely \(b\), such that \(b \in \llbracket P \rrbracket_{\{b,c\}}\) (and therefore \(b \in \llbracket P \rrbracket_{\{b,c\}}\) and \(b \notin \llbracket P \rrbracket_{\{a,b,c\}}\) (and \(b \notin \llbracket P \rrbracket_{\{a,b,c\}}\)).

\(^{18}\) Proof. Suppose for a contradiction that \(a_1 \notin \llbracket Q \rrbracket^X\) and \(a_2 \notin \llbracket Q \rrbracket^X\) in M, and, for another comparison class \(X' \subset X\), \(a_1 \in \llbracket Q \rrbracket^X\) and \(a_2 \notin \llbracket Q \rrbracket^X\). Since \(a_1 \in \llbracket Q \rrbracket^X\), there is some \(a_3 \in X'\) such that \(a_3 \sim_X a_1\) and \(a_3 \notin \llbracket Q \rrbracket_X.\) Since \(X' \subset X\), \(a_3 \in X\) and, by the AAA, \(a_3 \in \llbracket Q \rrbracket_X.\) Furthermore, by Granularity, \(a_3 \sim_X a_1.\) So \(a_1 \in \llbracket Q \rrbracket^X.\)\(\blacksquare\)
the effect of ensuring that AA\(^T\)'s have identical classical and strict denotations across CCs, while AA\(^P\)'s have identical classical and tolerant denotations across CCs. This is shown in Theorem 4.4.2.

**Theorem 4.4.2 Tolerant/classical and strict/classical identity.** Let \(Q^T\) be a total predicate and let \(Q^P\) be a partial predicate. Then,\(^{19}\)

1. For all \(X \subseteq D\), \(\llbracket Q^T \rrbracket_X^t = \llbracket Q^T \rrbracket_X\).
2. For all \(X \subseteq D\), \(\llbracket Q^P \rrbracket_X^t = \llbracket Q^P \rrbracket_X\).

The consequence is therefore that AAs have only a single denotation that shows non-trivial existential context sensitivity: tolerant for AA\(^T\)'s and strict for AA\(^P\)'s. This fact will have important implications for the non-trivial scales associated with these predicates, which will be discussed in Section 4.4.2.

The results presented in this subsection also give us an immediate explanation for another empirical pattern discussed in Chapter 3: the distribution of *for* phrases. Recall that *for* phrases were very natural with RAs; however, they were only natural with AAs on their “imprecise” or tolerant use; that is, (28-b) is only appropriate if the restaurant in question is not completely empty. Furthermore, we can observe that *for* phrases are completely unnatural with (precise uses of) non-scalars (28-c).

(28) a. John is tall *for* a jockey.  
    (Always appropriate)  
    b. This restaurant is empty *for* a Friday.  
    (Appropriate on the tolerant use)  
    c. ??This number is prime *for* a natural number.  
    (Never appropriate)

This distributional pattern can be easily understood if we make the common assumption that *for* phrases (at least partially) specify the content of the CC parameter of the evaluation of an adjective. Furthermore, we might suppose that there exists a pragmatic constraint on their use

\(^{19}\) 1. By Cobreros et al.'s (2012b) Lemma 1. \(\llbracket Q^T \rrbracket_X^t \subseteq \llbracket Q^T \rrbracket_X\). Show \(\llbracket Q^T \rrbracket_X^t \subseteq \llbracket Q^T \rrbracket_X\). Let \(a_1 \in \llbracket Q^T \rrbracket_X\) and suppose for a contradiction that \(a_1 \notin \llbracket Q^T \rrbracket_X^t\). Then there is some \(a_2 \sim_{Q^T} a_1\) such that \(a_2 \notin \llbracket Q \rrbracket_X\). But, by the total axiom, since \(a_2 \notin \llbracket Q \rrbracket_X\), \(a_2 \not\sim_{Q^T} a_1\). \(\bot\)

2. By Cobreros et al.'s (2012b) Lemma 1. \(\llbracket Q^T \rrbracket_X^t \subseteq \llbracket Q^T \rrbracket_X\). Show \(\llbracket Q^T \rrbracket_X^t \subseteq \llbracket Q^T \rrbracket_X\). Let \(a_1 \in \llbracket Q^T \rrbracket_X^t\) and suppose for a contradiction that \(a_1 \notin \llbracket Q^T \rrbracket_X\). Then there is some \(a_2 \sim_{Q^T} a_1\) such that \(a_2 \notin \llbracket Q \rrbracket_X\). But, by the PA, since \(a_2 \notin \llbracket Q \rrbracket_X\), \(a_2 \not\sim_{Q^T} a_1\). \(\bot\)
limiting their distribution to situations in which specifying content of the CC parameter is informative, i.e. in which the CC plays some non-trivial role in the interpretation of the adjective. In my analysis, all the denotations of RAs are free to vary in an (almost) unconstrained manner depending on CC, so the use of for phrases with these predicates would always be informative and so the use of a for phrase would be predicted to always be natural with these adjectives. However, the semantic denotations of AAs are CC-independent: they are fixed across contexts by the AAA. Therefore, we would expect for phrases to be strange in contexts targeting the semantic denotations of these predicates, i.e. contexts that favor precision. However, in my proposal, CCs do make a non-trivial contribution to the meaning of AAs on their tolerant (or strict) interpretations through the context-dependence of the ∼Q relations. So we (correctly) predict that sentences with AAs and for phrases improve in contexts where the AA is being used imprecisely. Finally, all three of the interpretations of NSs are context-independent, so we (correctly) expect that it should always be strange to use a for phrase with these adjectives. This being said, the precise analysis of the semantics and pragmatics of for phrases is still an open problem in the field and goes beyond the scope of this paper (cf. recent works by Bylinina (2011), Bale (2011), and Frazee et al. (2013) for recent proposals).20

Finally, given that I proposed that non-scalar predicates are subject to the very strong axiom BP, which is the conjunction of the TAs and PAs, we correctly predict that both the strict and tolerant denotations

20 An interesting test-case for an analysis of for phrases in this vein involves occurrences of a predicate like full in a context in which a wine glass has been filled up to a particular socially dictated standard line which is lower than the full capacity of the glass (or, alternatively, an espresso cup that has been filled a quarter with coffee). As observed by McNally (2011), for phrases are natural in examples like (ii).

(ii) This glass is full for a wine glass.

One possible analysis of the acceptability of the for phrase in (ii) is that full is being interpreted tolerantly in this example, which seems reasonable since the glass is not, technically, completely full. A possible objection to this view is that it is possible to say something like (iii), in which case, since full is a AA, we would expect completely to be picking out its semantic/classical extension. Nevertheless, the for phrase is still possible.

(iii) This glass is completely full for a wine glass.

Another possibility for (ii) is that the semantic denotations of certain AAs like full are actually subject to constraints that, while being stronger than those obeyed by RAs, are actually a bit weaker than incredibly powerful AAA. Thus, there might be some limited CC-based variation and, in these limited cases, we should expect for phrases. However, I leave choosing between these two analyses to future research.
of these adjectives are always identical with their classical denotations. This is stated as Corollary 4.4.3, which is a direct corollary of Theorem 4.4.2. Since the classical denotations of NSs also satisfy the AAA, we (correctly) predict that such predicates should be neither universally nor existentially context sensitive.

**Corollary 4.4.3**  If S is a non-scalar predicate (that is, if S satisfies the AAA and the relevant axioms in Table 4.1), then for all X ⊆ D, [S]_X = [S]_X.

I therefore conclude that the DelTCS correctly captures the context sensitivity patterns found in the adjectival domain that were argued for in Chapter 3.

### 4.4.2 Gradability results

The analysis given in Table 4.1 also provides us an analysis of the observed gradability patterns discussed in Chapter 3. Consider first the analysis of AA^P_s and AA^T_s. The constraints placed on the definition of the ~ function with AAs have a similar effect to van Benthem’s “coherence” constraints with RAs. In particular, these constraints allow us to extract non-trivial strict weak order relations from the existential context sensitivity of the pragmatic denotations of AAs. In other words, the proposals that I made about how indifference relations can be established and change across CCs give us a pragmatic solution to the paradox of AAs.

The idea is conceptually similar in some sense (although extremely different in its execution and its implications for the structure of the lexicon) to a suggestion made by Récanati (2010), with respect to how an adjective like empty can be both absolute and gradable. He proposes that there are two homophonous property-denoting predicates: empty^1, which is gradable, and empty^2, which is absolute. He says (pp. 118–19),

(29) So the property which admits of degrees, and which the measure function measures, is not the basic property of emptiness which the adjective ‘empty’ primarily expresses, but a distinct property that can be defined in terms of it: the property of (as I said) approximating emptiness... If this is right then there are two properties associated with an adjective such as ‘empty’. There is the basic property of emptiness, corresponding to the primary sense (empty^2). It is absolute and does not admit of degrees. In terms of that property, however, we can define another predicate and
generate a scale corresponding to the degrees to which that other predicate applies.

Récanati suggests constructing the scale based on Lasersohn’s (1999) Pragmatic Halo framework that was outlined in Chapter 2. Recall that, in the halo system, objects of the same logical type as a constituent are partially, not strictly, ordered with respect to that constituent’s precise semantic meaning. Therefore, some crucial other proposal must be made to show how to generate the necessary strict weak orders and linear ‘degree’ orders from underlying partial orders. More importantly, Récanati proposes that the relation between empty₁ and empty₂ is that of homophony, meaning that, in the lexicon, for every AA, there are two versions, a non-scalar version (the absolute version) and a scalar version (presumably an RA).²¹ In addition to being rather inelegant, this analysis would seem to make wrong predictions with respect to the distribution of the RA empty₁ in a variety of syntactic constructions. For example, if there was a homophonous relative empty, we would expect to be able to shift its standard in the definite description construction to distinguish between two moderately full containers. Thus, the sentence in (30) could have the reading in (30-a).

(30) Pass me the empty one.
   a. Pass me the empty₁ one. (Less full)
   b. Pass me the empty₂ one. (Completely empty)

But, clearly, it is not possible to use (30) in this way.²² I therefore conclude that a homophony analysis of the gradability of AAs is empirically insufficient. However, the analysis that I presented in this chapter does not rely on homophony to explain the gradability of AAs. The relation between what Récanati calls empty₁ (gradable empty) and empty₂ (nongradable empty) is simply the difference between the pragmatic and semantic denotations of a single lexical item, empty.

The particular constraints that we imposed on the ∼Q relations in Section 4.4.1 immediately give us important results concerning the

²¹ Of course, this is not necessary. Récanati could adopt a more articulated analysis of scale structure than he does (along the lines of Kennedy and McNally (2005) or Kennedy (2007) perhaps) and propose that empty₂ is an AA; however, if he did so, it would be unclear why we would need two emptys in the first place.

²² Note that Récanati’s (2010) main proposal, which is that the gradability of AAs can be analyzed as the result of a pragmatic modulation process (see also Récanati (2004)), is unaffected by this conclusion, if we view the construction of tolerant adjectival meanings as a type of modulation.
pragmatic scales associated with AAs. In particular, the scales that are associated with their tolerant and strict denotations are strict weak orders. These facts are shown in Theorems 4.6.3 and 4.6.4, and these results are proved in Section 4.6 (appendix).

**Theorem 4.4.4**  If \( Q \) is a \( \text{AA}^T \) (subject to the AAA and the relevant constraints in Table 4.1), \( \prec_Q \) is a strict weak order.

**Theorem 4.4.5**  If \( R \) is a \( \text{AA}^P \) (subject to the AAA and the relevant constraints in Table 4.1), \( \prec_R \) is a strict weak order.

Thus we now have a solution to the paradox of AAs; that is, we now have an explanation for how an adjective like *empty* can be, at the same time, gradable and not (universally) context sensitive. The solution is stated as in (31).

(31)  **Solution to the paradox of absolute adjectives:**

Although AAs have neither context sensitive nor gradable semantic denotations, they have both context sensitive and gradable pragmatic denotations.

Note that this solution relies on the adoption of a more interactive view between semantic meaning and pragmatic meaning than is sometimes assumed in the literature. On the one hand, tolerant/strict denotations could be considered pragmatic, since they are constructed making reference to indifference relations, which are aspects of the extra-linguistic context. On the other hand, tolerant/strict denotations could be considered semantic, since they are the objects from which the non-trivial scales associated with AAs are constructed and which license the use of the comparative and other degree morphology. In other words, the theory developed here cross-cuts the traditional *semantics/literal meaning vs pragmatics/meaning-in-context* distinction, and this approach is situated firmly in the class of what we might call *radical pragmatic* theories. This aspect of the theory is discussed at greater length in Chapter 6.

Finally, in the proposal presented here, we assign two pragmatic denotations to each adjectival predicate: a tolerant and a strict one. At first glance, we might therefore think that all AAs are associated with (at least) two non-trivial pragmatic scales. However, the possibility that

---

23 See Récanati (2004, 2010) for a general introduction to and argumentation for this kind of framework.
a particular absolute predicate may be associated with both a tolerant and a strict scale is actually ruled out by the TAs and PAs proposed in (21). In particular, we can prove that if $Q$ is an $AA^T$, then it is necessarily associated with a trivial strict scale, and if $R$ is a $AA^P$, then it is necessarily associated with a trivial tolerant scale. Thus, the analysis accounts for the fact that scalar modifiers target the tolerant scale with adjectives dry and the strict scale with adjectives like wet: the tolerant scale and the strict scale are the only articulated orderings assigned to total and partial predicates respectively.

The axioms in Table 4.1 generate possibly non-trivial tolerant strict weak orders out of trivial semantic orders with predicates of the absolute class. So we might think that by adopting them with the RA class, we would arrive at tolerant relative scales with the same properties as their tolerant absolute counterparts. However, this would be naive. In fact, with the analysis that I proposed, we arrive at $>^T_P$ relations that are not even transitive. This fact is stated below and proved in Burnett (2012a).

As shown in this work, the transitivity of $>^T_P$ for RAs goes through in domains of less than six individuals; however, it is possible for three individuals $a_1, a_2, a_3$ to all be equivalent with respect to $>^T_P$, but the variation in the semantic denotation of $P$ across CCs containing only two of the three individuals and the indifference relations in those CCs can be such that $a_1 >^T_P a_2$ and $a_2 >^T_P a_3$, but $a_1 \approx^T_P a_3$.

**Fact 4.4.6:** For $P \in RA$, $>^T_P$ is not necessarily transitive.

This proposal has the welcome consequence that the only non-trivial strict weak orders that are associated with relative predicates are those derived from their semantic denotations. In other words, my analysis predicts that RAs are uniquely associated with scales derived from their semantic denotations, $AA^T$’s are uniquely associated with scales derived from their tolerant denotations, and $AA^P$’s are uniquely associated with scales derived from their strict denotations.

Finally, since the pragmatic denotations of NSs are always identical to their semantic denotations, which are not context sensitive (see Corollary 4.4.3), NSs will not be associated with any non-trivial scales. The scalarity predictions made by the theory presented in this section are summarized in Table 4.3.

---

24 These facts are proved in Burnett (2012a).
25 The proof of this statement follows the proof of Theorem 4.2.18.
4.4.3 Potential vagueness results

Within the framework developed in this chapter, we can now have a formal definition of the P-vague property that was introduced in Section 3.4 (the property of being able to construct a Soritical series in some context):

**Definition 4.4.7 Potentially vague adjective (formal)**  An adjective P is P-vague just in case there is some model M such that there is some X ⊆ D such that,

1. **Clear case:** There is some a₁ ∈ X such that \( \| P(a_1) \|_X = 1 \).
2. **Clear non-case:** There is some aₙ ∈ X such that \( \| \neg P(a_ₙ) \|_X = 1 \).
3. **Sorites series:** There are a₁, ..., aₙ ∈ X such that \( a_1 \sim_X a_2, a_2 \sim_X a_3, ..., aₙ₋₁ \sim_X aₙ \).

Recall that AA\(^T\)'s have P-vague positive forms and not P-vague negative forms, while AA\(^P\)'s display the reverse pattern. These patterns are direct consequences of the analysis presented so far. I illustrate the pertinent results below for AA\(^T\)'s, noting that the corresponding results for the partial predicates follow from the duality of >\(_P\) and >\(_R\).

**Corollary 4.4.8**  If Q is a total predicate, then Q is P-vague.\(^{26}\)

---

\(^{26}\) **Proof.** Consider the model M such that D = \{a₁, a₂, a₃, a₄, a₅\}. Consider X ⊆ D such that X = \{a₁, a₂, a₃, a₄, a₅\}. Suppose \( \| Q \|_X = \{a₁\} \). Suppose \( \sim_X = \{a₁, a₂, a₃, a₄\} \) and closure under reflexivity and symmetry. Suppose furthermore that \( a₁ >_Q a₂ >_Q a₃ >_Q a₅ \) in M. Therefore,

1. **Clear case:** \( \| Q(a₁) \|_X = 1 \).
2. **Clear non-case:** \( \| \neg Q(a₁) \|_X = 1 \).
3. **Sorites series:** The sequence \( (a₁, a₂, a₃, a₄, a₅) \).

Therefore, Q is P-vague.
Corollary 4.4.9 If \( Q \) is a total predicate, then its negation is not P-vague.\(^{27}\)

Furthermore, since the indifference relations associated with RAs are symmetric, both positive and negative forms of these predicates are correctly predicted to have the potential vagueness property. NSs, on the other hand, are not associated with any non-trivial scales and are therefore predicted to have only precise forms. I therefore conclude that the analysis provided above within DelTCS correctly predicts the vagueness patterns that we see in the adjectival domain.

4.4.4 Other empirical consequences

The proposals made in this chapter make certain empirical predictions that, I argue, are borne out in the data associated with absolute comparatives and gradable uses of non-scalar predicates. Comparatives and the properties of the scales associated with AAs will be the major focus of Chapter 5; however, we can observe that both total and partial comparatives display certain patterns associated with the semantic property of evaluativity,\(^{28}\) which, I argue, follow naturally from the analysis given in this chapter.

Firstly, it is often observed that comparatives formed from most RAs make no claims about whether the greater of the two individuals satisfy the positive form of the predicate to a high degree.\(^{29}\)

\(^{27}\) Proof. Suppose the negation of \( Q \) is P-vague. Then there is some model and some CC \( X \) and some sequence \( a_1 \succ_Q a_2 \succ_Q \ldots \succ a_n \) such that \( \mathcal{I}_X(Q(a_1)) = 1 \), \( \mathcal{I}_X(\neg Q(a_n)) = 1 \), and \( a_n \sim_Q a_{n-1} \sim_X Q a_1 \). Since \( \mathcal{I}_X(Q(a_1)) = 1 \), \( \mathcal{I}_X(Q(a_n)) = 1 \), and \( \mathcal{I}_X(\neg Q(a_n)) = 1 \), \( \mathcal{I}_X(\neg Q(a_n)) = 1 \). Finally, since \( a_n \cdots a_1 \) form a Soritical series, there are some \( a_i, a_{i+1} : a_i \geq_Q a_{i+1} \geq_Q a_n \), and \( \mathcal{I}_X(Q(a_i)) = 1 \) and \( \mathcal{I}_X(\neg Q(a_{i+1})) = 1 \). Furthermore, \( a_{i+1} \sim_Q a_i \). But, by the total axiom, \( a_{i+1} \not\sim_X a_i \). \( \bot \) So the negation of \( Q \) is not P-vague.

\(^{28}\) Rett (2008, p. 9) defines evaluativity as follows: "A construction is evaluative if it makes reference to a degree which exceeds a contextual standard." Note that, strictly speaking, in my analysis, neither total nor partial comparatives are truly "evaluative," since (unlike RAs) neither the semantics nor the pragmatics of AAs involves a contextual standard. However, I keep the term to describe the data discussed in this subsection because the idea of a contextual standard is evoked in other discussions of this data set (cf. Rett (2008) and also Kennedy (2007), who uses different terminology).

\(^{29}\) Exceptions to this generalization are so-called extreme RAs like beautiful and brilliant. I have nothing new to add to the discussion of evaluativity in relative comparatives. For an account of these patterns within DegS, see Rett (2008).

(iv) a. Mary is more beautiful than Sue. (\( \Rightarrow \) Mary is beautiful.)
   b. Mary is more brilliant than Sue. (\( \Rightarrow \) Mary is brilliant.)
(32) a. This dwarf is taller than that dwarf.
   b. This really short stick is longer than that really short stick.
   c. This student in the remedial class is smarter than that student,
      also in the remedial class.

However, comparatives formed from $AA^T$s differ from their relative counterparts in that they seem to be evaluative: the greater individual of the two needs to be at least somewhat close to satisfying the AA's semantic denotation, as shown by the weirdness of comparatives in (33).

(33) a. Storage closet A is emptier than storage closet B.
   # If they both have tons of objects in them, even if B has a couple fewer objects in it.
   b. This twisty staircase is straighter than this other twisty staircase.
   # If they are both really twisty, even if one has one fewer twists than the other.

The “pseudo evaluativity” of total comparatives is straightforwardly predicted by the theory. By the definition of the tolerant comparative (Definition 4.3.5) $a$ will only be tolerantly greater than $b$ if there is some context and some CC in which $a$ is indifferent from some semantically $Q$ individual, and $b$ is not indifferent from any semantically $Q$ individuals. So if there is no CC in which $a$ is indifferent from some individual at the top endpoint of the scale, then for all $b \in D$, $⟨a, b⟩ \notin _Q$.

A second set of data involving evaluativity concerns similar effects with $AA^P$s. As discussed by Rotstein and Winter (2004), Kennedy and McNally (2005), Kennedy (2007), and Rett (2008), the subject of a partial comparative is always understood to have the property denoted by the $AA^P$. That is, the inferences in (34) hold:

(34) a. This shirt is dirtier than that shirt ⇒ This shirt is dirty.
   b. This towel is wetter than that towel ⇒ This towel is wet.
   c. John is sicker than Mary ⇒ John is sick.
   d. This stick is more bent than that stick ⇒ This stick is bent.

Again, these data are predicted by the theory. As discussed in Section 4.4.3, the only non-trivial scale that can be associated with a $AA^P$ is the strict one. By the definition of the tolerant/strict comparative (Definition 4.3.5), subjects of strict comparatives must be in the strict denotation of the predicate in some CC. Since strict denotations are
subsets of semantic denotations, subjects of strict comparatives must be in the semantic denotation of the predicate in that CC. Since the semantic denotations of AAs are invariant across CC, if a comparative with a $AA^P Q$ is true under its strict interpretation, then its subject must always be in $Q$’s semantic denotation. Thus, the inferences in (34) go through.

Finally, I argue that the account provided here can shed light on some empirical disagreements in the literature concerning the evaluativity of the object of total comparatives. As highlighted by Toledo and Sassoon (2011, p. 7), there is some disagreement about whether or not the object in the than clause of a total absolute comparative can be in the extension of the positive form of the $AA^T$. For example, Kennedy and McNally (2005) (following Unger (1975)) claim that (35-a) is necessarily false; whereas, Rotstein and Winter (2004) say that the almost identical example in (35-b) can be true.

\begin{enumerate}
\item The red towel is cleaner than the blue one, but both are clean. 
\hspace{2em} (Kennedy and McNally, 2005)
\item Both towels are clean, but the red one is cleaner than the blue one.
\hspace{2em} (Rotstein and Winter, 2004)
\end{enumerate}

The approach developed so far has a straightforward account of why the examples in (35-a) and (35-b) could alternatively be viewed as both contradictions and non-contradictions: The only non-trivial comparative relation for a $AA^T$ is the tolerant comparative ($>_t clean$). Thus, since the blue towel is not the cleanest object (i.e. the red one is cleaner), we know that the blue towel is not in the semantic denotation of clean, no matter what the contextually given CC is (i.e. blue $\notin \llbracket clean \rrbracket_X$, for all $X \subseteq D$). Thus, if we are speaking precisely, both towels cannot both be clean and one be cleaner than the other. However, if we are speaking loosely, it might be possible to consider the blue towel tolerantly clean even if it is not maximally clean (i.e. blue $\in \llbracket clean \rrbracket_X$, for some contextually given $X$). The account given in this chapter predicts that, for people who are a bit more hard-nosed (who require a higher level of precision, cf. Lewis (1979)), (35-a) and (35-b) should be contradictions; however, these sentences are not contradictions for those of us who are a bit more laissez-faire. In other words, the arguments in (35-a) and (35-b)

\footnote{See also Toledo and Sassoon (2011) for an account of the contrast in (35-a) and (35-b) in similar terms.
are contradictory in $\models^{tc}$, $\models^{sc}$, and $\models^{cc}$, but are not contradictions in $\models^{tt}$, $\models^{st}$, or $\models^{ct}$. I therefore conclude that the model makes correct predictions with respect to the “evaluativity” and entailments of AA$^P$s and AA$^T$s in comparatives.

Turning now to NSs: In Section 4.3, I proposed that pragmatic denotations of NSs were governed by an additional constraint that forces them to coincide with their semantic denotations. A simple way of describing this proposal is to say that NSs are conventionally associated with a higher degree of precision than either their relative or absolute counterparts. Since AAs and NSs were given the same classical semantic analysis in Section 4.2, I propose that what differentiates NSs from AAs is in their pragmatics, not their semantics.

(36) **The AA/NS distinction:**

The differences between AAs and NSs are purely pragmatic: at the level of their semantic denotations, they are identical.

As such, I propose the following analysis of gradable uses of NSs (these uses are sometimes called *coerced* NSs):

(37) **Scalar “coercion”:**

Coerced/gradable NSs are subject to all the same constraints as regular NSs, except BP.

This approach is very different from certain other current views (such as the ones in the popular DegS framework) that propose a semantic and even a syntactic difference between these predicates. In this section, I give three arguments in favor of the position that the AA/NS distinction should be reduced to facts about the use of these predicates, in particular, how precisely we tend to use them. I call the arguments: 1) The technical nature of non-scalars, 2) The ease of “coercion”, and 3) The NS$^C \rightarrow$ AA dependency.

### 4.4.4.1 The technical nature of non-scalar adjectives

The first argument that the AA/NS distinction has to do with precision concerns the nature of the inventory of NSs. In particular, if we look at the (extended) inventory of NSs discussed throughout this work (38), we can notice that the vast majority (if not all) of them come from domains in which precision is important: logic and mathematics (*atomic, hexagonal, square, even, odd, prime*), biology (*pregnant, dead, male, female*),
physics (opaque, transparent, visible, invisible), and the law (legal, illegal, Canadian, French).

(38) **Non-scalar adjectives:**
atomic, geographical, polka-dotted, pregnant, legal, illegal, dead, hexagonal, square, male, female, even, odd, prime, Canadian, French, perfect, imperfect, opaque, transparent, visible, invisible . . .

The connection between the register and communicative domain in which a term is used and whether or not it is scalar is straightforwardly expected in a theory in which scalarity is a pragmatic matter. Although one could perhaps invent a historical explanation for them, these lexicalization patterns are somewhat puzzling for an analysis in which AAs like empty and straight have an inherently gradable meaning, but NSs like illegal and perfect do not.

Of course, since the test for being a member of the NS class is whether or not, out of the blue, you sound “weird” in a comparative construction (i.e. ?This shape is more hexagonal that that one vs ✓ This room is emptier than that one), some readers may have different judgments about the non-scalar status of some of the words in (38). So they might not find the generalization concerning scalar/non-scalar lexicalization patterns so convincing. But this observation about variation in judgments of non-scalarity brings me to the second argument: the ease of “coercion.”

4.4.4.2 **The ease of coercion**  As I mentioned, a characteristic property of NSs is their strangeness in comparative constructions; however, another characteristic property of these predicates is the ease with which, given an appropriate context, they can become gradable. In other words, although we noted that the adjectives in (38) sound strange in the comparative out of context, it is perfectly natural to use many of them as follows:

(39) a. This dress is more polka-dotted than that one; it has more dots on it.
b. This room is more square than that room.
c. Sarah is more pregnant than Sue; Sarah is showing more.
d. Murder is more illegal than smoking pot.
e. Zombie A is more dead than zombie B.
f. France is more hexagonal than Canada.
and so on . . .
A striking example of a scalar use of the adjective *one-armed*, which is a prototypical example of a non-scalar, is shown in (40) from the television show *Community*.

(40) Annie: Where did you get this?
    Abed: Some *one-armed* guy with a scar dropped it off. He said he
    was Starburns’ lawyer.
    Troy: How *one-armed* was he? Tell me when to stop.
       *(indicates increasingly high cut-off points on his left arm).*
       *(Community. Season 3, episode 18 (NBC).)*

Even some of the more “mathematical” terms in (38) can acquire agradable meaning. For example, Armstrong et al. (1983) show that even
if they admit that a particular well-defined concept like *odd* or *even* is not
inherently gradable, participants can still order individuals with respect
to how well they exemplify the concept. With this in mind, we could
form comparatives like those in (41-a) and (41-b), as well as in (41-c)
from Rett (2013, p. 9), which is also inspired by the results of Armstrong
et al. (1983).

(41) a. 4 is more even than 34.
    b. 3 is more odd than 447.
    c. 7 is more prime than 2.

In sum, it seems to be a general property of NSs that, with very little
effort, they can appear in degree constructions and (as discussed in
Chapter 3) when they do so, they become context sensitive and vague.31
Of course, an analysis in which there was a scalar lexical transformation
process in the grammar (perhaps some sort of degree-argument-adding
operation) could account for the ease of coercion (this process could be
highly productive). However, the pragmatic analysis that I have given
needs no such morpho-semantic operation: NSs are simply AAs that

31 In fact, this seems to be a property of precise expressions more generally. As observed
by Russell (1923), even logical expressions can easily become vague when used in contexts
in which a lower level of precision is permissible. He says (p. 86):

There is, however, less vagueness about logical words than about the words of daily life,
because logical words apply essentially to symbols, and may be conceived as applying
rather to possible than to actual symbols. We are capable of imagining what a precise
symbolism would be, though we cannot actually construct such a symbolism. Hence we
are able to imagine a precise meaning for such words as “or” and “not”. We can, in fact,
see precisely what they would mean if our symbolism were precise. All traditional logic
habitually assumes that precise symbols are being employed. It is therefore not applicable
to this terrestrial life, but only to an imagined celestial existence.
tend to be used with a higher level of precision. Alternatively, in my analysis, we could describe AAs as simply NSs that tend to be used loosely.

4.4.4.3 The NS\(^c\) → AA dependency  My final argument in favor of an analysis in which NSs are semantically identical to AAs comes from an empirical observation (already discussed in Chapter 3) about the properties of gradable NSs. The generalization is the following:

(42) The NS\(^c\) → AA dependency:
When NSs become gradable, they become AAs.

For example, we saw in Chapter 3 that, although they are existentially context sensitive, gradable non-scalars uniformly fail the definite description test:

(43)  a. Pass me the hexagonal one.
    (But both/neither are/is hexagonal!)
    b. Show me the illegal one.
    (But both/neither are/is illegal!)
    c. Show me the dead one.
    (But both/neither are/is dead!)

An analysis in which scalar coercion is a morpho-semantic operation would have to build (42) into the operation, and this would raise the question of why we cannot coerce an NS into an RA. However, the dependency between coerced NSs and AAs is a consequence of the pragmatic analysis that I have given: NSs have context-independent semantic denotations, just like AAs. I therefore conclude that a pragmatic analysis of the AA/NS distinction has certain empirical advantages over a morpho-semantic analysis, particularly when we look at data associated with gradable NSs.

Finally, we can observe that the pragmatic loosening operation that NSs are proposed to undergo is predicted to have important consequences on the scalarity of these elements: in principle, these predicates can be associated with both non-trivial tolerant and strict scales (note that the models that demonstrate the scalarity of both AA\(^T\)s and AA\(^P\)s are models for coerced NSs). Thus, the results concerning the (non-)scalarity of adjectival predicates are shown in Table 4.4, where ♦✓ stands for possibly associated with a non-trivial SWO (depending on context).
Table 4.4 Scality patterns

<table>
<thead>
<tr>
<th>Adjective</th>
<th>( &gt;_p : \text{non-trivial SWO?} )</th>
<th>( &gt;_p^T : \text{non-trivial SWO?} )</th>
<th>( &gt;_p^L : \text{non-trivial SWO?} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>AA(^T)</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>AA(^P)</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>NS</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Gradable NS</td>
<td>×</td>
<td>◊✓</td>
<td>◊✓</td>
</tr>
</tbody>
</table>

Table 4.5 Proposed constraints for (non-)scalar adjectives

<table>
<thead>
<tr>
<th>Constraint</th>
<th>RA</th>
<th>AA(^T)</th>
<th>AA(^P)</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constraint on ( [\cdot] )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Reversal (NR)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Upward Difference (UD)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Downward Difference (DD)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Absolute Adjective Axiom (AAA)</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Constraint on (~)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reflexivity (R)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tolerant Convexity (TC)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Strict Convexity (SC)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Granularity (G)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Minimal Difference (MD)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Contrast Preservation (CP)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Symmetry (S)</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Total Axiom (TA)</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Partial Axiom (PA)</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Be Precise (BP)</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

4.5 Conclusion

This chapter presented a new logical framework for the analysis of adjectival vagueness, context sensitivity, and gradability: the DelTCS system. Within this framework, I set an analysis of the classical, tolerant, and strict semantics of different classes of adjectival predicates. The bulk of the analysis consisted in proposing a series of constraints on the interpretations of predicates across CCs and the definition of the \(~\) as it applies to adjectives of different classes. The constraints are summarized in Table 4.5.

As the table suggests, the constraints associated with \(~\) can be grouped into three distinct classes. The first class \((R, TC, SC, G, MD,\)
and CP) is meant to characterize cognitive indifference relations and their basic distribution in different CCs. As such, these constraints apply to all adjectival predicates. The second class of constraints (S, TA, and PA) deals with the symmetry of the indifference relations. I argued that, although the \( \sim_p \)s are generally symmetric, indifference relations with AAs display a limited asymmetry with respect to whether they can relate individuals along the border of the semantic denotation of the adjective. I hypothesized that these asymmetries could be instances of similar “prototypicality” effects that have been independently observed in the psychological literature. Finally, the third class of constraints (consisting solely of BP) deals with how adjectival predicates are used in conversation. As such, unlike the others, BP is easily violable if the context allows for a lower level of precision than usual.

I showed that, with this axiom set, the multi-valued Delineation logical system that I proposed correctly derives the context sensitivity and potential vagueness patterns that were argued for in Chapter 3. Furthermore, I showed that my analysis provides a solution to the puzzle of the gradability of AAs that was presented in Section 4.2: the non-trivial scales that are associated with AAs are constructed from their pragmatic meanings not their classical semantic meaning. Additionally, we saw that the analysis makes other predictions concerning the pragmatic scales associated with RAs and the non-gradability of NSs. In Chapter 5, I will show that the analysis developed so far makes even finer predictions about the properties of the orders that are associated with the various classes of adjectives. In particular, I will show that my proposals to account for the context sensitivity and vagueness patterns give us a full account of a large and theoretically important data set: the adjectival scale structure patterns.

4.6 Appendix: Longer proofs

In this section, I show that the constraints placed on the definition of the \( \sim \) function with AAs have a similar effect to van Benthem’s constraints with RAs. In particular, these constraints allow us to extract non-trivial strict weak order relations from the type 2 context sensitivity of the pragmatic denotations of AAs.

We first note a series of facts about the \( \succ^\iota_Q \) and \( \succ^\epsilon_Q \) relations and their relationship to \( \succ_Q \). For example, MD ensures that semantic absolute denotations are subsets of tolerant denotations. This lines up with the general fact in the basic TCS system (as presented in Cobreros et al.’s
Theorem 4.6.1 Tolerant subset. If $Q$ is an AA, then $>Q \subseteq >^t_Q$.

Proof. Let $a_1, a_2 \in D$ such that $a_1 >_Q a_2$ to show that $a_1 >^t_Q a_2$. Since $a_1 >_Q a_2$, there is some $X \subseteq D$ such that $\|Q(a_1)\|_X = 1$ and $\|Q(a_2)\|_X = 0$. Now consider $\{a_1, a_2\} \subseteq D$. By DD, $\|Q(a_1)\|\{a_1, a_2\} = 1$ and $\|Q(a_2)\|\{a_1, a_2\} = 0$. By the definition of $\|\cdot\|_t$, $\|Q(a_1)\|\{a_1, a_2\} = 1$. Furthermore, by MD, $a_1 \not\sim^Q \{a_1, a_2\} Q a_2$. So, by the definition of $>^t_Q$ (Definition 4.3.5), $a_1 >^t_Q a_2$.

Additionally, we can show that classical denotations of absolute comparatives are also included in the strict denotations of these constituents. This is a difference with the basic TCS system, since generally strict denotations are subsets of classical ones in this framework. Although Theorem 4.6.2 is slightly at odds with the properties of basic TCS, we might note, firstly, that basic predicate strict denotations are still subsets of classical semantic denotations in DeTCS and secondly that the language of the framework presented in Cobreros et al. (2012b) does not even contain comparative relations whose interpretation is constructed in the manner proposed in DeTCS, so it is not clear that Cobreros et al.’s (2012b) Lemma 1 should necessarily apply to constituents other than basic predicates in DeTCS. Thus, I do not believe that, at the end of the day, Theorem 4.6.2 goes against the spirit of the original TCS system.

Theorem 4.6.2 Strict subset. $>Q \subseteq >^s_Q$.

Proof. Let $a_1 >_Q a_2$ to show $a_1 >^s_a \{a_1, a_2\}$. Since $a_1 >_Q a_2$, by DD, $\|Q(a_1)\|\{a_1, a_2\} = 1$ and $\|Q(a_2)\|\{a_1, a_2\} = 0$. Therefore, by MD, $a_2 \not\sim^Q \{a_1, a_2\} Q a_1$. So, by the definition of $\|\cdot\|_s$, $\|Q(a_1)\|\{a_1, a_2\} = 1$ and $\|Q(a_2)\|\{a_1, a_2\} = 0$. So $a_1 >^s_Q a_2$.

Secondly, with only Tolerant/Strict Convexity, we can prove that van Bentham’s NR holds at the tolerant and strict levels. It is in this sense that, as I mentioned, Tolerant/Strict Convexity can be viewed as the tolerant/strict correspondent of NR.

Lemma 4.6.3 No tolerant reversal (T-NR): For $X \subseteq D$, if $\|Q(a_1)\|_X = 1$ and $\|Q(a_2)\|_X = 0$, then there is no $X' \subseteq D$ such that $\|Q(a_2)\|_{X'} = 1$ and $\|Q(a_1)\|_{X'} = 0$. 
Proof. Suppose \( \|Q(a_1)\|_X^I = 1 \) and \( \|Q(a_2)\|_X^I = 0 \). Suppose, for a contradiction, that there is an \( X' \subseteq D \) such that \( \|Q(a_1)\|_{X'}^I = 1 \) and \( \|Q(a_2)\|_{X'}^I = 0 \). Therefore, \( a_1 >'_Q a_2 \) and \( a_2 >'_Q a_1 \). Furthermore, by assumption and the definition of \( \|Q\|_X \), there is some \( a_3 \sim_X Q a_1 \) such that \( \|Q(a_3)\|_X = 1 \), and \( a_3 \not\sim_Q a_2 \). Thus \( a_3 >'_Q a_2 \) and so \( a_3 >'_Q a_2 >'_Q a_1 \). Since \( a_3 \sim_X Q a_1 \), by TC, \( a_3 \sim_Q a_2 \). \( \square \)

**Theorem 4.6.4 No strict reversal.** For some \( X \subseteq D \), let \( \|Q(a_1)\|_X^I = 1 \) and \( \|Q(a_2)\|_X^I = 0 \). Then, there is no distinct \( X' \subseteq D \) such that \( \|Q(a_1)\|_{X'}^I = 1 \) and \( \|Q(a_2)\|_{X'}^I = 0 \).

Proof. Suppose, for a contradiction, that there is an \( X' \subseteq D \) such that \( \|Q(a_1)\|_{X'}^I = 1 \) and \( \|Q(a_2)\|_{X'}^I = 0 \). Since \( \|Q(a_1)\|_{X'}^I = 1 \), there is some \( a_3 \sim_X Q a_1 \) such that \( \|Q(a_3)\|_{X'} = 0 \). Therefore, since \( \|Q(a_2)\|_{X'} = 1 \), \( a_2 >'_Q a_3 \). By assumption, \( a_3 >'_Q a_2 >'_Q a_3 \), and \( a_3 \sim_X Q a_4 \). Therefore, by SC, \( a_3 \sim_X Q a_2 \). But \( \|Q(a_2)\|_{X'}^I = 1 \). \( \square \)

Using the complete axiom set, we can show that, for all absolute predicates \( Q >'_Q \) is a strict weak order (irreflexive, transitive, and almost-connected).

**Lemma 4.6.5 Irreflexivity.** For all \( a_i \in D \), \( a_i \not>'_Q a_i \).

Proof. Since it is impossible, for any \( X \subseteq D \), for an element to be both included in \( \|Q\|_X^I \) and not included in \( \|Q\|^I_{X'} \), by the definition of \( \|\cdot\|^I \), \( >'_Q \) is irreflexive. \( \square \)

We now prove transitivity for \( >'_Q \).

**Lemma 4.6.6 Transitivity.** For all \( a_1, a_2, a_3 \in D \), if \( a_1 >'_Q a_2 \) and \( a_2 >'_Q a_3 \), then \( a_1 >'_Q a_3 \).

Proof. Suppose \( a_1 >'_Q a_2 \) and \( a_2 >'_Q a_3 \) to show that \( a_1 >'_Q a_3 \). Then there is some \( X \subseteq D \) such that \( \|Q(a_1)\|_X^I = 1 \) and \( \|Q(a_2)\|_X^I = 0 \). Thus, there is some \( a_4 \in X : \|Q(a_4)\|_X = 1 \) and \( a_4 \sim_X Q a_1 \). Now consider \( X \cup \{a_3\} \). By the AAA and the assumption that \( a_1 >'_Q a_2 \) and \( a_2 >'_Q a_3 \), \( \|Q(a_2)\|_{X \cup \{a_3\}} = 0 \) and \( \|Q(a_3)\|_{X \cup \{a_3\}} = 0 \). \textbf{Case 1:} \( X \cup \{a_3\} = X \). Since \( \|Q(a_1)\|_X^I = 1 \) and \( \|Q(a_3)\|_X^I = 0 \), \( x >'_Q z \). \textbf{Case 2:} \( X \subseteq X \cup \{a_3\} \). Since \( X \subseteq X \cup \{a_3\} \) and \( a_4 \sim_X Q a_1 \), by Granularity, \( a_4 \sim_Q X_{\cup \{a_3\}} a_1 \). By the AAA, \( \|Q(a_4)\|_{X \cup \{a_3\}} = 1 \). So \( \|Q(a_1)\|_{X \cup \{a_3\}} = 1 \). Suppose, for a contradiction that \( \|Q(a_3)\|_{X \cup \{a_3\}} = 1 \). Then there is some \( a_5 \in X \cup \{a_3\} \) such that \( a_5 \sim_X Q a_5 \). But \( a_5 \sim_Q X \cup \{a_3\} a_5 \). Therefore, since \( a_1 >'_Q a_2 \) and \( a_2 >'_Q a_3 \), \( a_1 >'_Q a_3 \). \( \square \)
\{a_3\} : \|Q(a_3)\|_{X \cup \{z\}} = 1 and a_5 \sim_{X \cup \{a_3\}} a_5. By assumption and since \|Q(a_2)\|_X = 0, by MD and Tolerant Subset, a_5 \succ^Q a_2 \succ^Q a_3. Since a_5 \sim_{X \cup \{a_3\}} a_3, by TC, a_5 \sim_{X \cup \{a_3\}} a_2. Since \|Q(a_2)\|_X = 0, a_5 \not\sim_X a_2. So by CP, since \(X \cup \{a_3\} \sim X = \{a_3\}, a_5 \not\sim_{X \cup \{a_3\}} a_3. \) So \(\|Q(a_3)\|_{X \cup \{z\}} = 0,\) and \(a_1 \succ^Q a_3, \checkmark\)

Finally, we can prove almost connectedness.

**Lemma 4.6.7 Almost connected.** For all \(a_1, a_2 \in D,\) if \(a_1 \succ^Q a_2\) then for all \(a_3 \in D,\) either \(a_1 \succ^Q a_3\) or \(a_3 \succ^Q a_2.\)

**Proof.** Let \(a_1 \succ^Q a_2\) and \(a_3 \not\sim^Q a_2\) to show \(a_1 \succ^Q a_3.\)

**Case 1:** \(\|Q(a_1)\|_D = 1.\) Since \(a_1 \succ^Q a_2\) and \(a_3 \not\sim^Q a_2,\) \(\|Q(a_3)\|_D = 0.\) So \(a_1 \succ^Q a_3\) and, by Theorem 4.6.1, \(a_1 \succ^Q a_3. \checkmark\)

**Case 2:** \(\|Q(a_1)\|_D = 0.\) Since \(a_1 \succ^Q a_2,\) there is some \(X \subseteq D\) such that \(\|Q(a_1)\|_X = 1\) and \(\|Q(a_2)\|_X = 0.\) So there is some \(d \in X : \|Q(a_4)\|_X = 1\) and \(a_4 \sim^X a_1.\) Consider \(X \cup \{a_3\}.\) Since \(a_3 \not\sim^Q a_2, a_1, a_2, a_3 \notin \|Q\|_{X \cup \{a_1\}}.\) Since \(a_4 \sim^X a_1,\) by Granularity, \(a_4 \sim^Q a_1\) and by the AAA, \(\|Q(a_4)\|_{X \cup \{a_1\}} = 1.\) So \(\|Q(a_1)\|_{X \cup \{z\}} = 1.\) Now suppose for a contradiction that \(\|Q(a_3)\|_{X \cup \{a_1\}} = 1.\) Then there is some \(a_5 \in X \cup \{a_3\} : \|Q(a_5)\|_{X \cup \{a_1\}} = 1 and a_5 \sim^X a_3.\) Since \(\|Q(a_5)\|_{X \cup \{a_1\}} = 1\) and \(\|Q(a_2)\|_{X \cup \{a_1\}} = 0, a_5 \succ Q a_2;\) so by Theorem 4.6.1, \(a_5 \succ^Q a_2.\)

Furthermore, since, by assumption, \(z \not\sim^Q y, y \succ^Q z.\) Since \(d' \succ^Q y \succ^Q z\) and \(d' \sim^X z, by TC, d' \sim^X z y.\) However, since \(\|Q(a_2)\|_X = 0,\) and by the AAA, \(\|Q(a_5)\|_X = 1, a_5 \not\sim_X a_2.\) Since \(X \subset X \cup \{a_3\}\) and \(a_5 \sim^X a_3, by CP, there is some \(a_6 \in X \cup \{a_3\} - X\) such that \(a_5 \not\sim_X a_6.\) Since \(X \cup \{a_3\} - X = \{a_3\}, a_5 \not\sim^X a_3.\) \(\Downarrow\) So \(\|Q(a_3)\|_{X \cup \{z\}} = 1 and a_1 \succ^Q a_3, \checkmark\)

We can now prove one of the two main theorems of this section:

**Theorem 4.6.8** If \(Q\) is an absolute predicate (i.e. satisfies the AAA and the pertinent axioms in Table 4.1), \(<^Q\) is a strict weak order.

**Proof.** Immediate from Lemmas 4.6.5, 4.6.6, and 4.6.7. \(\checkmark\)

Using the same axiom set, we can also show that the strict scale \(<^Q\) is a strict weak order.
Lemma 4.6.9 Irreflexivity. If $Q$ is an absolute predicate, $>_Q^s$ is irreflexive.

Proof. Immediately from the definition of $>_Q^s$.

Lemma 4.6.10 Transitivity. If $Q$ is an absolute predicate, $>_Q^s$ is transitive.

Proof. Suppose $a_1 >_Q^s a_2$ and $a_2 >_Q^s a_3$ to show $a_1 >_Q^s a_3$. Since $a_2 >_Q^s a_3$, there is some $X \subseteq D$ such that $\|Q(a_2)\|_X^s = 1$ and $\|Q(a_2)\|_X = 0$. So there is some $a_4 \sim_X a_3$ such that $\|Q(a_4)\|_X = 0$. Clearly, $a_2 >_Q a_4$. Now consider $X \cup \{a_1\}$. By the AAA, $\|Q(a_4)\|_{X \cup \{a_1\}} = 0$ and, by Granularity, $a_4 \sim_{X \cup \{a_1\}} a_1$. So $\|Q(a_3)\|_{X \cup \{a_1\}} = 0$. Suppose for a contradiction that $\|Q(a_1)\|_{X \cup \{a_1\}} = 0$. So there is some $a_5 \sim_{X \cup \{a_1\}} a_1$ such that $a_5 >_Q^s a_2, a_2, a_1$ and, by Theorem 4.6.2, $a_1 >_Q^s a_5$ and by Theorem 4.6.2, $a_1 >_Q^s a_2, a_2, a_1$. So $a_5 >_Q a_5$, and by Theorem 4.6.2, $a_1 >_Q^s a_5$, $a_5 >_Q a_5$. Since $a_5 \sim_{X \cup \{a_1\}} a_1$, by SC, $a_5 \sim_{X \cup \{a_1\}} a_2$. Since $\|Q(a_5)\|_X = 1$, $a_5 >_Q a_2$. So, by CP, $a_4 \sim_{X \cup \{a_1\}} a_1$, $\|Q(a_1)\|_{X \cup \{a_1\}} = 0$. So $\|Q(a_1)\|_{X \cup \{a_1\}} = 0$.

Case 1: $\|Q(a_1)\|_{X \cup \{a_1\}} = 1$. Then $a_1 >_Q^s a_2$. √ Case 2: $\|Q(a_1)\|_{X \cup \{a_1\}} = 0$.

Then there is some $a_4 \sim_{X \cup \{a_1\}} a_1$ such that $\|Q(a_4)\|_{X \cup \{a_1\}} = 0$. Suppose for a contradiction that $a_4 \neq a_1$. So $a_4, a_4 \in X$. Since $\|Q(a_4)\|_X = 1$, $a_4 >_Q a_1$. So, by CP, $a_4 \sim_{X \cup \{a_1\}} a_3$. Since $\|Q(a_4)\|_{X \cup \{a_1\}} = 0$, for all $a_6 \in X \cup \{a_3\}$, $a_6 \geq a_4$. So $a_2 \geq a_4$ and $a_3 \geq a_4$. Since $a_1 >_Q a_2$, $a_2 \geq a_4$ and $a_4 \sim_{X \cup \{a_1\}} a_1$, by SC, $a_4 \sim_{X \cup \{a_1\}} a_2$. Since $a_3 >_Q a_2$, by Theorem 4.6.4, $a_3 \geq a_4$, $a_2 \geq a_4$. Since $a_4 \sim_{X \cup \{a_1\}} a_2$, by SC, $a_4 \sim_{X \cup \{a_1\}} a_3, a_3 \sim_{X \cup \{a_1\}} a_4$. Since $a_4 \sim_{X \cup \{a_1\}} a_2$, by SC, $a_4 \sim_{X \cup \{a_1\}} a_3$. So $a_4 = a_3$. Since $a_1 >_Q a_2$, $\|Q(a_1)\|_{X \cup \{a_1\}} = 1$, so $a_1 >_Q a_3$, and, by Theorem 4.6.2, $a_1 >_Q a_3$. √

The second main theorem of this section is the following:

Theorem 4.6.12 If $Q$ is an absolute predicate, $>_Q^s$ is a strict weak order.

Proof. Immediately from Lemmas 4.6.9, 4.6.10, and 4.6.11.
5

Scale structure in Delineation semantics

5.1 Introduction
This chapter presents both new and previously discussed data associated with the scale structure of members of the four principle classes of adjectives that are studied in this work. Following much previous work, I argue that the adjectives in each of the studied classes are associated with scales that have different properties. In particular, as we will see, there are empirical arguments for proposing that total absolute adjectives (AA^T) (like empty and straight), are associated with scales that have maximal elements, partial absolute adjectives (AA^P) (like wet and dirty), are associated with scales that have minimal elements, and relative adjectives (RAs) (like tall and expensive) are associated with scales that have neither minimal nor maximal elements. Additionally, we will see that, when non-scalar adjectives (NSs) (like dead and hexagonal) are used as scalar adjectives, they can be associated with scales that have both minimal and maximal elements.

Furthermore, we will see in this chapter that the association of an adjective with a scale with the correct properties is already predicted by the analysis presented in Chapter 4. In other words, once we have an (independently necessary) analysis of context sensitivity and (potential) vagueness in the adjetival domain, we get an analysis of adjetival scale structure “for free.”

The chapter is laid out as follows: In Section 5.2, I present the data associated with the scale structure patterns found in the adjetival domain. I first introduce tests for the presence of scalar endpoints and I apply these tests to relative scalar adjectives (RAs), AA^T's and AA^P's. Then, I apply them to gradable uses of NSs. In Section 5.3, I present the scale structure results of my analysis. I show that the patterns argued for in Section 5.2 are predicted by the theory given in Chapter 4. I therefore conclude that the Delineation Tolerant, Classical,
Strict (DelTCS) framework provides a broad and general perspective on the analytical relationships between context sensitivity, vagueness, and scalarity that exist in natural language.

5.2 Scale structure patterns

This section presents an illustrative subset of data concerning the boundedness or the unboundedness of the orders associated with adjectival predicates. I will first consider only scalar adjectives, leaving a discussion of non-scalars to Section 5.2.1.

Although there are many different diagnostics for scale structure in the literature (see Rotstein and Winter, 2004; Kennedy and McNally, 2005; Kennedy, 2007, among others), the way that we will diagnose scalar properties of different kinds of predicates is through looking at how these predicates pattern in constructions whose semantics appears to require some notion of measurement based on a non-trivial order, possibly with a maximal or minimal element.\(^1\)

In what follows, I will argue (following others) that AA\(^T\)'s are associated with scales (i.e. non-trivial orderings) that have maximal elements, AA\(^P\)'s are associated with scales that have minimal elements, and RAs are associated with scales that have neither maximal nor minimal elements.

A first argument that AA\(^T\)'s are associated with scales with a maximal element comes from the distribution of scalar modifiers. For example, as observed by Cruse (1986), Rotstein and Winter (2004), and Kennedy and McNally (2005) (among others), while *almost* is perfectly fine with AA\(^T\)'s, it is strange with AA\(^P\)'s (see (i)).\(^2\)

\(^{\text{i}}\) John is *almost* hungry: four hours after breakfast, he is no longer satiated from breakfast; he is not yet hungry, but he is already starting to think about lunch.

\(^{\text{1}}\) The formal notion of a scale will be defined for the DelTCS framework in Section 5.3, and for the Degree Semantics (DegS) framework in Chapter 6. We use the intuitive notion of order in the empirical discussion in this section.

\(^{\text{2}}\) Rotstein and Winter (2004) notice that, to the extent that *almost* with partial adjectives is OK for some speakers, *almost* with AA\(^P\)'s requires a somewhat strange context (see their example (p. 266)).

(i) a. This towel is *almost* dry/*wet.
   b. The stick is *almost* straight/*bent.
   c. The table is *almost* clean/*dirty.

\(\text{(1)}\) a. This towel is *almost* dry/*wet.
   b. The stick is *almost* straight/*bent.
   c. The table is *almost* clean/*dirty.
scale structure patterns 107

d. The metal is almost flat/*curved.
e. John is almost bald.

Furthermore, almost is generally much less acceptable with RAs than with total adjectives, as in (2).

(2)  
   a. John is almost *fat/*tall/*wide.
   b. This watch is almost *expensive/*attractive/*fashionable.

If we adopt a simple and intuitive analysis of almost as an item that picks out just those individuals that are close to (but not at) the top endpoint of a scale, then we can explain why only AA\textsuperscript{T}s are possible with this modifier: only they have top endpoints.

Likewise, as discussed by Kennedy and McNally (2005) and Sauerland and Stateva (2007) (among many others), when completely appears with AA\textsuperscript{T}s, it restricts the extension of the predicate to only those individuals that satisfy the adjective to the highest possible degree. Thus, we find a maximal interpretation with completely in sentences like (3).

(3)  
   a. John is completely bald ≈ John has the highest degree of baldness.
   b. This room is completely empty ≈ This room has the highest degree of emptiness.

On the other hand, when it appears with other kinds of scalar adjectives, the maximal interpretation disappears: with RAs, either completely is ungrammatical (4), or it receives a mereological interpretation; that is, it can be paraphrased by ‘in all parts/aspects’ (see also Moltmann (1997) for the mereological reading of completely) (5).

(4)  
   a. *John is completely tall.
   b. *Mary is completely short.

(5)  
   a. John is completely happy.
   b. Susan is completely red.

We find the same pattern with AA\textsuperscript{P}s: in the examples in (6), only the mereological reading (not the maximal one) is available.

(6)  
   a. The cat is completely wet.
   b. The cat is completely dirty.
c. For a student who just moved here, she is very familiar with the class routines and her teachers’ expectations. In fact, she’s completely familiar.

McNally (2011) (p. 6)

These patterns can immediately be accounted for if AA\textsuperscript{T}s are associated with orderings with maximal elements, but AA\textsuperscript{P}s and RAs are not.

A second argument that the scales associated with AA\textsuperscript{T}s have a maximal (unlike the scales associated with RAs and AA\textsuperscript{P}s) comes from the aspectual properties of strong adjectival resultative secondary predication constructions. These are complex verbal predicates composed of an atelic activity verb like hammer or wipe (7), and a secondary adjectival predicate that specifies the result of the action described by the main verb. With the addition of an adjective (of the appropriate type), the construction as a whole gets a telic (i.e. bounded) interpretation, as shown in (8).

(7) a. John hammered the metal (?in an hour/for an hour).
    b. John wiped the table (?in an hour/for an hour).

(8) a. John hammered the metal flat (in an hour/*for an hour).
    b. John wiped the table clean (in an hour/*for an hour).

As observed by Green (1972) and Dowty (1979), not all adjectives can be the secondary predicate of a resultative construction. While clean, dry, and smooth are acceptable, damp, dirty, stained, and wet are unacceptable. When it was first observed, this distribution pattern appeared puzzling, since there is nothing incoherent about the action of wiping something and having that action cause it to be damp/dirty/stained/wet.\footnote{See Dowty (1979) and Goldberg and Jackendoff (2004) for the claim that contrasts such as (10) show that resultative secondary predicates are idiomatic expressions. See Wechsler (2005b) for a criticism of this approach.}

(9) He wiped it clean/dry/smooth/*damp/*dirty/*stained/*wet.

Green (1972) (her (5b–6b)).

(10) John hammered the metal flat/straight/*long/*expensive.\footnote{Imagine that the metal increases in value the more that it is worked.}

More recently, authors such as Wechsler (2005a) and Beavers (2008) (among others) have proposed that the following generalization governs
the distribution of adjectives in resultative secondary prediction constructions, which is empirically supported by a corpus investigation in Boas (2003) and Wechsler (2005a):

(11) Wechsler’s generalization:
Only AA$^T$s are licensed as strong resultative secondary predicates.

Example (11) is a robust generalization. As shown in (12) and (13), Italian (a Romance language that allows such constructions) and Dutch show the same distinctions as English.

(12) Italian:

| a. Gianni ha battuto il ferro piatto piatto. |
| Gianni has beaten the iron flat flat       |
| ‘Gianni beat the iron flat.’ (AA$^T$s)     |
| b. *Gianni ha battuto il ferro lungo lungo |
| Gianni has beaten the iron long long.      |
| ‘*Gianni beat the iron long.’ (RA)         |

(13) Dutch:

| a. Jan heef het metaal plat gehamerd. |
| Jan has the metal flat hammered       |
| ‘Jan hammered the metal flat.’         |
| b. *Jan heef het metaal lang gehamerd. |
| Jan has the metal long hammered.      |
| ‘*Jan hammered the metal long.’        |

Why should (11) hold? The scale structure-based explanation of the distribution of adjectival secondary predicates makes use of certain common assumptions about the calculus of telicity. I will not go into details about how telic interpretations arise; however, it is generally proposed that the construction of a durative telic event (i.e. an accomplishment like in (8)) requires both the presence of an incremental structure and an upper bound to this structure (cf. Krifka (1989, 1998), Rothstein (2004), Kratzer (2004), among very many others). The fact that the simple transitive verb phrases (VPs) in (7) are atelic strongly suggests that it is the total adjective that is providing the upper bound that is required to create the telic interpretation. Furthermore, the fact

---

5 For reasons that are still mysterious to linguists, Italian requires doubling (or some modification) of a resultative adjective in order for the construction to be grammatical, see Folli and Ramchand (2005) for discussion.
that AA$^T$s alone can create such atelic/telic alternations suggests that only these adjectives have the required upper bounds to their scales.

Finally, similar observations about the relationship between AA$^T$s and telicity in degree achievements (causative verbs formed from scalar adjectives: ex. to lengthen, to straighten) are made by Hay et al. (1999) and Kennedy and Levin (2008) (among others). Although the exact patterns that show this link are too complicated to reproduce succinctly here, these authors argue that degree achievement verbs formed from adjectives like straight and empty (i.e. to straighten/empty) are generally telic, whereas the corresponding verbs formed from adjectives like long and wet (i.e. to lengthen/wet) are generally atelic.

In sum, the link between AA$^T$s and telic VP interpretations constitutes a strong empirical argument that these (and only these) adjectives are associated with scales that have maximal endpoints.

In the literature on scale structure, it turns out that there are many fewer tests for the presence of a minimal element than for the presence of a maximal element. One such test involves the distribution of the modifiers slightly and a little.$^6$ These modifiers can combine with a scalar adjective of any class; however, with AA$^P$s, they can receive an additional interpretation that is impossible with both AA$^T$ and RAs (see Solt (2012) for an in-depth discussion). With all scalar adjectives, slightly or a little can have an ‘excessive’ interpretation: the degree to which the property holds of the subject exceeds our expectations.

(14) **Relative adjectives**
    a. John is slightly/a little tall (for his age).
    b. He’s a little friendly $\approx$ He’s a little too friendly.

(15) **Total adjectives**
    a. The bar is slightly/a little empty/full (for my taste).
    b. John is slightly/a little bald (for me).

(16) **Partial adjectives**
    a. This towel is slightly/a little wet (for me to use).
    b. Your dress is slightly/a little dirty (to wear outside).

However, partial adjectives with slightly/a little can also have an existential interpretation: the sentences in (17) can be said if there is some amount of wetness on the towel or some amount of dirt on your dress, even if this amount does not exceed our expectations.

$^6$ For others, see Burnett (2012a).
(17)  a. This towel is slightly/a little wet.
    ≈ There is some wetness on the towel.

    b. Your dress is slightly/a little dirty.
    ≈ There is some dirt on your dress.

Similar observations are made for French *un peu* 'slightly/a little' by Martin (1969).\(^7\)

(18)  a. Jean est un peu bête.
    Jean is a little stupid
    Only 'John is a little (too) stupid.'

    b. La boîte est un peu vide.
    The nightclub is a little empty.
    Only 'The nightclub is a little (too) empty.'

    c. Ta robe est un peu sale.
    Your dress is a little dirty.
    'Your dress is a little (too) dirty,' or
    'Your dress has some dirt on it.'

One appealing explanation for the interpretative patterns shown above is that *slightly* or *a little/un peu* pick out the set of individuals that lie on an adjective's scale higher than a particular standard. The standard can be given by the context, or it can be set at the bottom endpoint of the scale. Since AA\(^P\)s are the only adjectives with bottom endpoints, they are the only ones that can give *slightly/a little/un peu* an existential interpretation. In other words, it is desirable for our theory to be able to associate scales with bottom endpoints with AA\(^P\)s.

Thus, I propose (following many authors) that AA\(^T\)s are associated with scales with maximal endpoints, AA\(^P\)s are associated with scales with minimal endpoints, and RAs are associated with open scales: scales that have no endpoints. These empirical patterns are summarized in Table 5.1.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>RA</th>
<th>AA(^T)</th>
<th>AA(^P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximal element?</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Minimal element?</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

\(^7\) I thank Francis Corblin for bringing this to my attention.
5.2.1 Non-scalar adjectives

When they are used precisely, NSs are not associated with any non-trivial scale; that’s why (I argued) they sound strange in the comparative. However, when we loosen our standards of precision, we can turn these predicates into scalar ones. So now we might investigate what the properties of these coerced scales are.

In what follows, note that the judgments presented below are either my own or taken from other sources whose judgments I share. Whether or not a particular adjective can be coerced and how it can be coerced depends on both the context and how willing the speaker is to apply the predicate in an unusual way in that context. So, a certain amount of contextual and speaker variation is expected. For example, as discussed in Chapter 4, some participants in Armstrong et al.’s (1983) experiment could associate a scale with the predicate even. I find scalar coercion of even extremely difficult to do and find a comparative like (19) (based on the results of the Armstrong et al. (1983) study) incoherent.

(19) 4 is more even than 18.

Contextual and speaker variation is expected under my analysis of scalar coercion as the overriding of a pragmatic precision constraint; nevertheless, even though such variation exists, I believe we can make some generalizations about how coerced NSs can differ from “true” AAs.

For some adjectives, coercion creates a AA^T: this is what we see with coerced hexagonal.

(20)  Top endpoint tests
   a. France is almost hexagonal.
   b. *This shape is slightly hexagonal.

Pregnant, on the other hand, is more easily coerced into a partial adjective.

(21)  a. *Mary is almost pregnant.
   b. Mary is slightly pregnant.
       (She’s showing, but not very much.)

However, many coerced NSs seem to be able to be associated with both a scale with a maximal element and a scale with a minimal element. As shown in (22) and (23), dead is one of these adjectives.
(22) a. DEA agent 1: So bring me up to speed on Tuco Salomar.
   DEA agent 2: Dead.
   DEA agent 1: Still?
   DEA agent 2: Completely.
   (Breaking Bad. Season 2, episode 5, “Breakage.”)
b. The coma patient is almost dead.
c. Loosely speaking, people in vegetative states are dead.

(23) a. ‘Dead Person is Actually Only Slightly Dead.’
b. Strictly speaking, Zombie John is dead, but he’s still chasing after us to eat our brains.

Nationality terms like Canadian are also adjectives that permit both maximal endpoint and minimal endpoint scales:

(24) a. Erin is slightly Canadian. (She is 1/8th Canadian)
   From http://sluttymuffins.blogspot.com/2010/03/fun-fact-erin-is-slightly-canadian.html
   See also the 1949 film Slightly French (http://www.imdb.com/title/tt0041885/).
b. Naoko is completely/almost Canadian.
   (She has completely/almost gone through the citizenship process.)

Finally, we can see the same pattern with illegal:

(25) a. Your accountant’s tax practices are almost/completely illegal.
b. Your accountant’s tax practices are slightly/a little illegal.

In summary, although there is some variation, with many NSs it is relatively easy to coerce them into both partial and AA\(^T\)s.

5.2.2 Summary of scale structure data

In this section, I have argued for the following scale structure patterns (shown in Table 5.2, where the diamond indicates coercion): RAs are associated with scales with no endpoints; AA\(^T\)s are associated with scales with a top endpoint; AA\(^P\)s are associated with scales with a bottom endpoint; and NSs, when they are coerced, are associated with scales with a top endpoint, a bottom endpoint, or both.
A certain caveat is in order concerning Table 5.2. Although $AA^T$'s like *straight* and *dry* are associated with scales that have no bottom endpoint (I will call these predicates properly $AA^T$), it has been observed that some other $AA^T$'s also pass the tests for having a scale with a bottom element. As shown in (26) and (27), adjectives such as *closed* and *open* are compatible with both *almost* and the existential interpretation of *slightly*.

(26) a. The door is *almost* closed.
   b. The door is *almost* open.

(27) Examples from Solt (2012)
   a. He'd lean his head back, his eyes *slightly* closed...
   b. Helene...had been in the bathroom, door cracked *slightly* open, peaking out through the small gap.
      (Analog Science Fiction & Fact, 122:10, p. 108)

Therefore, it is common in the literature (ex. Kennedy and McNally, 2005; Kennedy, 2007; Toledo and Sassoon, 2011; Lassiter, 2011 among others) to propose that such adjectives, which I will call *fully closed scale adjectives*, are associated with scales that have both a top and bottom endpoint. I will make some remarks concerning these predicates below; however, in Section 5.3, I show that the patterns in Table 5.2 are exactly predicted by the analysis in Chapter 4.

### 5.3 Scale structure in Delineation semantics

In the delineation approach to gradable predicates, “degrees” on a scale are equivalence classes of individuals that are related by the $\approx_p$ or $\approx_{p/s}$ relations (cf. Cresswell, 1976; van Rooij, 2011c; Bale, 2011, among others). Since, in this framework, degrees are equivalence classes of individuals $([x]_{\approx_p})$, if the domain $D$ is finite, then the scales associated with all adjectival predicates will have endpoints and the number of distinct
degrees will be necessarily limited by the cardinality of $D$. Thus, by definition, all scales over finite domains (be they associated with RAs, AA$^P$s, or AA$^T$s) must have both a top element and a bottom element. So how can we account for the open, top-closed and bottom-closed distinction in the framework developed here?

One way of expressing the “infinite” nature of open and partially closed scales while looking only at finite domains is to think about how a scale in a particular domain associated with a scalar adjective $P$ might be extended, should we add in other individuals. If we extend the scale associated with $P$ to include such individuals, some kinds of extensions may be blocked by the semantic or pragmatic axioms that $P$ obeys. As we will see, AAs will allow only a subset of the possible extensions that RAs allow. Thus, we can provide delineation-compatible definitions of top-closed scales (i.e. scales with maximal elements), bottom-closed scales (i.e. scales with minimal elements) and open scales as in the following definitions: 

**Definition 5.3.1 Top-closed scale.** For a predicate $P$ in a model $M$, $>_p$ is a top-closed scale iff for all extensions of $M$, $M'$, there is no $a_1 \in D_{M'} - D_M$ such that $a_1 >_p a_2$ in $M'$, for $a_2 : \neg \exists a_3 : a_3 >_p a_2$ in $M$.

In other words, we’ll say that a scale in a model is top-closed just in case its maximal elements remain maximal under all extensions of the model.

**Definition 5.3.2 Bottom-closed scale.** For a predicate $P$ in a model $M$, $>_p$ is a bottom-closed scale iff for all extensions of $M$, $M'$, there is no $a_1 \in D_{M'} - D_M$ such that $a_2 >_p a_1$ in $M'$, for $a_2 : \neg \exists a_3 : a_2 >_p a_3$ in $M$.

Thus, we will say that a scale in a model is bottom-closed just in case its minimal elements remain minimal under all extensions of the model.

**Definition 5.3.3 Open scale.** For a predicate $P$ in a model $M$, $>_p$ is an open scale iff $>_p$ is neither top-closed nor bottom-closed in $M$.

---

8 I thank Denis Bonnay for suggesting this strategy to me.

9 A more general statement of the top-closed (and, in fact, of all the scale structure properties) would be the existential statement in 29. The definitions are given for the $>_p$ relations, but the top-closed/bottom-closed/open properties can be defined from the $>_t$ and $>_s$ relations in a parallel way.

(ii) **Top-closed scale.** For a predicate $P$, $>_p$ is a top-closed scale iff there is some model $M$ such that, for all extensions of $M$, $M'$, there is no $x \in D_{M'} - D_M$ such that $x >_p d$ in $M'$, for $d : \neg \exists d' : d' >_p d$ in $M$.

However, it is easy to show that, given the constraints imposed on comparison classes (CCs) in my system (i.e. the absolute adjective axiom (AAA)), 29 and Definition 5.3.1 are equivalent. Thus, I state all the definitions of scale structure properties as universals, since I find them to be more intuitive in the context of the proposals made in this book.
Finally, a scale will be open in a model just in case some extensions allow for new maximal members, and some extensions allow for new minimal members.

We have already seen a first set of results concerning scalarity (i.e. whether or not an adjective is associated with a non-trivial scale) in Chapter 4. These results (which, I argued, are borne out in the data) are repeated in Table 5.3.

Using Definitions 5.3.1–5.3.3, we can show a further series of results concerning the properties of those non-trivial scales in Table 5.3. Firstly, we can prove that being associated with scales that have endpoints is a consequence of membership in the absolute adjective class. We can note that the top endpoint of a predicate’s tolerant scale is the predicate’s semantic denotation.

**Lemma 5.3.4** Total top endpoint. If $Q$ is an absolute adjective (i.e. satisfies the AAA), all models $M$, and $a_2 \in D$,

- If $\|Q(a_2)\|_D = 1$ then there is no $a_3 \in D$ such that $a_3 >^1_{Q} a_2$.

In other words, we predict (correctly) that the elements that are at the top endpoint of the empty/straight/clean scale are those that are completely empty/straight/clean, since those are the individuals that were proposed to be in the predicate’s semantic denotation. Now we

---

**Proof.** ⇒ Let $\|Q(a_2)\|_D = 1$ and suppose there is some $a_3$ such that $a_3 >^1_{Q} a_2$. Then, there is some $X \subseteq D$ such that $\|Q(a_1)\|_X = 1$ and $\|Q(a_3)\|_X = 0$. But $\|Q(a_2)\|_D = 1$, so by the AAA, $\|Q(a_3)\|_X = 1$. \(\perp\) Suppose there is no $a_3 \in D$ such that $a_3 >^1_{Q} a_2$ and suppose for a contradiction that $\|Q(a_2)\|_D = 1$. Since $>^1_{Q}$ is non-empty, there is some $a_4$: $\|Q(a_4)\|_X = 1$, for some $X \subseteq D$. By the AAA, $\|Q(a_4)\|_D = 1$. So $a_4 >^1_{Q} a_2$. Now consider the CC \{a_2, a_4\}. By the AAA, $\|Q(a_2)\|_{\{a_2, a_4\}} = 0$ and $\|Q(a_4)\|_{\{a_2, a_4\}} = 1$. By Minimal Difference (MD), $a_4 >^0_{Q} a_2$. So $\|Q(a_2)\|_{\{a_2, a_4\}} = 0$ and $\|Q(a_4)\|_{\{a_2, a_4\}} = 1$, so $a_4 >^1_{Q} a_2$. \(\perp\)
show that, given the fact in Lemma 5.3.4, an AA's tolerant scale ($>^t_Q$) is top closed (i.e. has a maximal element).

**Theorem 5.3.5**  If $Q$ is an AA$^T$, then $>^t_Q$ is a top-closed scale.\(^{11}\)

Secondly, we can show that the anti-extension of an absolute adjective is the bottom endpoint of its strict scale.

**Lemma 5.3.6** Partial bottom endpoint. For all $Q \in AA$, all models $M$, and $d, d' \in D$,

- If $d \notin \llbracket Q \rrbracket_D$, then there is no $d' \in D$ such that $d >^s_Q d'$.\(^{12}\)

In other words, the analysis correctly predicts that the minimal element of the non-trivial scale associated with a AA$^p$'s like dirty/wet/bent consists of those individuals that are not at all dirty/wet/bent, since those are the members of the predicate's semantic anti-extension. Correspondingly, we can show that an AAs strict scale is a bottom-closed scale:

**Theorem 5.3.7** If $Q \in AA$, then $>^s_Q$ is a bottom-closed scale.\(^{13}\)

In other words, based on Lemmas 5.3.4 and 5.3.6, we predict that the scales associated with total adjectives end at the same point where the scales associated with the scales associated with total adjectives. This result is schematized in Figure 5.1.

This result replicates exactly a proposal made by Rotstein and Winter (2004, p. 260) to account for the scale structure patterns discussed in this section. We can note however that, while this coincidence between the endpoints of partial and total scales is part of Rotstein and Winter's (2004) main proposal, it is a consequence of the analysis of context sensitivity and potential vagueness patterns developed in this work.

---

\(^{11}\) Proof. Let $M$ be a model and let $Q(a_1) \models_D = 1$. Therefore, by Lemma 5.3.4, there is no $a_2 \in D_M$ such that $a_2 >^t_Q a_1$. Now consider the extension of $M$, $M'$, such that $D_{M'} = D_M \cup \{a_1\}$. Show $a_3 \npreceq^t_Q a_1$. Suppose for a contradiction that $a_3 >^s_Q a_1$. Then there is some $X \subseteq D_{M'}$ such that $\llbracket Q(a_1) \rrbracket_X^t = 1$ and $\llbracket Q(a_1) \rrbracket_X^s = 0$. But, since $M'$ is an extension of $M$, $\llbracket Q(a_1) \rrbracket_D^t = 1$. So, by the AAA, $\llbracket Q(a_1) \rrbracket_X^s = 1$ and $\llbracket Q(a_1) \rrbracket_X^t = 1$. \(\bot\)

\(^{12}\) The proceeds by duality with the proof of Lemma 5.3.4.

\(^{13}\) Proof. Let $M$ be a $t$-model and let $x \notin \llbracket Q \rrbracket_D$. Therefore, by Lemma 5.3.6, there is no $y \in D_M$ such that $x >^s_Q y$. Now consider the $t$-model $M'$ such that $D_{M'} = D_M \cup \{z\}$. Show $x \succ^t_Q z$. Suppose for a contradiction that $x >^s_Q z$. Then there is some $X \subseteq C_{M'}$ such that $x \notin \llbracket Q \rrbracket_X^s$. But since $x \notin \llbracket Q \rrbracket_D$ in $M$ and $M'$ extends $M$, $x \notin \llbracket Q \rrbracket_D$ in $M'$ and, by the AAA, $x \notin \llbracket Q \rrbracket_X^s$. \(\bot\)
While AAs are subject to the AAA, which imposes very strong constraints on the semantic denotations of absolute constituents, I proposed in Chapter 4 that RAs were only subject to van Benthem’s axioms (No Reversal (NR), Upward Difference (UD), and Downward Difference (DD)). These conditions are very weak (they just ensure a strict, weak ordering), and, as such, very many more models will be models for RAs than for AAs. Thus, the scales built from the semantic denotations of RAs (\(>_{ps}\)), which are the only non-trivial strict, weak orders associated with these predicates, will permit extensions where their maximal and minimal elements do not remain maximal/minimal. This is unlike with AAs, where we saw that maximal elements on tolerant scales must remain maximal and minimal elements on strict scales must remain minimal. In other words, relative semantic scales are open scales: they have no non-accidental endpoints.

**Theorem 5.3.8**  
If \(P\) is an RA, then \(>_{p}\) is an open scale.\(^{14}\)

We therefore correctly predict that RAs should pass neither the tests for having a maximal element nor the tests for having a minimal element. Finally, note that, since coerced NSs are simply AAs that can have both non-trivial tolerant and strict scales, we predict that these predicates should be able to have both a non-trivial top-closed scale and a non-trivial bottom-closed scale.

In summary, in this section, we saw that the appropriate association of scales of particular types with particular types of adjectives is a consequence of the analysis of the context sensitivity and potential

\(^{14}\) **Proof.** (Not top closed:) Let \(M\) be a model and let \(a_1 \in D_M\) such that there is no \(a_2 \in D_M\) such that \(a_2 >_p a_1\). Now consider the proper extension, \(M'\), such that \(D_{M'} = D_M \cup \{a_1\}\). Suppose that \(\llbracket P(a_1) \rrbracket_{(x,y)} = 1\) and \(\llbracket P(a_1) \rrbracket_{(x,y)} = 0\). This is permitted (provided \(P\) still satisfies NR, UD, and DD) because \(\llbracket P\rrbracket\) can vary across CCs. So \(a_1 >_p a_1\).

(Not bottom closed:) Let \(M\) be a model and let \(a_1 \in D_M\) such that there is no \(a_2 \in D_M\) such that \(a_1 >_p a_2\). Now consider the proper extension of \(M, M'\), such that \(D_{M'} = D_M \cup \{a_3\}\). Suppose \(\llbracket P(a_1) \rrbracket_{(x,y)} = 1\) and \(\llbracket P(a_3) \rrbracket_{(x,y)} = 0\). So \(a_1 >_p a_3\).
vagueness of scalar and NSs that was presented in Chapter 4. In other words, given the analysis presented in Chapter 4, there is no need to stipulate that an adjective like clean is associated with a top-closed scale; all that we require to see this fact is an appropriate definition of what it means to be a top-closed scale within a delineation framework.

It was suggested in the first part of this chapter that there may be reasons to think that the class of AA’s should be divided into two subclasses: “proper” total adjectives like dry and straight, which are only compatible with modifiers sensitive to scales with top endpoints, and “fully closed scale” adjectives, like open and closed, which are compatible with both top-closed and bottom-closed endpoints. It is easy to show that, based on the architecture of the framework, it is impossible for an adjective to be associated with an articulated scale that has both a top and bottom endpoint. Nevertheless, I suggest that we can still arrive at an appropriate analysis of this limited class of predicates within a delineation system. One possibility is to propose that some fully closed-scale absolute adjectives are subject to neither the Total nor the Partial axioms that create asymmetries between absolute predicates and their negations. This would have the consequence that these predicates could be associated with two non-trivial scales: a top-closed tolerant scale and a bottom-closed strict scale. This analysis is particularly appealing for predicates like open, where it is possible to find a single basic semantic denotation (e.g. the set of objects having some degree of aperture) that would constitute the top endpoint of a tolerant scale and whose complement would constitute the bottom endpoint of a strict scale (28).

(28)  a. The door is almost open. ⇒ The door is not open.
      b. The door is slightly open. ⇒ The door is open.

Another possibility is to propose that certain other fully closed-scale adjectives are in fact ambiguous between a total (i.e. universal) version and a partial (i.e. existential) version. Such an analysis is appealing for predicates like closed, where the semantic denotation of the total predicate closed₁ would pick out those objects that are completely closed and be associated with an articulated top-closed tolerant scale (29-a) and the semantic denotation of the partial predicate closed₂ would pick out those objects that have some amount of closure and be associated with an articulated bottom-closed strict scale (29-b).\(^{15}\)

\(^{15}\) Note that in such an analysis we predict that antonyms empty and full would not be directly associated with “the same scale” as in, for example, DegS analyses of such cases of
However, I leave the question of how one or both of these styles of analyses should be applied to the complete list of “fully closed-scale” adjectives to future research.

5.4 Conclusion

In this chapter, I have shown that the scale structure patterns that have been previously observed and argued for in the literature are predicted by the analysis that I gave of the context sensitivity and potential vagueness patterns of RAs, AAs, and NSs, in the previous parts of the monograph. Furthermore, in this chapter and in Chapter 4, we have seen that the theory developed in this work makes predictions about the basic scalarity of adjectives. Thus, while the proposals that I made in Chapter 4 concern how to properly analyze the context dependence and the vagueness of RAs, AAs, and NSs, they have straightforward implications for their gradability. For each adjective, two aspects of its meaning are predicted:

1. Whether or not the adjective is scalar.
2. If the adjective is scalar, whether its scale has maximal elements, minimal elements, both or neither.

I therefore conclude that, from a simple and independently necessary theory of context sensitivity and vagueness, we can arrive at a full theory of gradability and scale structure in the adjectival domain. In Chapter 6, I give a comparison between the properties and predictions of my theory and some other existing theories of adjectival scale structure in the literature within the DegS framework, and, in Chapter 7 of this work, I explore extending DeTCS outside the adjectival domain to show how such an extension gives us new insight to old puzzles concerning distributivity, existential and definite plural determiner phrases (DPs).

antonymy (Kennedy, 1997). However, in the framework presented here, this property is not limited to pairs of fully closed-scale adjectives; rather, since scales are derived from cross-contextual variation associated with predicates, no two distinct predicates are associated with exactly the same scale. As such, in the DeTCS framework, relations between synonyms and antonyms must be established through meaning postulates (see Burnett, 2012a).
6

Beyond Delineation semantics

6.1 Introduction

This chapter explores how the enrichment of non-Delineation theories of the semantics of gradable expressions with the structure of a non-classical logic such as Tolerant, Classical, Strict (TCS) can be useful to understanding the complex relationships between context sensitivity, vagueness, and scale structure that were discussed in the first part of the book (Chapters 2–5). An important conclusion of the previous chapters is that the use of TCS (or some system like it) to analyze the vague characteristics of both relative scalar adjectives (RAs) and absolute scalar adjectives (AAs) has the potential to solve certain empirical puzzles associated with the analysis of the absolute/relative distinction that face many versions of a simpler Delineation semantics (DelS) system. Thus, I argued that the TCS extension of DelS developed in this work has important consequences for the empirical coverage of DelS as a general framework for the semantic/pragmatic analysis of gradability in natural language.

In this chapter, I examine whether the usefulness of TCS is limited to DelS or whether there is a natural TCS extension of other current frameworks that can yield an interesting account of the data described in Chapter 5. In this way, one of the aims of this chapter is to serve as a comparison between the analysis of the absolute/relative distinction developed in Chapters 1–5 and other influential analyses in the literature set within different frameworks.

This being said, I suggest that finding the most illuminating way to compare different analyses is not a trivial matter. Both gradability and vagueness are enormous areas of linguistic syntax/semantics/pragmatics and analytical philosophy, and research into these phenomena has generated a wide range of accounts set in frameworks that, in many cases, have drastically different underlying assumptions and characterizations of the syntax, semantic, and pragmatic interfaces. For example, a characterizing property of the proposal developed in
this work is that it is set within DelS. However, many (if not most) of its possible competitors adopt some other analysis of the scales associated with scalar constituents, for example a semantics based on degrees \((\text{DegS})\), to be further discussed in Section 6.2, see also Bartsch and Vennemann, 1973; Cresswell, 1976; Heim, 1985; Bierwisch, 1989; Kennedy, 1997, among many many others or a semantics based on tropes (Moltmann, 2009). Some proposals (particularly those set in the DegS framework) involved a highly articulated theory of the syntax–semantics interface; that is, how the scales associated with predicates are accessed by certain morphosyntactic constructions. Other theories (particularly those set within DelS, like the proposal developed in this account) focus more on inferential relations between full sentences and make no particular detailed claims concerning how exactly the sentences are constructed. Additionally, some proposals (such as DelTCS) are set in static frameworks; that is, frameworks that encode no notion of temporal progression. Other proposals, such as Barker (2002) and recent work on the interpretation of adjectives within Bayesian pragmatics (such as Potts, 2008; Franke, 2012; Qing and Franke, 2014; Lassiter and Goodman, 2013, 2015, among others) are dynamic: important features of meanings of predicates are created through updating beliefs and contexts from one time to another. The Bayesian theories are also different from many other existing proposals in that they are probabilistic: the use and interpretation of linguistic expressions is created through rules with probabilities assigned to them. DelTCS (like many other accounts) is a deterministic theory: the interpretation of an expression is categorically determined by grammatical and interpretive rules. Furthermore, as discussed in Chapter 2, when it comes to the analysis of vagueness, some theories (like DelTCS) adopt a non-classical logical approach; whereas, many other theories preserve the properties of classical logic. Finally, proposals also differ in how they view the relationship between the literal truth conditional meaning of linguistic expressions (i.e. semantics) and the meanings of expressions in context (i.e. pragmatics). In DelTCS and in other highly contextualist theories

---

1 Note that there has been some work within DelS on the semantics of modifiers (by, for example, Klein, 1986) and other more complicated syntactic constructions such as differentials (cf. McConnell-Ginet, 1973); however, the syntactic realization of scales and scale structure is an area that is largely unexplored within DelS theories, particularly when we compare them to their DegS siblings. This is an area in which much future work would be desirable.

2 Other recent probabilistic proposals for the semantics of vague expressions are Lassiter (2011) and Sutton (2013).
such as Récanati (2010), aspects of an expression’s use in context can contribute to the calculation of the literal meaning of expressions. For example, in DelTCS, the contextually determined indifference relations between individuals associated with AAs are crucial in the calculation of the context-independent meaning of comparatives formed with these kinds of predicates. In other frameworks, such as Williamson’s (1994) Epistemicism or Kennedy’s (2007) DegS theory (to be discussed in section 6.2), there is a very clear separation between semantic and pragmatic aspects of meaning.

In summary, many of the existing analyses of vagueness and scale structure in the literature start from different sets of basic assumptions, which can vary along (at least) the parameters shown in (1). As such, directly comparing theories that differ along more than one dimension in (1) can often be difficult.

(1) Theoretical profile parameters
   a. Scalar source: DelS vs DegS vs Tropes vs ...
   b. Semantics–Pragmatics Interface: Interactive vs separate
   c. Logical Foundations: Classical vs Non-classical
      • If non-classical, what kind of non-classical?
   d. Dynamnicity: Static vs dynamic
   e. Variability: Deterministic vs probabilistic

DelTCS therefore has the profile shown in Table 6.1.

An additional complication is that, even if different frameworks do share common assumptions, teasing apart their empirical predictions can often be difficult because of their respective modes of presentation. This issue is particularly salient with the study of vagueness, which draws in work from both logical traditions and the philosophy of language more broadly. Work on the logic of vagueness (such as TCS and, indeed, this work) is often highly formal, involving much notation and elaborate proofs. Proposals coming from the philosophy of language tradition, on the other hand, are often stated in plain language. This certainly increases readability, but such proposals are often compatible with

<table>
<thead>
<tr>
<th>Theory</th>
<th>Scales</th>
<th>Semantics–Pragmatics Interface</th>
<th>Logic</th>
<th>Dynamicity</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>DelTCS</td>
<td>Delineation</td>
<td>Interactive</td>
<td>Multi-valued</td>
<td>Static</td>
<td>Deterministic</td>
</tr>
</tbody>
</table>
Thus, to make true and complete comparisons across proposals laid out with different levels of formality, we would have to look at all the reasonable formalizations.

Of course, going in and carefully establishing the existence of (non-)equivalences across frameworks with very different forms would constitute a whole other book in itself. Therefore, in this chapter, we will pick one proposal that is (relatively) close to DelTCS, and I will translate it into the notation and general set-up of DelTCS. This will allow us to see, point by point, where the accounts differ, where they are the same, and what the empirical consequences of these similarities and differences are. The proposal I will focus on is Kennedy’s account, set within Degree semantics (DegS); however, I believe it to be an interesting and worthwhile project to use extensions of this method with other kinds of analyses set in other kinds of frameworks.

The chapter is therefore set out as follows: in Section 6.2, I present Kennedy’s analysis of the relationship between context sensitivity, vagueness, and scale structure set within DegS, and I show how certain key aspects of his proposal, once translated into the notation adopted in this book, look very similar to proposals that I made in the previous chapters. I suggest that these similarities open the door for a larger exploration of the use of a TCS-style framework outside of DelS. In Section 6.3, I examine in greater detail one of the most discussed components of Kennedy’s theory: the Interpretive Economy (IE) axiom relating scale structure and the use of the positive form across contexts. I outline its original formulation within static deterministic DegS and also its reconceptualization within recent Bayesian pragmatics approaches to adjectival meaning. With these proposals in mind, in Section 6.4, I give a DegS extension of TCS (called DegTCS), based on the proposals in Kennedy (2007). I show how within this new framework we can prove (an appropriate version of) Interpretative Economy as a theorem. I therefore conclude that the general structure of TCS can be useful in the analysis of gradability and vagueness outside of the DelS framework.

6.2 Scale structure in Degree semantics

This section presents a detailed discussion of Kennedy’s (2007) analysis of the relationship between context sensitivity, vagueness/imprecision,
and scale structure. This particular DegS analysis is not the only one available in the literature; however, I choose to focus on this proposal because it both builds on previous work (such as Kennedy and McNally, 2005) and, like DelTCS, it specifically aims to account for the interaction between vagueness and scale structure (unlike, for example, Rotstein and Winter (2004), which does not mention vagueness).

6.2.1 Kennedy’s (2007) Degree analysis

As in many other works on the syntax and semantics of gradable predicates (such as Bartsch and Vennemann, 1973; Seuren, 1973; Bierwisch, 1989; Cresswell, 1976; von Stechow, 1984; Heim, 1985; Kennedy, 1997; Kennedy and McNally, 2005; Rett, 2008, 2014, among others), Kennedy assumes that gradable adjectives map their arguments onto abstract representations of measurement called *degrees*. That is, the ontology adopted by researchers in the DegS framework is proposed to contain (along with individuals, truth values, events, etc.) individuals of the type *degree* ($d$). Furthermore, the set of degrees is assumed to be totally ordered with respect to some dimension $D$ (height, cost, colour etc.). The structure consisting of a set of degrees and a total ordering on those degrees according to a dimension is called a *scale*. Here is a more formal definition:

**Definition 6.2.1** Let $\text{Dim} = \{\delta_1, \ldots, \delta_n\}$ be a finite set of dimensions. Then for all $\delta_i \in \text{Dim}$, the *scale* associated with $\delta_i$ is a tuple $(D, >_{\delta_i})$, where $D = \{d_1, d_2, \ldots\}$ is a possibly infinite set of degrees and $>_{\delta_i}$ is a total ordering on $D$.

In Kennedy (2007), gradable adjectives, such as *tall*, are analyzed as measure functions: functions from subsets of the domain of individuals that have some value on a dimension, to degrees on the scale lexically associated with the adjective. They can equivalently be represented as relations between individuals and degrees on the adjective’s scale. For example, the denotation of an adjective like *expensive* is given in functional notation in (2-a) and in relational notation in (2-b).

(2) "

a. $\lambda d \lambda x. \ x$ is an individual that has some cost and $d$ is $x$’s degree on the *cost* scale.

b. $\{(a_i, d_i) : a_i$ is an individual that has some cost and $d_i$ is $a_i$’s degree on the cost scale}\.
With the relational/measure-function view of the basic denotation ofgradable adjectives, the semantics of the comparative construction is quite straightforward: more P/P-er for an adjective P, establishes an ordering between individuals based on the degrees that they are mapped to by P. More specifically,

(3) \[[x \text{ is more expensive than } y] = 1 \text{ iff } \|\text{expensive}\| (x) > \text{cost} \|\text{expensive}\| (y)\]

Since adjectives are analyzed as measure functions/binary relations in the lexicon, we need a more complicated analysis of the “intransitive” use of the adjective in the positive construction (i.e. This car is expensive.). Following Bartsch and Vennemann, 1973; Cresswell, 1976; Kennedy, 1997, among others, Kennedy proposes that adjectives combine with a null morpheme, called pos, which restricts the denotation of the predicate to those objects that “stand out’ in the context of utterance, relative to the kind of measurement that the adjective encodes.” (Kennedy, 2007, p.17). The way in which it is ensured that the denotation of the positive form of expensive is restricted in the appropriate way is through a “standard” function s that is referred to in the denotation of pos (4) and is a “context sensitive function from measure functions to degrees that returns a standard of comparison based both on properties of the adjective g (such as its domain) and on features of the context of utterance.” (Kennedy, 2007, p. 16).

(4) \([\text{pos}] = \lambda g \lambda x. g(x) \geq s(g)\)

The s function is similar to both the delineation functions used by Lewis (1970) and Barker (2002), and the basic comparison-class (CC) relativized denotations adopted in Chapter 4, which serve to delineate the denotation of the positive form of the adjective in a CC. An important difference is that Kennedy’s s requires that the extension of a positive adjective in a context “stand out” from its anti-extension in that context (see Kennedy, 2007, fn. 14 for discussion of differences between his proposal and Lewis/Barker’s). Kennedy also proposes that s is subject to a further constraint associated with how it can partition the salient individuals in a context. Following Soames (1999), he adopts a “metalinguistic” principle that ensures that individuals that are viewed as indifferent with respect to a predicate in the context do not find themselves on opposite sides of the boundary created by s. He says (p. 19),
We are unwilling to commit to the position that one or two objects that differ minimally along the scalar continuum measured by \( g \) could stand out relative to \( g \) while the other one doesn't; as a result, whenever two such objects are evaluated together against the positive form of \( g \), we treat them the same.

According to Kennedy, it is this principle that makes us unwilling to reject the inductive premise of the Sorites argument for RAs because, by (6),

Two objects that differ very slightly on a graded continuum are judged the same relative to the positive form because treating them differently would involve a commitment to the unlikely position that such a small difference in degree could matter as to whether one stands out or not, relative to the measure on which the continuum is based. (Kennedy, 2007, p. 28)

In Kennedy (2007), the relationships between context sensitivity, vagueness, and scale structure discussed in Chapters 3 and 5 are derived through the interaction between scale structure, context sensitivity, and a series of principles including (5). More precisely, he proposes that adjectives can vary depending on the structure of the scale that they are associated with, i.e. whether their scale has a maximal endpoint, minimal endpoint, no endpoints, or two endpoints. According to Kennedy, the scalar endpoints provide what he calls, following Williamson (1992), natural transitions, which serve to create distinction between individuals at endpoints of the scale and those that are not. He says (p. 35),

What it means to stand out relative to the measure expressed by a closed scale adjective is to be on the upper end of a natural transition based on the scale: the transition from a non-maximal degree to a maximal one or the transition from a zero degree to a non-zero degree. Crucially, these transitions are inherent to a closed scale and to the kind of measure it represents, and so also to the meaning of a closed scale adjective.

In other words, for Kennedy, individuals at scalar endpoints cannot be indifferent from non-endpoint individuals in any context. Lacking endpoints, the scales associated with RAs have no “natural transitions,” and therefore there is no principled difference with respect to “standing out” between individuals at different degrees on an open scale.
Additionally, in order to capture the observation that the use of the positive form of an AA is largely limited to those individuals at the endpoints of the scale, Kennedy (p. 36) proposes a constraint, IE, stated as in (8).

\[(8) \quad \text{Interpretive Economy (IE):} \]

Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions.

This principle is very general and abstract; however, more concretely, Kennedy says that “the effect of Interpretive Economy on the positive form is to ensure that closed scale adjectives are absolute.” How exactly we go from the general statement in (8) to the more specific result of the constrained context sensitivity of AAs is not explored in detail in the 2007 paper; however, this idea is expanded upon in some recent work in Bayesian pragmatics, which will be discussed in Section 6.3.

Finally, as Kennedy notes, there are contexts in which an absolute predicate like empty may be more or less felicitously applied to an individual who is not, strictly speaking, at the top endpoint of the adjective’s scale (i.e. This nightclub is empty examples). In such cases, Kennedy suggests that a general pragmatic phenomenon of imprecision derives the appropriate interpretation.

6.2.2 Comparison between Kennedy (2007) and DelTCS

As presented in Section 6.2.1, the proposal in Kennedy (2007) appears quite different from the one developed in Chapters 2–5. The differences along the theoretical parameters discussed in the introduction are shown in Table 6.2.

That there are differences between Kennedy (2007) and DelTCS is certainly true of the form in which the proposals are laid out: important aspects of this DegS proposal are described in an intuitive and easily accessible manner; whereas, the emphasis in this book has been on the logical regimentation and formal results, which makes the shape of the

<table>
<thead>
<tr>
<th>Theory</th>
<th>Scales</th>
<th>Semantics–Pragmatics Interface</th>
<th>Logic</th>
<th>Dynamicity</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>DelTCS</td>
<td>Delineation</td>
<td>Interactive</td>
<td>Multi-valued</td>
<td>Static</td>
<td>Deterministic</td>
</tr>
<tr>
<td>Kennedy (2007)</td>
<td>Degree</td>
<td>Separate</td>
<td>Classical</td>
<td>Static</td>
<td>Deterministic</td>
</tr>
</tbody>
</table>
proposals look very different. But to what extent are the contents of the theories really different?

One easily observable difference between the two accounts is that Kennedy makes no explicit use of a paraconsistent logic in his semantics and explanation of the Sorites (see (5)); whereas, the TCS non-classical framework is an important part of the theory developed here. The use of this logic was motivated, on the one hand, by its easy analysis of borderline contradictions (9) (see Alxatib and Pelletier, 2011; Cobreros et al., 2012b; Alxatib et al., 2013; Cobreros et al., 2015, among others), and, on the other hand, its potential to provide a unified analysis of the context sensitivity of RAs and “loose” uses of AAs.

(9)  

a. Mary is neither tall nor not tall.

b. This glass (with a tiny bit of beer in it) is both empty and not empty.

This being said, we can observe that the statement in (5) is very similar to the definition of the tolerant extension of a predicate at a CC in DelTCS (and in TCS more generally): if two individuals are indifferent with respect to a predicate, and one is in the classical extension of the predicate at a CC, then the second is (at least) in the tolerant extension. In fact, these similarities can be made explicit if we translate (5) into the notation that we have adopted in the book, as (10).

(10)  

If \( a_1 \sim_p a_2 \) and \( a_1 \in \text{\{\text{pos} P\}} \), then \( a_2 \in \text{\{\text{pos} P\}} \).

b. Cobreros et al. (2012b)/this work:
If \( a_1 \sim_p a_2 \) and \( a_1 \in \text{\{\text{P}\}} \), then \( a_2 \in \text{\{\text{P}\}}^t \).

Thus, I suggest that the proposals do resemble each other when it comes to the attribution of the source of Sorites susceptibility, although the statement in (10-b) is presumably a bit weaker than (10-a).

Likewise, Kennedy makes use of the notion of natural transitions in his analysis of the use of the positive form of absolute predicates and the way in which AAs and RAs differ in how they display the characterizing properties of vague language. Natural transitions in his theory are a product of scale structure, where the existence of a top endpoint causes individuals at the endpoints to “stand out” from non-endpoint individuals. My theory makes no explicit reference to “natural transitions,” and does not attribute them directly to scale structure, particularly because scales are derived objects in this framework. So
this is another difference. However, once again, we can observe that the proposal that endpoint individuals are importantly different from non-endpoint individuals is quite similar (although not exactly identical) to the one that I made in Chapter 3 concerning the shape of indifference relations between members of the classical (anti-)extension of an absolute predicate. I proposed that members of the classical extension of a total predicate cannot be indifferent from members in its classical anti-extension. We saw in Chapter 5 that members of total classical extensions are just those members who are located at the top endpoint of the adjective’s scale; likewise, members of a total predicate’s classical anti-extension correspond to those individuals who are not at the top endpoint of the scale. So again we find a convergence between Kennedy’s proposal and DelTCS.

This being said, what I proposed is a bit weaker than (7): in DelTCS, we still allow for elements not at the top endpoint to be treated as indifferent from endpoint individuals in at least some contexts, and, therefore, we allow for a (restricted) amount of Sorites susceptibility with absolute predicates. In the interest of being explicit about the comparison between Kennedy (2007) and the theory presented in this book, I therefore formalize the constraints on indifference as shown in (11).

   Let \( Q \) be associated with a top-closed scale and suppose \( a_1, a_2 \in D \) such that \( \|Q\|(a_1) \) is \( d_{\text{max}} \) (the maximal element of \( Q \)'s scale) and \( d_{\text{max}} > Q \|Q\|(a_2) \). Then, there is no context \( C \) such that \( a_2 \sim_Q a_1 \) and \( a_1 \sim_Q a_2 \) in \( C \).

b. (This work):
   Let \( Q \) be associated with a top-closed scale and suppose \( a_1, a_2 \in D \) such that \( a_1 \in \|Q\|_D \) and \( a_2 \notin \|Q\|_D \). Then there is no \( X \subseteq D \) such that \( a_2 \sim_Q a_1 \) in \( X \).

So again, we see similarities between Kennedy (2007) and DelTCS, although with the proposals advocated in this work being again a bit weaker than in Kennedy (2007).

Another difference between Kennedy (2007) and this work is that, in this Degree analysis, context sensitivity of RAs and AAs have different sources: the context sensitivity of \( \text{pos} \) combined with lack of endpoints (in the case of RAs) and a general pragmatic imprecision (in the case of AAs). In my proposal, all classes of scalar adjectives vary their tolerant or strict denotations across contexts, which in the logic are instantiated
by CCs. Such an analysis is possible because this work adopts a very particular view of the relationship between literal meaning (i.e. lexical and compositional semantics) and meaning that arises from the use of an expression in some extra-linguistic context (i.e. what is often called pragmatics). In particular, we assume in this work that aspects of use in context can (and do) affect compositional semantics, the most pertinent example of which is the derivation of the scales associated with AAs from their “loose” uses. The proposal developed in this work therefore contributes to a number of current research programs that seek to unify certain aspects of what has traditionally been separated along the semantics–pragmatics border. The perspective adopted here most closely resembles the one found in the work of François Récanati (such as in Récanati, 2004 (Literal Meaning) and Récanati, 2010 (Truth Conditional Pragmatics)), but is consistent with much recent work in radical and moderate pragmatics (see Cappelen and Lepore, 2005, for discussion). Kennedy (2007) adopts a more traditional view of the semantics–pragmatics interface; therefore, as far as I can see, an account such as the one that I have provided for the gradability of AAs would not be possible in this framework. This being said, following Kennedy and McNally (2005), Kennedy suggests that one option for the analysis of the context sensitivity of AAs is through the use of Lasersohn’s (1999) Pragmatic Halos (PH) framework. We saw in Chapter 2 that there are many similarities between PH and TCS’s tolerant denotations. Thus, we can identify yet another similarity between the theory given in this book and the one described in Kennedy (2007).

A final (quite salient) difference between Kennedy (2007) and the present work concerns the analysis of the lexical semantics of gradable predicates. While my proposal is set within a Delineation approach to the semantics of scalar expressions, Kennedy uses DegS. As discussed in Section 6.2.1, the DegS analysis complemented with an IE principle that aims to restrict the use of the positive forms of absolute predicates. Both the use of DegS and IE constitute departures from my proposal, but the extent to which they are necessary departures is not so clear. This point will be further addressed in the next sections.

In summary, I have argued that the analysis of the RA/AA distinction presented in Chapters 2–5 and the one found in Kennedy (2007) have certain important differences. As I see it, the main ones are the following:

(12) a. The use of TCS (non-classical logic) to analyze the properties of vagueness (vs non-paraconsistent logic + contextualism/epistemicism).
b. The *Truth Conditional Pragmatics* approach to the semantics/pragmatics interface (vs clear separation between semantics and pragmatics).

c. The use of DelS to analyze the semantics of gradable predicates (vs DegS).

On the other hand, I have suggested that there are significant similarities between some of the proposals made in this work and some of the proposals in Kennedy (2007), similarities that are made apparent by the formalization and translation process described above. I believe that these similarities justify exploring to what extent we can formalize the correspondences between them.

In Section 6.4, I replace one of the assumptions in (12) with one from Kennedy (2007): DegS. I give a DegS extension of the TCS framework, and show how we can use it to capture (as theorems) some of the kinds of results that appear as axioms in Kennedy (2007). Some of the results given in Section 6.4 will be focussed on facts captured by the IE meta-principle (8), which is an axiom in Kennedy’s theory. For this reason, in Section 6.3, we take a closer look at IE and how it has been reinterpreted within Bayesian pragmatics.

### 6.3 Interpretive Economy and Bayesian pragmatics

Although the IE meta-principle is an important part of Kennedy’s analysis, as observed by Potts (2008), Chierchia (2010), Lassiter and Goodman (2013), its wording and the role it plays in Kennedy’s analysis give rise to questions concerning how it fits into a general theory of pragmatics. In this section, we review one approach that aims to derive a principle similar to IE from a broader theory of contextually based meaning enrichment: the account of RAs and AAs outlined in Lassiter and Goodman (2013, 2015), which is set within Gricean Bayesian pragmatics.  

#### 6.3.1 Bayesian pragmatics

The *Bayesian Pragmatics* framework (Franke, 2010; Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013, among many others) combines recent work in game-theoretic pragmatics with a Bayesian approach to inference under uncertainty (Pearl, 2000; Griffiths et al.,

---

4 See also Franke (2012), Qing and Franke (2014) for similar ideas.
5 See Benz et al. (2005) for an introduction to this field.
2008; Tenenbaum et al., 2011, among others). More precisely, this framework follows proposals by Lewis (1969) to the effect that it is enlightening to model linguistic communication as a signalling game, which is a game of pure co-ordination between two players: S, the sender/speaker, and R, the receiver/hearer. In this kind of game, S has a proposition $\phi$ (called their type, modelled here as a set of possible worlds) that they would like to communicate to R, and S's action is to choose a message, linguistic expression $m$, from the set of available messages, i.e. well-formed interpreted expressions in the language. R's action is to update their beliefs concerning the actual world based on three components:

1. The beliefs that they held prior to hearing what S said. These beliefs (called priors) are represented as a probability distribution $P(\cdot)$ over propositions (sets of possible worlds).
2. The semantic denotation of the expression that S used ($[m]$), which is a proposition.
3. R's beliefs concerning the strategies that S uses to pick the message $m$, given their beliefs about the way that the world is.

An important hypothesis adopted by many authors in this area (see Blutner, 2015, for a recent overview) is that the strategies of both the speaker and the hearer are guided by informativity; that is, the speaker aims to choose a message that is as informative as possible relative to the current topic of conversation (or Question under Discussion in the sense of Ginzburg, 1995a,b; Roberts, 1996). Furthermore, the hearer assumes that the speaker is also being as informative as possible. The speaker and hearer also factor in possible cognitive or social costs that using a message might have. These mutual beliefs of informativity are what allow the speaker and the hearer to co-ordinate on a (possibly pragmatically enriched) interpretation for a particular message.

To give a quick example, suppose we are interested in explaining why, in most cases, using a quantifier like some in a statement, like (13-a) strongly implies that the corresponding statement, where we substitute some for all, is false (13-b). As discussed in Grice (1975), the implicature that (13-b) is a puzzle for a naive semantic theory in which (13-a) is true just in case Sarah eats one or more (i.e., possibly all) of the cookies.

(13)  

- a. Sarah ate some of the cookies.
- b. Sarah ate all of the cookies.
The way the implicature is analyzed in the Bayesian pragmatics framework is as follows: the receiver/hearer R starts off with their prior beliefs concerning how many of the cookies have been eaten. As mentioned above, these beliefs are represented by a probability distribution over sets of possible worlds \( P(\cdot) \), i.e. the set of worlds in which Sarah eats no cookies are assigned probability \( p_i \), the set of worlds where she eats two of three cookies is assigned the probability \( p_j \), the set of worlds in which she eats all three of the cookies is assigned \( p_k \) etc. R then hears S say the sentence *Sarah ate some of the cookies*. In the first step in the interpretation process, R conditions their beliefs on the truth of *Sarah ate some of the cookies*; that is, they assign the probability 0 to the sets of the worlds in which Sarah eats no cookies, and then normalize the prior probability distribution over the worlds in which at least some cookies are eaten. This interpretation process, which Lassiter and Goodman call the *literal listener*, can be written more formally as in (14), where \( A \) is a set of possible worlds and \( u \) is an utterance/message.

\[
(14) \quad PL_0(A|u) = PL_0(A||u|| = 1)
\]

In their final calculation, R takes into account how exactly they believe S chooses the message. In particular, a crucial feature of this framework is that R has the belief that the speaker chooses the optimal message based on 1) the conditionalized prior \( PL_0 \), 2) informativity, and 3) message costs. For the full formal details of how speaker strategies (notated \( PS_1(\cdot) \)) are calculated, see (Frank and Goodman, 2012, among others); however, we can describe them informally as follows:

1. The conditionalized prior ensures that R believes that S would not say anything false (i.e. in this example, it assigns a zero probability to worlds in which Sarah eats no cookies).
2. The informativity constraint ensures that R thinks S’s type is the strongest meaning possible; that is, here is where the preference for the most informative interpretation is encoded. We can observe a similarity because the maximization of informativity in Bayesian pragmatics and the wording of IE: “Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions.”
3. Finally, the costs constraint serves to create a dispreference for the use of certain kinds of messages. For example, we might assume that *Sarah ate some but not all of the cookies* is more costly from
a processing and/or production point of view because it is longer than an expression like *Sarah ate some of the cookies*.

With optimized speaker strategies in hand, R creates their posterior beliefs based on their prior beliefs and their beliefs about the strategy that the speaker is using. More formally, R’s posterior beliefs are calculated as in (15) (notated $P_{L_i}$, which Goodman and Lassiter call the pragmatic listener).

$$P_{L_i}(A | u) = \frac{P_{L_i}(u | A) \times P_{L_i}(A)}{\sum_{a'} P_{L_i}(u | A') \times P_{L_i}(A')}$$

The upshot of all of this is that, since R believes the speaker to always say the most informative thing (modulo costs), in this example as in Grice (1975), they reason that if *Sarah ate all of the cookies* were true, then S would have picked this message. Since they did not, it is highly likely that it is false; therefore, in most cases, R ends up assigning a very high probability to the set of possible worlds in which Sarah eats some but not all of the cookies.

### 6.3.2 Adjectival interpretation

In Section 6.3.1, I gave an outline of how a Gricean approach to scalar implicatures could be captured within a general Bayesian game-theoretic framework. Scalar implicatures have been the focus of a significant portion of the research in Bayesian pragmatics; however, a number of other proposals have applied similar models to the interpretation of gradable adjectives and the absolute/relative distinction. These works include Franke (2012), Lassiter and Goodman (2013), Qing and Franke (2014), Lassiter and Goodman (2015). For succinctness and because they give an analysis that is most closely related to both Kennedy’s formulation of IE and the analysis of Gricean reasoning above, I will focus on the work of Lassiter and Goodman (Lassiter and Goodman, 2013, 2015). However, I will make some remarks concerning Franke (2012) at the end of the section.

In the Lassiter and Goodman account, the pragmatic interpretation works along the same lines as with scalar implicatures. Furthermore, these authors adopt a DegS analysis of the semantics of gradable adjectives in which these predicates relate individuals to threshold values (degrees). When adjectives combine with POS, the resulting DegPs contain a free variable: the threshold associated with $P$, which we will write $\theta_P$. Since we are interpreting expressions with free variables, we
need to extend the framework described above to deal with the extra complexity that this entails. There are a number of logically possible extensions; however, Lassiter and Goodman propose to conditionalize the interpretation of the message on the possible assignments of values for the threshold variable. Thus, at the first level of interpretation, we now have conditionalization over assignment functions, notated in (16) with the V series.

\[(16) \quad P_{L_0}(A|u, V) = P_{L_0}(A|[u]V^1 = 1)\]

The speaker model \((P_S)\) is the same as in the model for Gricean reasoning, modulo the free variables. The final function that determines the pragmatic interpretation (the pragmatic listener \(P_{L_1}\)) is also the same, except that the receiver takes into account the probability of a particular assignment of variables. As Lassiter and Goodman (2015, p. 18) say, “the pragmatic listener then derives a variable-sensitive interpretation by considering how likely it is that the speaker would have said \(u\) if the answer were \(A\) and the variables were as in \(V\)—and, as usual, multiplying this value by the prior probabilities of \(A\) and \(V\).” The final result of this iterated interpretation process is a kind of “balancing act” in the interpretation of RAs between informativity and plausability, which derives a ‘taller than average’ interpretation. In particular, according to (Lassiter and Goodman, 2013, p. 596),

Very weak interpretations, with \(\theta_{\text{tail}}\) falling in the lower region of the height prior, are probably true; however, the informativity preference entails that speaker would probably not have chosen to use the utterance in such a situation. Conversely, very strong interpretations (with \(\theta_{\text{tail}}\) in the extreme upper tail of the height prior) are dispreferred because they make the utterance very likely to be false—even though they would be extremely informative if true. The effect is a preference for interpretations which make \(A\) fairly tall, but not implausibly so.

In Lassiter and Goodman (2013), the relative/absolute distinction is modelled through proposing the existence of a difference in the shape of the prior beliefs associated with RAs vs AAs. They say (p. 599),

”Prototypical relative interpretations arise with priors with a relatively mild rate of change and little or no mass on the endpoints, while prototypical absolute interpretations arise with priors in which a significant portion of the prior mass falls close to an upper or lower bound.

If we consider the total/partial pair safe and dangerous, Lassiter and Goodman (2013, p. 599) propose that “the probability that something counts as dangerous increases rapidly as its degree of danger deviates
from the zero point, and the probability that it counts as safe decreases similarly (though slightly faster)."

Although these previous statements may sound a lot like the encoding of IE effects directly into the prior probability distributions, they are crucially different in the following way: the link between the application of the predicate and scale structure can be sensitive to properties of particular objects to which the predicate is being applied. This allows for possible exceptions to the 'closed scale $\Rightarrow$ endpoint threshold' generalization discussed in Chapter 3. One such exception, discussed by McNally (2011) and Lassiter and Goodman (2013) is the predicate full when it is applied to sets of wine glasses, since, in this case, we may indeed want to call a glass full once it has 5 oz of wine in it, no matter how large the actual physical glass is. This is a welcome consequence of the model because, as observed by Lassiter and Goodman (2013), such exceptions are not elegantly handled in either Kennedy's or my framework.

6.3.3 Summary

In summary, Lassiter and Goodman's recent work within Bayesian pragmatics aims to show how the effects of IE can be derived from the same principles as those underlying Gricean reasoning associated with quantity implicatures (i.e. informativity). This is an ambitious, interesting research program and the preliminary results presented in these papers suggest that it has potential. This being said, at the point that I am writing this, the area of Bayesian pragmatics is still new and, as researchers in this area point out, there are a number of open areas that need to be explored in a more detailed way. One of the most salient issues concerns how exactly we ought to define the parameters of the games, in particular the set of possible messages and interpretations. The results described above rely greatly on which particular syntactic and semantic alternatives happen to be in the models studied. As such, it is pressing to develop a general theory of how these sets of alternatives are constructed. Indeed, (Lassiter and Goodman, 2015, p. 10) state that

$Lassiter and Goodman suggest, following Fox and Katzir (2011) that the set of possible messages could be sentences denoting possible answers to the question under discussion. In Burnett (2016), I suggest that insights from variationist sociolinguistics (in the sense of Labov, 1966) could be used to give a more principled general answer to these questions. However, I leave further discussion of syntactic/semantic alternatives aside.
Table 6.3 Profiles of DelTCS, Kennedy (2007), and Lassiter and Goodman (2013)

<table>
<thead>
<tr>
<th>Theory</th>
<th>Scales</th>
<th>Semantics-Pragmatics Interface</th>
<th>Logic</th>
<th>Dynamicity</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>DelTCS Kennedy (2007)</td>
<td>Delineation</td>
<td>Interactive</td>
<td>Multi-valued</td>
<td>Static</td>
<td>Deterministic</td>
</tr>
<tr>
<td>Lassiter and Goodman (2014)</td>
<td>Degree</td>
<td>Separate</td>
<td>Classical</td>
<td>Static</td>
<td>Deterministic</td>
</tr>
</tbody>
</table>

A major desideratum in future work will be to get a clearer picture of how speakers and listeners choose a realistic but manageable set of alternatives for pragmatic reasoning, and more generally how people choose an action set under relatively unconstrained conditions.

With this in mind, we can add the Bayesian approaches discussed to our comparison table, as shown in Table 6.3.

Note that informativity is not the only proposed source of IE effects within Bayesian pragmatics found in the literature. Although I focused on Lassiter and Goodman’s work because of its interest in the unification of Gricean reasoning and more general scalar reasoning, Franke (2012) proposes another analysis of IE effects within a game-theoretic framework, which derives these effects from different principles. In particular, Franke proposes that, rather than motivated by pressure to make informative statements, IE effects are the product of the particular cognitive and contextual salience of scalar endpoints. In his games, Franke (2012, p. 7) proposes that,

The strategy in focus is one in which players simply choose whatever is most salient from their own perspective: the sender chooses the most salient property of the designated object; the receiver chooses the most salient object given that property.

He suggests that this strategy is (in his words, p. 7) “plausible and appealing because (i) it presupposes hardly any rationality on the side of the agents, as it merely exploits the agents’ cognitive make-up, but still (ii) it is remarkably successful.” In Section 6.4, I give a new static framework, Degree Tolerant, Classical, Strict, which also derives IE effects, and, as we will see, it does so through adopting (like Franke) a salience perspective rather than Lassiter/Goodman’s informativity perspective.8

7 Note that (Franke, 2012, p. 7) does end up using a notion of visual salience that appeals to a certain kind of informativity or surprise. However, to my knowledge, this is not supposed to be exactly the notion of informativity that is active in quantity implicatures, as with Lassiter and Goodman.

8 Note that Kennedy incorporates both elements of informativity maximization (in the Interpretive Economy proposal) and salience in the natural transitions proposal.
6.4 Degree Tolerant, Classical, Strict

This section presents a new logical framework, DegTCS, which marries certain aspects of DelTCS with certain aspects of Kennedy’s (2007) account of the absolute/relative distinction outlined in Section 6.2.1. We assume almost the same vocabulary as DelTCS: individual constants \((a_1, a_2, a_3 \ldots)\), RAs: \((P, P_1, P_2 \ldots)\), total absolute scalar adjectives \((AA^T)\)\((Q, Q_1, Q_2 \ldots)\) and partial absolute scalar adjective \((AA^P)\)\((R, R_1, R_2 \ldots)\).\(^9\) For every unary predicate \(P\), there is a binary predicate \(> P\). Furthermore, unlike in DelTCS, there is a unary function symbol: \(\text{pos}\).

The syntax of DegTCS is slightly different from DelTCS on account of the \(\text{pos}\) symbol:

1. Constants (and nothing else) are terms.
2. If \(P\) is a predicate symbol, then \(\text{pos} P\) is DegP.
3. If \(\text{pos} P\) is a DegP and \(t\) is a term, then \(\text{pos} P(t)\) is a wff.
4. If \(t_1\) and \(t_2\) are terms, and \(P\) is a predicate symbol, then \(t_1 > P t_2\) is a wff.
5. Nothing else is a wff.

In both DelS and DegS, context plays an important role in the calculation of scalar meaning, especially in the interpretation of the positive form. In DelTCS, this was modelled by having the positive form relativized to a particular property: a CC. The CC was implemented in the logic as a parameter on the evaluation of the positive form of the adjective. The role of CCs in DegS varies depending on the proposal. For example, although many authors propose that CCs are explicitly represented in the logical form of the sentence (Wheeler, 1972; Bartsch and Vennemann, 1973; Cresswell, 1976; Kennedy, 1997; Kennedy and McNally, 2005; von Stechow, 1984, among others), (Kennedy, 2007, section 2.2.) argues that having CCs referenced directly in the semantic denotation of an adjective is not particularly helpful to solve all the empirical puzzles associated with the vagueness and context sensitivity of RAs. Nevertheless, he admits that some contextually determined set of individuals must influence the standard in some way, saying (p. 16):

\(^9\) For the sake of space, we set aside non-scalar predicates in this chapter.
(17) although it is clear that a property that we can descriptively call a “comparison class” influences the computation of the standard of comparison by providing a domain relative to which this degree is computed, this property does not correspond to a constituent of the logical form.

Our theory presented in Chapters 2–5 is consistent with (17), since, as a parameter, the CC can be viewed as just one aspect of the extralinguistic context that influences the application of the positive form of the adjective. Therefore, in what follows, as in DelTCS, we will continue to assume that the positive form of expressions with gradable predicates are evaluated relative to sets of individuals, which provide the domain for the standard function.

We will start by directly interpreting expressions in \( T(\text{olerant}) \) Degree Models, defined below:

**Definition 6.4.1** A Degree t-model is a tuple \( M = (D, \langle D, > \rangle, \text{Dim}, \mu, \llbracket \cdot \rrbracket, \sim, \text{pos}) \), where

1. \( D \) is a non-empty finite set of individuals (entities).
2. \( \langle D, > \rangle \) is an infinite set of totally ordered individuals (degrees).
3. \( \text{Dim} \) is a set of dimensions \( \{\delta_1, \delta_2, \ldots\} \).
   - The set of scales is the set of possible pairings of subsets of degrees with dimensions: \( \mathcal{P}(D) \times \text{Dim} \).
   - We will notate a scale \( (D, > \delta) \), for \( D \in \mathcal{P}(D) \) and \( \delta \in \text{Dim} \).
4. \( \mu \) is a function that associates a scale with every adjectival predicate subject to the following constraints:
   - **Scale structure**
     - For all relative predicates \( P \), \( \mu(P) \) has neither a maximal nor a minimal element.
     - For all total predicates \( Q \), \( \mu(Q) \) has a maximal (but no minimal) element.
     - For all partial predicates \( R \), \( \mu(R) \) has a minimal (but no maximal) element.
     - For readability, we will notate the scale associated with a predicate \( P \) as \( >_P \); i.e. \( \mu(P) = >_P \).
5. \( \llbracket \cdot \rrbracket \) is an interpretation function such that:
   - For all constants \( a_i \), \( \llbracket a_i \rrbracket \in D \).

\(^{10}\) Note that by Cantor’s theorem, this set is uncountably infinite. However, many of the scales associated with a single dimension will be isomorphic to each other.
For all predicates $P$, $\lbrack P \rbrack$ is a function from members of $D$ and degrees on $\succ_p$ ($P$'s associated scale).

For all constants $a_1, a_2$ and all predicates $P$, $\lbrack a_1 \succ_p a_2 \rbrack = 1$ iff $\lbrack a_1 \rbrack \succ_p \lbrack a_2 \rbrack$.

We enforce a “crisp” interpretation of adjectives and comparatives:

**Crisp interpretations**

- $\lbrack P \rbrack^f = \lbrack P \rbrack = \lbrack P \rbrack^f$
- $\lbrack a_1 \succ_p a_2 \rbrack^f = \lbrack a_1 \succ_p a_2 \rbrack = \lbrack a_1 \succ_p a_2 \rbrack^f$

6. $\sim$ is a function from predicates $P$ and contexts/CCs $X \subseteq D$ to binary relations on $X$, notated $\sim^X$ that are reflexive and satisfy certain constraints (to be discussed below).

In DelTCS, we proposed that the distribution of indifference relations across CCs was constrained by a small set of general axioms: Tolerant Convexity (TC), Strict Convexity (SC), Contrast Preservation (CP), Granularity and Minimal Difference (MD) (see Chapter 4), which served to rule out the most unrealistic $\sim$ relations across contexts. There is nothing particularly Delineation-specific in the content of constraints like Granularity and CP (repeated in (18)), so we might assume that something like them holds in DegTCS as well. Of course, since we will not need to derive the scales associated with predicates from their interpretations across CCs, the constraints on $\sim$ will have a very different role in DegTCS than in DelTCS.

(18) a. **Granularity (G):** For all predicates $P_i$, all $X \subseteq D$, and all $a_1, a_2 \in X$, if $a_1 \sim^X_{P_i} a_2$, then for all $X' \subseteq D : X \subseteq X'$, $a_1 \sim^X_{P_i} a_2$.

b. **CP:** For all $X' \subseteq D$, and $a_1, a_2 \in X$, if $X \subseteq X'$ and $a_1 \succ^X_{P_i} a_2$ and $a_1 \sim^X_{P_i} a_2$, then $\exists a_3 \in X' - X : a_1 \succ^X_{P_i} a_3$.

TC and SC were defined with respect to the tolerant and strict scales associated with predicates. In DegTCS, these definitions are still appropriate, but can now be stated much more simply over the scale that is lexically associated with the adjective, as shown in (19).

(19) **Convexity:** For all predicates $P_i$, all models $M$, all $X \subseteq D$, and all $a_1, a_2 \in X$,

a. If $a_1 \sim^X_{P_i} a_2$ and there is some $a_3 \in X$ such that $a_1 \succeq^M_{P_i} a_3 \succeq^M_{P_i} a_2$, then $a_1 \sim^X_{P_i} a_3$.

b. If $a_2 \sim^X_{P_i} a_1$ and there is some $a_3 \in X$ such that $a_1 \succeq^M_{P_i} a_3 \succeq^M_{P_i} a_2$, then $a_2 \sim^X_{P_i} a_3$. 
In DelTCS, we also imposed limited non-symmetries on the indifference relations associated with total and partial predicates in the form of the *Total Axiom* and *Partial Axiom*, repeated in (20).

(20) a. **Total Axiom (TA):** For a total predicate \( Q \), a model \( M \), and \( a_1, a_2 \in D \), if \( \llbracket Q(a_1) \rrbracket_M = 1 \) and \( \llbracket Q(a_2) \rrbracket_M = 0 \), then \( a_2 \not\sim_X^Q a_1 \), for all \( X \subseteq D \).

b. **Partial Axiom (PA):** For a partial predicate \( R \), a model \( M \), and \( a_1, a_2 \in D \), if \( \llbracket R(a_1) \rrbracket_M = 1 \) and \( \llbracket R(a_2) \rrbracket_M = 0 \), then \( a_1 \not\sim_X^R a_2 \), for all \( X \subseteq D \).

I argued earlier in this chapter that Kennedy’s account also involves constraints such as (20) through his proposal of the existence of **natural transitions** that make distinctions between endpoint individuals and non-endpoint individuals. We therefore suppose, following this work, that all individuals who lie at the top endpoint of a scale associated with an AA\(^T\)'s are indifferent only from other individuals at the top endpoint of the scale. Likewise, all individuals who lie at the bottom endpoint of a scale associated with an AA\(^P\) are indifferent only from other individuals at the bottom endpoint. In other words, individuals at endpoints “stand out” in every context. This proposition is stated more formally in (21).

(21) **Natural transitions**

a. For all total predicates \( Q \) and all \( a_1 \in D \),

If \( \llbracket Q \rrbracket(a_1) = d_{\text{max}} \), then for all \( X \subseteq D : a_1 \in X \), if \( a_2 \not\sim_X^Q a_1 \),

then \( \llbracket Q \rrbracket(a_2) = d_{\text{max}} \).

b. For all partial predicates \( R \) and all \( a_1 \in D \),

If \( \llbracket R \rrbracket(a_1) = d_{\text{min}} \), then for all \( X \subseteq D : a_1 \in X \), if \( a_2 \not\sim_X^R a_1 \),

then \( \llbracket Q \rrbracket(a_2) = d_{\text{min}} \).

We now turn to the definition of Pos and degree phrases (DegPs), i.e. constituents of the form \([\text{pos} P]\). If we follow Kennedy, the denotation of a DegP containing the positive form of an adjective should be a subset of the pertinent individuals in the context, those that stand out with respect to a predicate. Observe that this is very similar to TCS’s **strict denotation**: an individual is included in the strict extension of a predicate just in case everything that they are indifferent from is in the classical extension of that predicate (22). Put another way, the strict denotation of a predicate consists of all the individuals that in the predicate’s classical extension “stand out” from all the members of the anti-extension.
(22) **DelTCS strict denotation**

\[
\langle P \rangle^s_X = \{ a_1 : \text{for all } a_2 \sim_p^X a_1, a_2 \in \langle P \rangle \}
\]

I therefore suggest that a natural interpretation of a DegP [pos P] in the style of Kennedy would be to take the strict interpretation of a predicate in a context (i.e. a CC).

This being said, since our adjectives are no longer simply interpreted as properties, their “classical” extension cannot be defined in the same way as in DelTCS. So what can we do? There are multiple solutions to this problem;\(^{11}\) the one we will adopt here is to incorporate basic divisions between the positive form’s extension and anti-extension into the pos morpheme.

(23) pos is a function that takes a predicate \( P \) and a CC \( X \subseteq D \) and returns a partition on \( X \), separating \( P \)'s extension in \( X \) (notated \( P^+_X \)) from its anti-extension (notated \( P^-_X \)) in \( X \).

As in Kennedy (2007), we put a constraint on pos which makes it sensitive to the degree orderings on the scale associated with the adjective.

(24) **Scalar constraint on pos:**

Let \( P \) be a predicate and let \( X \) be a CC. Then for all \( a_1 \in X \),

- If \( a_1 \in P^+_X \), then for all \( a_2 \in X \) if \( \langle P \rangle(a_2) \succeq_p \langle P \rangle(a_1) \), then \( a_2 \in P^+_X \).
- If \( a_1 \in P^-_X \), then for all \( a_2 \in X \) if \( \langle P \rangle(a_1) \succeq_p \langle P \rangle(a_2) \), then \( a_2 \in P^-_X \).

For reasons that will be made clear below, we need to impose some relation between the extension of a total predicate and the indifference relations associated with the elements in the extension. In particular, we propose a \( \sim \) closure condition on the indifference relations associated with individuals in the positive extension of the predicate\(^{12}\) (25). Keeping in line with the analysis of partial predicates as duals of total predicates, \( \sim \) closure applies to the anti-extension of \( AA^P \)’s.

---

\(^{11}\) See also works such as van Rooij (2011c), Burnett (2015) for alternatives to adopting (anti-)extensions.

\(^{12}\) One way of thinking about this constraint is as the DegS version of Chapter 4’s MD constraint, which ensured that \( \sim \) respected the boundaries of the classical extension of a predicate in minimal CCs.
(25) **(Anti-)Extension indifference closure**

a. Let $Q$ be a total predicate. Then, for all $X \subseteq D$ and $a_1 \in X$, if $a_1 \in Q_+^X$, then for all $a_2 \in Q_+^X$, $a_2 \sim_Q^X a_1$.

b. Let $R$ be a partial predicate. Then, for all $X \subseteq D$ and $a_1 \in X$, if $a_1 \in R_+^X$, then for all $a_2 \in R_+^X$, $a_2 \sim_R^X a_1$.

Now, with the $\text{pos}$ function, we require a more complicated rule for the interpretation of a DegP:

(26) **Interpretation of DegP:**

For all predicates $P$ and $X \subseteq D$,

$$\llbracket \text{Pos}P \rrbracket_X = \{a_1 : \text{for all } a_2 \sim_P^X a_1, a_2 \in P_+^X\}$$

Rule (26) has the effect of “extracting” what in DelTCS would have been the strict denotation of a predicate in a CC (compare with (22)).

For convenience, we force the strict interpretation of DegPs to be identical to their classical interpretation (27-a); however, it would be interesting to explore in the future to whether having different classical and strict interpretations of such constituents could be useful in the analysis of higher order vagueness. We also define the tolerant extension of the positive form of an adjective as in TCS, shown in (27-b).

(27) For all predicates $P$ and $X \subseteq D$,

a. $\llbracket \text{pos}P \rrbracket_X^t = \llbracket \text{pos}P \rrbracket_X$

b. $\llbracket \text{pos}P \rrbracket_X^s = \{a_1 : \text{there is some } a_2 \sim_P^X a_1 : a_2 \in P_+^X\}$

And the Tolerant, Classical, Strict truth(s) of a formula with the positive form of a predicate is defined as follows:

(28) For all predicates $P$, CCs $X$ and $a_1 \in X$,

$$\llbracket \text{pos}P(a_1) \rrbracket_X^{t/s/e} = \begin{cases} 1 & \text{if } \llbracket a_1 \rrbracket \in \llbracket P \rrbracket_X^{t/e/s} \\ 0 & \text{if } \llbracket a_1 \rrbracket \in X - \llbracket P \rrbracket_X^{t/e/s} \\ i & \text{otherwise} \end{cases}$$

With a degree extension of TCS in place, we now turn to the results.

6.4.1 **Results**

This section presents a number of results of the DegTCS framework, particularly associated with the denotation of the interpretation of the
positive form of adjectival predicate. In DeTCS, we imposed constraints on the distribution of the positive form of the predicate and the indifference relations across CCs and showed how this had consequences for scale structure. In DegTCS, we put constraints on the scales associated with predicates and the indifference relations across CCs. So now we will show how the appropriate generalizations concerning the distribution of the positive form of the predicate fall out from the system.

The first results hold across the full range of gradable predicates: relative, total, and partial. With the structure defined above and, in particular, the scalar constraint on pos (24), we can prove van Benthem’s No Reversal (NR) holds of the classical extension of the positive form ([pos P]). We note that, since the classical interpretation of the positive form of a predicate is the same as the strict interpretation, NR also holds at the strict level.

**Theorem 6.4.2 No Reversal.** Let P be a predicate, let \( X \subseteq D \), and let \( a_1, a_2 \in X \). And suppose \( \|\text{pos}(a_1)\|_X = 1 \) and \( \|\text{pos}(a_2)\|_X = 0 \). Then,

- There is no \( X' \subseteq D \) such that \( \|\text{pos}(a_1)\|_{X'} = 1 \) and \( \|\text{pos}(a_2)\|_{X'} = 0 \).

We now turn to results associated with AA\(^T\)s. For space considerations, we will limit our exposition to theorems associated with total AA\(^T\)s; however, the interested reader is encouraged to verify that the corresponding results hold dually for partial predicates.

Concerning total predicates, we can first show that if an individual \( a_1 \) is at the top endpoint of the scale associated with an AA Q, then it must always be included in the classical denotation of the positive form of Q, in every context.

\[ \text{Corollary: If } \|\text{pos}(a_1)\|_X = 1 \text{ and } \|\text{pos}(a_2)\|_X = 0, \text{ then } a_1 \geq_P a_2. \text{ Proof. Suppose for a contradiction that } a_1 \not\geq_P a_2, \text{ i.e. } a_2 >_P a_1. \text{ Since } \|\text{pos}(a_2)\|_X = 0, \text{ there is some } a_2 \in X \text{ such that } a_2 \sim_P a_2 \notin P^+_X. \text{ Since } \|\text{pos}(a_1)\|_X = 1, \text{ by (24), } a_1 >_P a_3. \text{ So } a_2 >_P a_3 >_P a_1. \text{ Since } a_3 \sim_P a_2, \text{ by convexity (19), } a_3 \sim_P a_1. \text{ So by (26), } \|\text{pos}(a_1)\|_X = 0. \bot \text{ So } a_1 \geq_P a_2. \]
Theorem 6.4.3 Top endpoint.\(^{14}\) Let \(Q\) be a total predicate and let \(a_1 \in D\). Then, if \(\| Q \|(a_1) = d_{\text{max}}\), then for all \(X \subseteq D\), if \(\| \text{pos} Q \|_X \neq \emptyset\), then \(a_1 \in X\), then \(\| \text{pos} Q(a_1) \|_X = 1\).

More generally we prove a version of what in DelTCS was our main constraint governing the distribution of the classical denotations of AAs across CCs: the absolute adjective axiom (AAA). In DegTCS, however, this is a theorem:\(^{15}\)

Theorem 6.4.4 The Absolute Adjective.\(^{16}\) Let \(Q\) be a total absolute predicate and let \(X \subseteq D\) such that there is at least some \(a_1 \in X\) such that \(\| Q \|(a_1) = d_{\text{max}}\). Then, for all \(a \in X\),

1. If \(\| \text{pos} Q(a) \|_X = 1\), then \(\| \text{pos} Q(a) \|_D = 1\).
2. If \(\| \text{pos} Q(a) \|_D = 1\) and \(a \in X\), then \(\| \text{pos} Q(a) \|_X = 1\).

Going beyond the classical denotation, we show that the tolerant use of the tolerant interpretation of the positive form of an AA\(^T\) can only hold of an individual if that individual is either at the top of the adjective's scale or considered indifferent from top endpoint individuals. In other words, from the basic definitions set out above and, most importantly, the natural transitions constraints on \(\sim\), we can prove what IE was designed to account for: the observation that the use of the positive form of an AA\(^T\) is limited to picking out those objects that lie at or very close to the top endpoint.

\(^{14}\) Proof. Suppose \(\| Q \|(a_1) = d_{\text{max}}, \| \text{pos} Q \|_X \neq \emptyset\), and \(\| \text{pos} Q(a_1) \|_X = 0\). Since \(\| \text{pos} Q \|_X \neq \emptyset\), there is some \(a_2 \in X\) such that \(a_2 \in P_X^\text{\#}\). Since \(\| Q \|(a_1) = d_{\text{max}}\), \(\| Q \|(a_1) \geq \| Q \|(a_2)\). So, by (24), \(\| \text{pos} Q(a_1) \|_X = 1\). \(\square\)

\(^{15}\) The subclause there is at least some \(a_1 \in X\) such that \(\| Q \|(a_1) = d_{\text{max}}\) is necessary for the proof to go through in this version of the system, since we have put only one constraint on pos (24) that restricts the application of the (anti-)extensions \((P_X^\text{\#} \text{ and } P_X^\text{\mbox{\#}})\). We could easily get rid of this clause by putting some constraints on pos that are similar to van Benthem’s discussed in Chapter 4, but exploring this option in detail will be left to future work.

\(^{16}\) Proof. 1. Suppose \(\| \text{pos} Q(a) \|_X = 1\), and (for a contradiction) that \(\| \text{pos} Q(a) \|_D = 0\). So there is some \(a_2 \in P_D^\text{\#}\) such that \(a_2 \notin P_D^\text{\mbox{\#}}\). By natural transitions, \(\| Q \|(a) \neq d_{\text{max}}\). Since \(\| \text{pos} Q(a) \|_X = 1\), \(a \in Q_X^\text{\#}\). By (24), \(a_1 \notin Q_X^\text{\mbox{\#}}\). By (25), \(a \sim _Q a_1\). But since \(\| Q \|(a_1) = d_{\text{max}}\) and \(\| Q \|(a) \neq d_{\text{max}}\), by (21), \(a \prec _Q a_1\). \(\square\)

2. Suppose \(\| \text{pos} Q(a) \|_D = 1\) and \(a \in X\). Suppose (for a contradiction) that \(\| \text{pos} Q(a) \|_X = 0\). Since \(\| \text{pos} Q(a) \|_X = 0\), \(a \notin Q_X^\text{\#}\); so by (24), \(a_1 \succ _Q a\). Since \(\| \text{pos} Q(a) \|_D = 1\), \(a \in Q_D^\text{\#}\), so \(a \sim _Q a_1\). But since \(\| Q \|(a) \neq d_{\text{max}}\), by (21), \(a \prec _Q a_1\). \(\square\)
Theorem 6.4.5 Interpretive Economy.\textsuperscript{17} Let $Q$ be a total absolute predicate and let $X \subseteq D$ such that there is at least some $a_1 \in X$ such that $\|Q\|_X(a_1) = d_{\text{max}}$. Then, for all $a \in X$,

\begin{itemize}
  \item $\|\text{pos}Q(a)\|_X^* = 1$ iff there is some $a_2 \in X : \|Q\|(a_2) = d_{\text{max}}$ and $a_2 \sim^X a$.
\end{itemize}

In summary, in this section we showed that, if we integrate some of the proposals in Kennedy (2007) within a DegS extension of the Tolerant, Classical, Strict framework for analyzing the puzzling properties of vague language, we can “flip” the DelTCS framework around and use the structure of TCS to capture relationships between scale structure and the context sensitivity of the positive form of a gradable predicate. In particular, I suggest that if a formalization of Kennedy’s (2007) proposal into the TCS system is done in the way outlined above, we gain another perspective on the IE meta-principle: in particular, theorem 6.4.5 shows that the addition of such an extra principle is not required in DegTCS since the facts that it derives are already provable from 1) the treatment of the positive form of the adjective as denoting the strict interpretation of the predicate, and 2) Pinkal/Kennedy’s independent proposal of natural transitions associated with the top/bottom endpoint of a total/partial predicate respectively.

6.5 Conclusion

This chapter gave a detailed comparison between the DelTCS system developed in the first part of the book (Chapters 2–5) and a couple of other current accounts of the absolute/relative distinction, focusing particularly on Kennedy’s (2007) account within a DegS approach to the semantics of gradable predicates. The range of frameworks compared in this chapter is shown in Table 6.4. I argued that, despite certain differences in the overall set-up of the theories, there were significant similarities between them and that Kennedy’s proposals could be

\textsuperscript{17} Corollary: Suppose there is some $a_1 \in X$ such that $\|Q\|_X(a_1) = d_{\text{max}}$. Then, for all $a \in X$, if $a \in Q^+_X$, then $\|Q\|(a) = d_{\text{max}}$. Proof: Immediately from (24), (21), and (25). 

Proof of Theorem: $\Rightarrow$ Suppose $\|\text{pos}Q(a)\|_X^* = 1$. Case 1: $a \in Q^+_X$. Then by the corollary above, $\|Q\|(a) = d_{\text{max}}$, and since $\sim$ relations are reflexive $a \sim^X a$. $\checkmark$ Case 2: $a \notin Q^+_X$. Then, by the definition in (27-b), there is some $a_2 \in Q^+_X$ such that $a_2 \sim^X a$. So, by the corollary above, $\|Q\|(a_2) = d_{\text{max}}$. $\checkmark$
Table 6.4  Profiles of DelTCS, Kennedy (2007), Lassiter and Goodman (2013), and DegTCS

<table>
<thead>
<tr>
<th>Theory</th>
<th>Scales</th>
<th>Semantics–Pragmatics Interface</th>
<th>Logic</th>
<th>Dynamicity</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>DelTCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DegTCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lassiter and Goodman (2014)</td>
<td>Degree</td>
<td>Interactive</td>
<td>Static</td>
<td>Deterministic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Degree</td>
<td>Separate</td>
<td>Multi-valued</td>
<td>Deterministic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Degree</td>
<td>Separate</td>
<td>Classical</td>
<td>Deterministic</td>
<td></td>
</tr>
</tbody>
</table>

incorporated into a novel DegS extension of the TCS framework that I developed in this chapter.

The DegTCS framework thus constitutes a “bridge” between DelTCS and the DegS framework that is more commonly adopted in natural language semantics. In addition to showing how the proposals made in the first five chapters of this work could largely be translated into the DegS framework, I argued that the DegTCS system is independently interesting since it allows us to prove a version of Kennedy’s much discussed IE principle within a static deterministic framework. In other words, in the first part of the book, I argued that adopting a TCS approach to the semantics/pragmatics of gradable predicates is particularly useful for researchers wishing to pursue a DelS treatment of scalar adjectives because it allowed for an analysis of the gradability of AAs that was not possible in non-TCS DelS. In this chapter, I argue that such an approach can be helpful in the analysis of the absolute/relative distinction independently of the theory of gradability that one adopts.

In Chapter 7, we continue our extension of the ideas developed in the first five chapters and we turn to context sensitivity, vagueness, and scale structure outside the adjectival domain.
7

Beyond the adjectival domain

7.1 Introduction
This chapter extends the framework developed in Chapters 2–5 to modelling data from outside the adjectival domain and, in particular, the context sensitivity, vagueness, and scale structure properties of determiner phrases (DPs). I give a mereological generalization (Simons, 1987; Hovda, 2008, among others) of the Delineation Tolerant, Classical, Strict (DeTCS) framework, and show how we can associate orderings with DPs in a manner parallel to the manner in which we constructed the scales associated with adjectives in the previous chapters. I therefore conclude that DeTCS (and its mereological extension M-DeTCS) constitutes a versatile architecture in which to set analyses of context sensitivity, vagueness, and gradability in natural language.

This chapter is laid out as follows: In Section 7.2, I give a brief overview of the basic vagueness and context sensitivity patterns that we see in the DP domain. I argue that we can decompose the set of determiner phrases (DPs) in a language like English into the same three basic semantic categories found in the adjectival domain: relative DPs, absolute DPs, and precise/non-scalar DPs. Then, in Section 7.3, I give a mereological extension of DeTCS, and I show how we can use this system to give an analysis of plural count noun phrases (girls, guys, townspeople, etc.) and existential bare plurals. Then, in Section 7.4, I give an analysis of the vagueness and gradability properties of definite plural DPs in distributive and collective contexts (i.e. expressions like the girls paired with predicates like are asleep, gathered, and are a group of four).

7.2 Context sensitivity and vagueness patterns
In this section, I present a brief review of the vagueness patterns associated with DPs. I argue that (following the presentation in Burnett, 2012b) we see three main classes of DPs: relative/“intensional,” absolute/imprecise, and non-scalar/precise DPs.
(1) Relative/Intensional DP
   a. Many girls arrived.
   b. Few boys left.

(2) Absolute/imprecise DP
   a. The/these girls are late.
   b. 30,000 spectators were at the game.

(3) Non-scalar/precise DP
   a. All the girls are late.
   b. No girls are late.
   c. 29,821 spectators were at the game.

It has often been observed (since at least Keenan and Faltz, 1985, if not earlier) that vagueness is a property that holds not only of scalar adjectives and nouns (like *heap*), but also of DPs. The first category of DPs that display the characterizing properties of vague language are what I refer to as *relative* quantifier phrases like *many people, few girls,* and *several boys* (cf. Keenan and Faltz, 1985; Lappin, 2000; Hackl, 2001; among others). Like relative adjectives (RAs), these constituents display borderline cases, fuzzy boundaries, and can be used in a Soritical argument in all (or almost all) contexts. For example, consider a context in which we are describing a party to which we expected about half the people (of a guest list of 100) to show up. In this context, (4) is clearly true if 90 or 80 people came, and clearly false if only 5 people came. However, what if 60 people came? 61?

(4) Many people came to the party.

Furthermore, in this context, at which number of guests does the sentence go from being true to being false? Thus, with *many people,* we can construct a Soritical series based on the number of guests at the party and form the appropriate paradoxical argument. We can easily think of other contexts in which *many people* displays the symptoms of vagueness, and, indeed, like adjectives such as *tall,* it is difficult to think of contexts in which this DP (or DPs like *few men* and *several people*) could be used precisely. Furthermore, we find the same pattern with the negation of *many:* *not many.* At which number of guests is (5) falsified?

(5) Not many people came to the party.

---
1 Keenan, Faltz, and Lappin refer to these DPs as *intensional* quantifier phrases.
Many people also display the same context sensitivity properties as RAs: just in the way that we can apply the predicate *tall* to objects that we would not consider tall in general, provided that they can be considered tall in comparison to some other contextual comparison class (CC), we can use (4) to describe situations in which only a small number of guests attended the party, provided that this number (significantly) exceeds our expectations. In sum, like RAs, DPs containing *many, few, or several* show the same context sensitivity and vagueness properties that RAs like *tall* or *expensive* do. I will therefore refer to these constituents as *relative DPs*.

The second kind of pattern that we see in the DP domain is one that parallels the vague/precise pattern displayed by absolute scalar adjectives (AAs). As discussed in many works, such as Dowty (1987), Yoon (1996), Lasersohn (1999), Brisson (2003), Malamud (2006), in contexts where it is important to be precise, sentences with definite descriptions and distributive predicates (like (6)) are true and appropriate just in case every member of the group denoted by the subject DP is affected by the predicate. Suppose (as in an example from Lasersohn, 1999, p. 523) that we are conducting a sleep experiment and that it is vital to our purposes that the people that we are studying actually fall asleep. In this context, not only is (6) true if all the subjects are asleep, but it is clearly false if one of the participants of the study is awake. In other words, like AAs such as *empty* and *straight*, definite plural DPs can be used precisely in some contexts.

(6) The subjects are asleep.

Furthermore, as discussed in Dowty (1987) and Brisson (2003), the precise use of a definite plural can be enforced by the linguistic context. For example, when they are paired with a member of a certain class of collective predicate (what is known (after Dowty, 1987) as a *pure cardinality* predicate, or (after Corblin (2008)) as a *holistic* predicate), the predicate must hold of the entire group denoted by the subject for the sentence to be felicitous. This can be seen in (7), where the predicate *are a group of four* must apply the group composed of every single girl picked out by the definite description, regardless of the extra-linguistic context.²

² The contribution of the linguistic context to the availability of the characteristic properties of vagueness with definite plurals will be examined at greater length in Section 7.4.
The girls are a group of four.

Despite the precision of (6) and (7), the aforementioned authors also observe that, in contexts where precision is not as important, sentences with definite plurals and distributive predicates can be said even if the plural predicate does not affect every single part of the subject. In the words of Brisson (2003), these DPs give rise to non-maximality effects with distributive predicates. Consider the case (also from Lasersohn, 1999) where, instead of describing an experiment, we are describing the state of a town at night. In this context, it is perfectly natural to use (8) even if a couple of insomniacs or night-watchmen are still awake. Note crucially, however, that, in contrast to relative DPs like many townspeople, the sub-group of the townspeople that are required to be sleeping in order to render (8) acceptable must at least be considered to differ only negligibly from the maximal group of townspeople. We now have an argument based on context sensitivity that plural definite DPs in distributive contexts show the same pattern as total absolute scalar adjectives (AA\textsuperscript{T}s) like empty and straight.

The townspeople are asleep.

Furthermore, we can observe that, in the contexts where non-maximality is allowed, definite plurals display the hallmark properties of vague language. For example, in the context described above, (8) is clearly true when all the townspeople are asleep, and clearly false when less than half of the townspeople are asleep. However, what if 75\% are asleep? 70\%? It is not clear: these are the borderline cases. Furthermore, once the context allows us to tolerate exceptions with a definite plural, exactly how many exceptions are we allowed to tolerate before the sentence becomes clearly false? It seems bizarre to think that, in this context, subtracting a single townsperson could make a difference to whether we would assent to (8), so how is it that our reasoning with definite plurals is not paradoxical? Again, we see similar asymmetries in the use of statements like (8) affecting the range of Sorites arguments we can construct with them as with AA\textsuperscript{T}s: while, depending on context, it may be appropriate to say (8) when not all the townspeople are asleep, there are no contexts in which it is appropriate to say negative forms such as (9) when all of the townspeople are asleep.\textsuperscript{3}

\textsuperscript{3} Negative sentences with definite plural DPs are a little bit tricky because the version of the sentence where sentential negation appears in the same clause (i) shows what is often
(9) It is not the case that the townspeople are asleep.

For these reasons, then, I will refer to definite plural descriptions such as the townspeople as total absolute DPs. Note importantly that, in all these examples, it is not that the reference of the definite description/numeral phrase is, itself, vague. In fact, non-maximality effects can be found even with plural demonstrative phrases in cases where the precise group to which we are attributing the plural property is completely identified. For example, the sentence in (10) could be said in a situation where we know exactly who the girls are and what it takes to be Canadian, but in this situation, for our purposes, it is not necessary that all the girls have that property.

(10) These girls are Canadian.

In other words, what is vague in a sentence like (10) is the actual predication of the property Canadian to the parts of the plural subject. I therefore propose that there exists an important distinction to be made between vague predicates (like tall and heap) on the one hand and the vagueness we see with definite plurals in distributive contexts, which I will call vague predication on the other hand. A consequence of this distinction is that sentences with the appropriate kinds of DPs and vague predicates will be potentially vague (P-vague) in different ways. For example, a sentence like (11) is vague with respect to the extent to which the individuals in the subject denotation count as heaps (predicate vagueness). It is also vague with respect to the extent to which the heaps (borderline or otherwise) count as tall (also predicate vagueness). Furthermore, the sentence is vague with respect to how the predicate distributes over the subject (i.e. how many of the (possibly borderline) heaps are (possibly borderline) tall).

(11) The heaps are tall.

An account for this kind of “vague predication” and how it differs from instances of “vague predicates” will be given in Sections 7.3 and 7.4.

The AA^T’s pattern is not limited to definite or demonstrative plurals. For example, Krifka, 2007; Smith, 2008; Cummins et al., 2012, among call a homogeneity effect (see Fodor, 1970, and much subsequent work in the field), where the sentence is understood to communicate that all of the townspeople were awake.

(i) The townspeople are not asleep.

Homogeneity will be (briefly) discussed in Section 7.4.
others, discuss some similar observations about DPs containing “round” numeral expressions like 30,000 and 100. Although we can use these terms precisely, in many contexts, sentences like (12) can be said even if slightly fewer than 30,000 spectators attended the game or if the stop sign is not quite 100 m away. As with AAs and definite plurals, in these contexts, the expression is vague with respect to how much deviation from the quantity denoted by the numeral phrase is allowed before the sentence is clearly false.

(12) a. There were 30,000 spectators at the football game.
   b. There is a stop sign 100 m down the road.

Thus, I further suggest that the DPs containing round numerals constitute a (particular subclass) of total absolute DPs.

In the adjectival domain, we saw reasons for distinguishing between two classes of AAs: $AA^T$s (like empty, dry, clean etc.) and partial absolute scalar adjectives ($AA^P$s) (like wet, dirty, bent etc.). However, (although this is up for debate) it is not clear that we find such a distinction in the DP domain. One possible candidate for the partial pattern is the existential bare plural, such as the DP dogs in example (13).

(13) Dogs are in the yard.

Although these DPs have existential force (like the $AA^P$s), they do not seem to have the kind of default evaluative meaning that we get with the use of $AA^P$s in the positive construction (This towel is wet). Additionally, as far as I can see, there do not seem to be partial scalar modifiers (such as slightly) that combine with existential bare plural DPs, and if we apply the partial modifier strictly speaking to the subject DP in (14), it seems to only target the property of being a dog (i.e. the NP constituent dog), rather than how many dogs need to be in the yard for the sentence to be acceptable.

(14) Dogs, strictly speaking, are in the yard.

I therefore suggest that, at least when it comes to vague predication, existential DPs are precise. Thus, although there may be some vagueness concerning what counts as a dog in (13), the distributive predication conditions of this sentence are precise: (modulo implicatures) the sentence is true iff the predicate in the yard affects at least one dog.

* Maybe Fido has very cat-like behavior...
There are other kinds of DPs and constructions that do not give rise to vague predication. Consider, for example, the contrast in the sentences in (15) (from Lasersohn, 1999).

(15)  a. The townspeople are asleep.
     b. All the townspeople are asleep.

As discussed above, given an appropriate context, (15-a) is a vague utterance. However, as observed by Dowty (1987) and Lasersohn (1999), (15-b) is precise: it is true just in case every single townsperson is asleep. Other DPs that force a precise use\(^5\) are those headed by logical expressions like every and no, and explicit co-ordination structures like (17).

(16)  a. Every girl in this room is asleep.
     b. No one is asleep.

(17)  John and Mary are asleep.

Finally, as discussed in Krifka (2007), very small and “unround” numeral phrases like those in (18) also enforce precision.

(18)  a. Three girls are asleep.
     b. 29,871 spectators were at the game.
     c. There is a stop sign 103 m down the road.

### 7.2.1 Summary

In summary, we have seen that the main context sensitivity and vagueness pattern that characterized the non-scalar/relative/absolute distinction are also found in the DP domain.

In the rest of the chapter, we will look more closely at the gradability and scale structure properties of two classes of constituents in Table 7.1:

---

\(^5\) It should be noted that it may still be possible to find some exceptional context in which logical expressions and expressions like all the girls can be used vaguely, particularly in cases of very high granularity (consider a situation where two people are in a 3,000-seat theatre for (ii)).

(ii) No one was in the theatre.

However, as discussed in Lasersohn (1999), it is markedly more difficult to find such contexts than with simple definite descriptions. I believe that vague uses of logical expressions bear some similarity to cases of vague “coerced” non-scalar adjectives (NSs) discussed in Chapters 2–5.
Table 7.1  Context sensitivity/vagueness classes in the adjective and DP domains

<table>
<thead>
<tr>
<th>Class</th>
<th>Adjectival Domain</th>
<th>DP Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>tall, expensive…</td>
<td>Many people, few people, several people…</td>
</tr>
<tr>
<td>AA(\star)</td>
<td>empty, dry…</td>
<td>The people, these people, 100 people…</td>
</tr>
<tr>
<td>AA(\prime)</td>
<td>wet, dirty…</td>
<td>??</td>
</tr>
<tr>
<td>Precise</td>
<td>prime, hexagonal…</td>
<td>All the people, 103 people, people…</td>
</tr>
</tbody>
</table>

existential bare plurals (which I suggested are precise DPs) and definite plurals (which I suggested are total absolute DPs). I will show that we can arrive at appropriate analyses of the relationship between the vagueness and context sensitivity properties of these constituents discussed above and the orderings associated with them within a mereological extension of Delineation Tolerant, Classical, Strict. Clearly, another important class of vague DPs are the relative DPs (i.e. those containing “Q-adjjectives” like many or few). These constituents will not be analyzed in this monograph; however, a treatment of these constituents and the relation between the scales associated with the subject in (19-a) and quantity comparatives such as (19-b) is given in Burnett (2015).

(19)  a. Many linguists came to the party.
     b. More linguists than philosophers came to the party.

7.3 Mereological Delineation Tolerant, Classical, Strict

We start with a simple language for modelling the behavior of P-vague predicates paired with singular definite subjects and existential bare plural subjects. In the rest of the chapter, I will group all the different classes of adjectival predicates together alongside nominal and verbal predicates as singular predicates, not making any lexical class-based distinctions between them. This is, of course, a simplification. There may indeed be interesting interactions between the potential vagueness associated with different kinds of DPs and the potential vagueness of relative, total, and partial predicates. However, I leave the modelling of these interactions within my framework to future research.

As such, in this section, we will be interested in modelling the context sensitivity, vagueness, and scale structure properties of the following kinds of sentences:

\[\text{See, for example, Yoon (1996), Rotstein and Winter (2004), and Malamud’s (2006) discussion of how distributivity interacts with the total/partial distinction.}\]
(20)  a. Mary is tall.
    b. (Some) girls are tall.
    c. Girls gathered.

7.3.1 Language

I first define the language and then give it a semantics.

Definition 7.3.1 Vocabulary. The language of Mereological Delineation Tolerant, Classical, Strict (M-DelTCS) M-DelTCS is comprised of (at least) the following basic expressions:

1. A set of constants $a, a_1, a_2, a_3 \ldots$
2. Two classes of unary predicates:
   • Singular predicates: $P, P_1, P_2, P_3 \ldots$
   • Complex Collective predicates: $Q, Q_1, Q_2, Q_3 \ldots$
3. A function symbol on the set of singular predicates: $\ast$
4. A function symbol on the set of singular predicates: $\exists$

The singular predicates $P_1, P_2 \ldots$ will be acting like P-vague singular adjectival, nominal, and verbal predicates such as *is tall, is a heap,* and *dances,* respectively. The complex collective predicates $Q_1, Q_2 \ldots$ will be acting like P-vague plural collective predicates like *to gather* and *to meet.* In line with the work of Link (1984), distributive plural predicates are constructed in the syntax through the addition of a pluralizing distributivity operator $\ast$. Complex collective predicates, on the other hand, are lexically plural.

Definition 7.3.2 Plural predicates.

1. If $P$ is a singular predicate, then $P^\ast$ is a plural predicate.
2. Complex collective predicates are plural predicates.

Since the DP subjects that interest us in this chapter are plural, we pluralize singular predicates when they appear within the DP in a similar way to the distributive plural predicates shown above. We thus introduce a nominal pluralization operator: $\ast$.

Definition 7.3.3 Plural noun phrases (NPs). If $P$ is a singular predicate, then $P^\ast$ is a plural NP.
For the moment, we are interested in modelling the behavior of existential bare plurals. Therefore, NPs are prefixed with a $\exists$ operator, which will have a very similar semantics to the existential quantifier.

**Definition 7.3.4** *Plural determiner phrases.* If $P^*$ is a plural NP, then $\exists P^*$ is a plural DP.

Finally, we define the set of well-formed formulas (wffs) of the language as follows:

**Definition 7.3.5** *Well-formed formulas.*

1. If $a$ is a constant and $P$ is a singular predicate, then $P(a)$ is a wff.
2. If $a$ is a constant and $P$ is a singular predicate, then $P^*(a)$ is a wff.
3. If $a$ is a constant and $Q$ is a complex collective predicate, then $Q(a)$ is a wff.
4. If $\exists P^*$ is a DP and $P^*$ (or $Q$) are plural predicates, then $\exists P^*(P^*)$ is a wff, as is $\exists P^*(Q)$.

In other words, we have formulas in our language aiming to model English sentences as shown in Table 7.2.\footnote{As discussed in Section 7.3.2, constants will range over both singular and plural individuals. This is done to keep the language as simple as possible, and I trust this will not be misleading.}

### 7.3.2 Classical semantics

As in DelTCS, every expression in the language is assigned three semantic values: a classical semantic value, a tolerant value and a strict value. However, unlike basic DelTCS, in which we supposed our domain to be an unordered set of individuals as in the models for first-order logic, we now interpret the expressions of our language into a domain that encodes **mereological** (i.e. part-structure) relations...
between its individuals. More precisely, we define our model structures as follows:

**Definition 7.3.6 Model structure.** A model structure \( M \) is a tuple \( \langle D, \preceq \rangle \), where \( D \) is a finite set of individuals, \( \preceq \) is a binary relation on \( D \).

Furthermore, we stipulate that \( \langle D, \preceq \rangle \) satisfies the axioms of classical extensional mereology (CEM). First, some definitions:

**Definition 7.3.7 Overlap (\( \circ \)).** For all \( a_1, a_2 \in D \), \( a_1 \circ a_2 \) iff \( \exists a_3 \in D \) such that \( a_3 \preceq a_1 \) and \( a_3 \preceq a_2 \).

**Definition 7.3.8 Fusion (Fu).** For \( a_1 \in D \) and \( X \subseteq D \), \( Fu(a_1, X) \) (\( a_1 \) fuses \( X \)) iff, for all \( a_2 \in D \),

\[ a_2 \circ a_1 \] iff there is some \( a_3 \) such that \( a_3 \in X \) and \( a_2 \circ a_3 \).

We now adopt the following constraints on \( \langle D, \preceq \rangle \):

1. **Reflexivity.** For all \( a_1 \in D \), \( a_1 \preceq a_1 \).
2. **Transitivity.** For all \( a_1, a_2, a_3 \), if \( a_1 \preceq a_2 \) and \( a_2 \preceq a_3 \), then \( a_1 \preceq a_3 \).
3. **Anti-symmetry.** For all \( a_1, a_2 \in D \), if \( a_1 \preceq a_2 \) and \( a_2 \preceq a_1 \), then \( a_1 = a_2 \).
4. **Strong supplementation.** For all \( a_1, a_2 \in D \), for all atoms \( a_3 \), if \( a_3 \preceq a_1 \), then \( a_3 \circ a_2 \), then \( a_1 \preceq a_2 \).
5. **Fusion existence.** For all \( X \subseteq D \), if there is some \( a_1 \in X \), then there is some \( a_2 \in D \) such that \( Fu(a_2, X) \).

We can note that, in CEM, for every subset of \( D \), not only does its fusion exist, but it is also unique (cf. Hovda, 2008, p. 70). Therefore, in what follows, I will often use the following notation:

**Definition 7.3.9 Join/sum/fusion (\( \vee \)).** For all \( X \subseteq D \), \( \vee X \) is the unique \( a_1 \) such that \( Fu(a_1, X) \).

\[ \cdot \text{ Occasionally, we will write } a_1 \vee a_2 \text{ for } \vee \{a_1, a_2\}. \]

Finally, since we stipulated that every domain \( D \) is finite, every structure \( \langle D, \preceq \rangle \) is atomic. Thus, the structures that we are interested in

\[ ^8 \text{This particular axiomatization is taken from Hovda (2008, p. 81). The version of fusion used here is what Hovda calls “type 1 fusion.”} \]
are those of atomistic CEM. We define the notion of an atom as follows:

**Definition 7.3.12** Atom. \( a_1 \in D \) is an atom iff there is no \( a_2 \in D \) such that \( a_2 < a_1 \).

- We write \( AT(D) \) for the set of atoms of \( \langle D, \leq \rangle \).

In other words, the expressions in our language will denote in structures that are complete atomic boolean algebras minus the bottom element, also known as join-semilattices, as shown in Figure 7.1.

Finally, we can observe that, in atomistic CEM, there is a very simple condition on the identity of individuals: two individuals are identical just in case they have the same atoms.

**Proposition 7.3.13** Simons’s (1987) SF8 (p. 87). For all \( a_1, a_2 \in D \),

- \( a_1 = a_2 \) iff for all atoms \( a_1, a_1 \preceq a_1 \iff a_1 \preceq a_2 \).

We are now ready to give a classical semantics for the language described above.

As in DelTCS, the interpretation of singular predicates is relativized to CCs, i.e. distinguished subsets of the domain. Furthermore, we stipulate that both the denotations of singular predicates and the CCs according to which they are interpreted are restricted to the set of atoms of \( D (AT(D)) \).

**Definition 7.3.14** Classical model. A c-model is a tuple \( \mathcal{M} = \langle D, \preceq, \llbracket \cdot \rrbracket \rangle \), where \( \langle D, \leq \rangle \) is a model structure and \( \llbracket \cdot \rrbracket \) is a function satisfying:

1. If \( a \) is a constant, then \( \llbracket a \rrbracket \in D \).
2. If \( P \) is a singular predicate and \( X \subseteq AT(D) \), then \( \llbracket P \rrbracket_X \subseteq AT(D) \).

\(^9\) Where identity and proper part are defined as follows:

**Definition 7.3.10** Identical (=). For all \( a_1, a_2 \in D \), \( a_1 = a_2 \) iff \( a_1 \preceq a_2 \) and \( a_2 \preceq a_1 \).

**Definition 7.3.11** Proper part (\( \preceq \)). For all \( a_1, a_2 \in D \), \( a_1 \prec a_2 \) iff \( a_1 \preceq a_2 \) and \( a_1 \neq a_2 \).
So far, the interpretation follows the set-up described in Chapters 2–5. However, we now have many more kinds of expressions in our language than in the simple DelTCS system. Therefore, I will propose semantic interpretations for the additional predicates and operators that were introduced in the previous section. As I mentioned, the analysis of the (classical) semantics of distributive predicates will follow that of Link (1984). A major insight of Link’s paper is to propose that there exists a non-arbitrary relationship between the interpretation of a singular predicate and the interpretation of its plural counterpart. More precisely, he proposes that plural distributive predicates (e.g. *are asleep, are tall etc.*) are derived from singular predicates \( P \) through the addition of a * operator, which generates all the individual sums/joins of members of the extension of \( P \). In addition to a main verbal/adjectival predicate pluralizer \( * \), we have an nominal pluralizer \( * \) in the language. I have separated the two operators because it makes the syntax of our small language much simpler, but we assign to them exactly the same interpretation, as shown in Definition 7.3.15. In other words, the idea is that these two operators should be thought of as two separate occurrences of an interpretable [+plur] marking.

\[
(21) \quad \text{Girls are asleep}
\]

\[
\text{Girls}_+[\text{plur}] \quad \text{asleep}_+[\text{plur}]
\]

At a basic level, the interpretation of \( P^* \) or \( P^* \) at a CC \( X \subseteq AT(A) \) is just the closure of the interpretation of \( P \) at \( X \) under join, and a consequence of this analysis is that distributive predicates, themselves, will have a join-semilattice structure.

**Definition 7.3.15 Interpretation of plurality \((*/*)\).** For all singular predicates \( P \) and \( X \subseteq AT(D) \),

1. \( \llbracket P^* \rrbracket_X = \{ a_1 : Fu(a_1, A), \text{ for some } A \subseteq \llbracket P \rrbracket_X \} \).
2. \( \llbracket P^* \rrbracket_X = \{ a_1 : Fu(a_1, A), \text{ for some } A \subseteq \llbracket P \rrbracket_X \} \).

In addition to (stage-level) distributive plural predicates, we can combine English bare plurals and conjunctions with (stage-level) complex collective predicates as shown in (22).

\[10\] Observe that \( \llbracket P^* \rrbracket_X \not\subseteq X \), but this is not problematic since (in this analysis) the denotation of a plural distributive predicate is just dependent on the CC chosen for the singular predicate.
(22) a. Girls met.
b. Girls gathered.
c. John and Mary met.
d. John and Mary gathered.

Recall that these predicates are inappropriate when applied to singular subjects.

(23) a. # Mary met.
b. # Mary gathered.

Based on examples like (23), I propose (as is common in the literature on these predicates) that the interpretation of complex collective predicates (i.e. predicates of the $Q$ series) is limited to the non-atomic members of $D$. Correspondingly, the CCs according to which these predicates are evaluated are not subsets of atoms, but rather subsets of non-atomic individuals. For readability, I will notate these CCs using the $Y$ series.

**Definition 7.3.16** Complex collective predicates. If $Q$ is a complex collective predicate and $Y \subseteq D - AT(D)$,

- $\llbracket Q \rrbracket_Y \subseteq D - AT(D)$.

Finally, we need to assign some interpretation to the existential expression $\exists$. In the spirit of Generalized Quantifier Theory (GQ theory: Barwise and Cooper, 1981; Keenan and Stavi, 1986, among others), I analyze the expression $\exists$ as denoting a determiner in GQ theory; that is, a relation between properties or, alternatively, a function from a property to a generalized quantifier, which itself is a function from a property to a truth value. In particular, two properties are in the $\exists$ relation just in case their intersection is non-empty. Furthermore, in order to model the vague predication that existential DP constituents give rise to (argued for in Section 7.2), the interpretation of $\exists$ will also be relativized to a CC on its second argument. In parallel to the CCs associated with singular and complex collective predicates, the CCs associated with generalized quantifiers (i.e. DP denotations) will be made up of possible values for the quantifiers’ argument: sets of properties. These sets of properties should be thought of as constituting the range of relevant possible cases or situations\footnote{Indeed, to keep things (relatively) simple and the parallelism between the nominal and adjectival domains complete, I am assuming that CCs associated with DPs are made up of...} that might influence whether or not the...
main predicate denotation would be categorized as satisfying the DP subject denotation. Thus, $\exists$ will be analyzed as a function that takes a pluralized predicate ($P^*$, whose interpretation is itself relativized to a distinguished CC made up of atomic individuals) and a CC ($Z$, which is a distinguished subset of the powerset of the domain), and yield the family of properties in $Z$ whose intersection with $P^*$ is non-empty. This is stated more formally in Definition 7.3.17.

**Definition 7.3.17 Existential quantifier ($\exists$)** If $P$ is a singular predicate, $X \subseteq AT(D)$ and $Z \subseteq \mathcal{P}(D)$,

$$\exists P^* \downarrow_{Z, X} = \{ A : A \in Z \& A \cap [\exists P^*]_X \neq \emptyset \}$$

Observe that, unlike in basic DelTCS, formulas can now be evaluated with respect to more than one CC. In fact, the number of CCs with respect to which a formula $\phi$ will be evaluated corresponds exactly to the number of $P$-vague expressions in $\phi$. As a convention, so that we know exactly which CC is associated with which predicate, I will order the CCs linearly. Thus, for a formula $\exists P^*_1(P^*_2)$ with three $P$-vague constituents, we write $\exists P^*_1(P^*_2) \downarrow_{Z, X_1, X_2}$ to indicate that the whole DP $\exists P^*_1$ is interpreted with respect to $Z \subseteq \mathcal{P}(D)$, $P^*_1$ is interpreted with respect to $X_1 \subseteq AT(D)$ and $P^*_2$ is interpreted with respect to $X_2 \subseteq AT(D)$.

Thus, we now arrive at a typology of CCs for different kinds of constituents as in Table 7.3.

We now assign interpretations to formulas as follows:

**Definition 7.3.18 Interpretation of formulas** For all constants, predicates, $X, X_1, X_2 \subseteq AT(D)$, $Y \subseteq D - AT(D)$, and $Z \subseteq \mathcal{P}(D)$,

<table>
<thead>
<tr>
<th>Table 7.3 Typology of Comparison Classes in M-TCS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expression</strong></td>
</tr>
<tr>
<td>Singular predicate</td>
</tr>
<tr>
<td>Complex collective predicate</td>
</tr>
<tr>
<td>Determiner phrase</td>
</tr>
</tbody>
</table>

properties of individuals. Another intriguing possibility would be to pursue a mereological event semantics extension of the DelTCS system and have CCs composed of events rather than properties. I leave this option to future work.
1. \[ [P(a)]_X = \begin{cases} 1 & \text{if } [a] \in [P]_X \\ 0 & \text{if } [a] \in X - [P]_X \\ 1 & \text{if } [a] \in [P^*]_X \\ 0 & \text{if } [a] \in (X - [P]_X)^* \\ 1 & \text{otherwise} \\ 0 & \text{otherwise} \\ \end{cases} \]

2. \[ [P^*(a)]_X = \begin{cases} 1 & \text{if } [a] \in [P^*]_X \\ 0 & \text{if } [a] \in (X - [P]_X)^* \\ 1 & \text{otherwise} \\ 0 & \text{otherwise} \\ \end{cases} \]

3. \[ [Q(a)]_Y = \begin{cases} 1 & \text{if } [a] \in [Q]_Y \\ 0 & \text{if } [a] \in Y - [Q]_Y \\ 1 & \text{otherwise} \\ 0 & \text{otherwise} \\ \end{cases} \]

4. \[ [\exists P^*_1(P^*_2)]_{Z,X_1,X_2} = \begin{cases} 1 & \text{if } [P^*_1]_{X_1} \in [\exists P^*_1]_{Z,X_1} \\ 0 & \text{if } [P^*_1]_{X_1} \in Z - [\exists P^*_1]_{Z,X_1} \\ 1 & \text{if } [Q]_Y \in [\exists P^*_1]_{Z,X_1} \\ 0 & \text{if } [Q]_Y \in Z - [\exists P^*_1]_{Z,X_1} \\ 1 & \text{otherwise} \\ 0 & \text{otherwise} \\ \end{cases} \]

5. \[ [\exists P^*_1(Q)]_{Z,X_1,Y} = \begin{cases} 1 & \text{if } [Q]_Y \in [\exists P^*_1]_{Z,X_1} \\ 0 & \text{if } [Q]_Y \in Z - [\exists P^*_1]_{Z,X_1} \\ 1 & \text{otherwise} \\ 0 & \text{otherwise} \\ \end{cases} \]

Some lines from Definition 7.3.17 deserve comment. Concerning 1.: again, as in DelTCS, \([P(a)]_X\) has an indefinite truth value if \(a\) is not even in \(X\). This can happen because \(a\) denotes an atomic individual that just happens to not be under consideration in the particular context, as in DelTCS. However, in M-DelTCS, \(a\) can also fail to be in \(X\) because \(a\) happens to denote a non-atomic individual. Thus, we are analyzing the strangeness of (24) as a kind of presupposition failure rather than a syntactic agreement failure. Nothing particular hinges on this analysis and examples like (24) could be ruled out completely if we slightly changed the syntactic categories and rules that we are using.

(24) # John and Mary is tall.

Concerning 2: the notation \((X - [P]_X)^*\) is meant to indicate the closure of the complement of the interpretation of \(P\) in \(X\) under join. Crucially, we want to make a distinction between sentences like (25-a) and (25-b). Supposing that Mary and Sue both identify as female and are very short people, we would like (25-a) to be false, while having (25-b) be indefinite.

(25) a. Mary and Sue are tall for female basketball players.
    b. # Mary and Sue are tall for male basketball players.
Interpretations 4 and 5 are exactly parallel to 1, except that now the pluralized predicate is the argument of a higher order existential quantifier.

To make the relation between the properties of the logic being developed and examples of natural language more explicit, consider how we might compositionally derive the classical interpretation of a sentence such as (26) containing two P-vague predicates empty and cup\(^{\text{12}}\) and an existential determiner.

(26) Some cups are empty.

Suppose, as suggested in Definition 7.3.14, that \([-\mathbf{s}] = [\ast]\) and that \([\text{are}] = [\ast]\). Then the full sentence will be evaluated with respect to CCs provided by empty and cup, as well as a higher order CC provided by the existential determiner. Then, the composition proceeds as in Figure 7.2.

7.3.2.1 Constraints on \([\cdot]\) As mentioned in Chapter 4, the definitions in 7.3.17 are extremely weak if we do not restrict the application of the predicates in some way. In Chapters 2–5, we adopted van Benthem's (1982, 1990) axiom set to characterize how relative predicates are applied across CCs. We will adopt them again here for the interpretation of singular predicates, as shown in Axioms 7.3.1–7.3.3; although I highlight again that this is a particular analysis of the constraints obeyed by relative predicates and (infinitely many) others can be adopted within the general architecture that I am developing here.

For all \(a_1, a_2 \in AT(D)\) and \(X \subseteq AT(D)\) such that \(a_1 \in [P]_X\) and \(a_2 \notin [P]_X\).

**Axiom 7.3.19 No reversal (NR):** There is no \(X' \subseteq AT(D)\) such that \(a_2 \in [P]_{X'}\) and \(a_1 \notin [P]_{X'}\).

\(^{12}\) For example, at exactly which width/height ratio does a cup become a bowl? See Labov (1973) for experimental results concerning the vagueness of cup and bowl.
Axiom 7.3.20 **Upward difference (UD):** For all $X'$, if $X \subseteq X'$, then there is some $a_3, a_4 : a_3 \in \mathcal{P} X'$ and $a_4 \notin \mathcal{P} X'$.

Axiom 7.3.21 **Downward difference (DD):** For all $X'$, if $X' \subseteq X$ and $a_1, a_2 \in X'$, then there is some $a_3, a_4 : a_3 \in \mathcal{P} X'$ and $a_4 \notin \mathcal{P} X'$.

Observe that (at least some) complex collective predicates can be context sensitive. For example, consider the predicate *meet* in a situation in which we have a group of four professors in a department that is looking to appoint a new head, which university policy mandates is done through voting at a faculty meeting in which each professor has to be physically present. Thus, suppose one of the professors is out of town and can only be present via Skype. In this situation, we might say that (27) is false.

(27) The professors met.

This being said, suppose we slightly change the context and the professors are simply getting together to discuss who they would like to elect head next week when they are all back in town. In this case, since it does not matter in the context whether the participants in the meeting event are physically contiguous, we might consider (27) to be true even if one of the professors is on Skype. In other words, we see contextual variation in what counts as ”meeting”: how close the participants in the event have to be, through what medium the action takes place, etc. Nevertheless, despite the existence of contextual variation, a predicate like *meet* still displays a certain amount of monotonicity when it comes to its application across CCs. For example, suppose we are in a context in which *meet* does not hold of a particular plural individual $a_1$ because the atoms that make up the individual do not share a physical location and there is another individual $a_2$, whose atoms do share a physical location, of which *meet* holds. Since $a_2$ is a better example of *meeting* than $a_1$,\(^\text{13}\) it seems reasonable to think that, while there may be contexts in which *meet* could be applied to both $a_1$ and $a_2$, there will never be any contexts in which *meet* holds of $a_2$ but not of $a_1$. Thus, some principle such as *No reversal* also appear to be at work with complex collective predicates.

\(^\text{13}\) Although I believe that the discussion here gives a decent idea of the kind of contextual variation that we see with complex collective verbal predicates, when we move to the verbal domain, it often is more useful to speak in terms of verbal predicates holding of events (in the sense of Davidson, 1967; Parsons, 1990, among many others). In this work, I do not give an event semantics extension of M-DelTCS; however, I believe that such an extension is desirable for modelling context sensitivity and gradability within the verbal domain.
I therefore adopt the same constraints to characterize relative complex collective predicates such as *meet* and *gather* as with RAs like *tall*.

For all $a_1, a_2 \in D - AT(D)$ and $Y \subseteq D - AT(D)$ such that $a_1 \in \llbracket Q \rrbracket_Y$ and $a_2 \notin \llbracket Q \rrbracket_Y$,

**Axiom 7.3.22 NR:** There is no $Y' \subseteq D - AT(D)$ such that $a_2 \in \llbracket Q \rrbracket_{Y'}$ and $a_1 \notin \llbracket Q \rrbracket_{Y'}$.

**Axiom 7.3.23 UD:** For all $Y'$, if $Y \subseteq Y'$, then there is some $a_3, a_4 : a_3 \in \llbracket Q \rrbracket_{Y'}$ and $a_4 \notin \llbracket Q \rrbracket_{Y'}$.

**Axiom 7.3.24 DD:** For all $Y'$, if $Y' \subseteq Y$ and $a_1, a_2 \in Y'$, then there is some $a_3, a_4 : a_3 \in \llbracket Q \rrbracket_{Y'}$ and $a_4 \notin \llbracket Q \rrbracket_{Y'}$.

Finally, we must consider the CCs associated with existential DPs like $\exists P^*$. Do we need to impose some constraints on how the value of $\llbracket \exists P^* \rrbracket$ is assigned across CCs, or does the way that the interpretation of this expression is defined restrict its contextual variability?

What we see is that the classical semantic denotations of existential DPs already show a similar context sensitivity pattern to that of the classical semantic denotation of AAs and NSs. Recall that it has been proposed that AAs and NSs obeyed what I called the Absolute Adjective Axiom (AAA), repeated in **Axiom 7.3.25**, which forced the classical denotations of these predicates to be constant across CCs.

**Axiom 7.3.25 AAA.** For all absolute and non-scalar predicates $Q$, all interpretations $\llbracket \cdot \rrbracket$, all $X \subseteq D$ and $a_1 \in X$,

1. If $\llbracket Q(a_1) \rrbracket_X = 1$, then $\llbracket Q(a_1) \rrbracket_D = 1$.
2. If $\llbracket Q(a_1) \rrbracket_D = 1$, and $\llbracket Q(a_1) \rrbracket_X \neq i$, then $\llbracket Q(a_1) \rrbracket_X = 1$.

With adjectival predicates, the AAA is an axiom: it does not follow from any definitions or other features of the logic. However, we have adopted a very particular definition for the interpretation of $\exists$, namely one that is parallel to the definition of the existential quantificational determiner

---

14 Again, complex collective predicates can belong to different classes. For example, the predicate *be alike* appears to be total (as shown by its compatibility with modifiers like *completely* and *loosely speaking* (iii)), while the predicate *disagree* appears to be associated with a fully closed scale, as shown by compatibility with modifiers like *completely* and *slightly* (iv).

(iii) a. Mary and John are completely alike.
   b. Mary and John are alike, loosely speaking.

(iv) a. You and I completely disagree.
   b. You and I slightly disagree.
in generalized quantifier theory. This definition has the effect that we do not need some DP correspondent to the AAA. Rather, the statements that form part of the AAA in the adjectival domain can be directly proved as theorems for the DP domain with existential DPs:

**Theorem 7.3.26 Absolute Existential Theorem.** For all singular predicates $P_1, P_2$, all $X_1, X_2 \subseteq AT(D)$, and all $Z \subseteq P(D)$,

1. If $\exists P_1^*(P_2^*)||_Z X_1, X_2 = 1$, then $\exists P_1^*(P_2^*)||_{P(D)} X_1, X_2 = 1$.
2. If $\exists P_1^*(P_2^*)||_{P(D)} X_1, X_2 = 1$ and $\exists P_1^*(P_2^*)||_Z X_1, X_2 \neq i$, then $\exists P_1^*(P_2^*)||_Z X_1, X_2 = 1$.

Thus, we see that we already derive the context-independence of the classical interpretations of existential DPs directly from their definitions.

The system presented above also gives us another set of results associated with how distributive plural predicates are related to their singular counterparts. As shown by Link (1984) within his *Logic of Plural and Mass Nouns* system, a pluralized distributive predicate holds of a plural individual $a_i$ just in case the singular version of the predicate holds of all the atoms underneath $a_i$.

**Theorem 7.3.27 Classical distributivity.** For all constants $a_i$, singular predicates $P$ and $X \subseteq AT(D)$,

- $\| P^*(a_i) \|_X = 1$ iff $\| P(a_i) \|_X = 1$, for all atoms $a_2 \leq a_1$.

Finally, I turn to the question of the scales associated with classical denotations of predicates. Recall that in the version of Delineation semantics (DelS) that I presented in Chapter 4, the scale associated with a predicate was defined as follows:

---

15 **Proof.** 1. Let $X \subseteq AT(D)$ and let $Z \subseteq P(D)$. Suppose $\exists P_1^*(P_2^*)||_Z X_1, X_2 = 1$. Then $\| P_1^* \|_X, \| P_2^* \|_X \in Z$ and $\| P_1^* \|_X \cap \| P_2^* \|_X \neq \emptyset$. Since $Z \subseteq P(D)$, $\| P_1^* \|_X, \| P_2^* \|_X \in P(D)$. Since $\| P_1^* \|_X \cap \| P_2^* \|_X \neq \emptyset$, by Definition 7.3.17, $\exists P_1^*(P_2^*)||_{P(D)} X_1, X_2 \neq 1$.

2. Suppose $\| P_1^*(P_2^*)||_{P(D)} X_1, X_2 = 1$ and $\exists P_1^*(P_2^*)||_Z X_1, X_2 \neq i$. Since $\| P_1^*(P_2^*)||_{P(D)} X_1, X_2 = 1$, $\| P_1^* \|_X \cap \| P_2^* \|_X \neq \emptyset$. Since $\| P_1^*(P_2^*)||_Z X_1, X_2 \neq i$, $\| P_2^* \|_X \in Z$. So, by Definition 7.3.17, $\exists P_1^*(P_2^*)||_Z X_1, X_2 = 1$. □

16 **Proof.** Suppose $\| P^*(a_i) \|_X = 1$ and let $a_2$ be an atom such that $a_2 \leq a_1$. Suppose, for a contradiction that $\| P(a_i) \|_X \neq 1$. Since $\| P^*(a_i) \|_X = 1$, by Definition 7.3.15, there is some set of atoms $A \subseteq P \subseteq X$ such that $Fu(a_1, A)$. Since $\| P(a_1) \|_X \neq 1$, $a_2 \neq A$. So $A \neq A \cup \{a_i\}$. Therefore, by Definition 7.3.15, $\nabla A \neq \nabla A \cup \{a_i\}$. However, since, by assumption, $a_2 \leq a_1$, and $\nabla A = a_1$, by the definition of Fusion, $\nabla A \cup \{a_i\} = a_1$. So $\nabla A = \nabla A \cup \{a_i\}$. Therefore $\| P(a_2) \|_X = 1$. □
Definition 7.3.28 Semantics for formulas with $>$. For all $X \subseteq D$, all individuals $a_1, a_2$, and predicates $P$,

$$\llbracket a_1 >_P a_2 \rrbracket_X = \begin{cases} 
1 & \text{if there is some } X' \subseteq D : \llbracket P(a_1) \rrbracket_{X',M} = 1 \text{ and } \\
\llbracket P(a_2) \rrbracket_{X',M} = 0 \\
0 & \text{otherwise}
\end{cases}$$

I have two remarks concerning how this definition translates into the DP domain. Firstly, observe that we no longer have a “comparative” symbol $>$ in our language for M-DelTCS. This is because, as summarized by (Wellwood et al., 2012, among others), DP comparatives have certain properties that make them more complicated than adjectival comparatives, and I believe that these properties suggest that it is reasonable to separate the scale associated with non-adjectival constituents from the syntactic expression of this scale using comparative morphology. How to give an appropriate analysis of DP comparatives within M-DelTCS is explored in Burnett (2014b, 2015). However, in this work, we will still use a very similar definition to 7.3.28 to obtain the scale associated with a predicate, just meta-linguistically. Secondly, Definition 7.3.28 defines an ordering between singular individuals with respect to a predicate. Now we want to associate orderings with plural predicates and even DPs, so we need a more general definition.

In the rest of the book, I therefore adopt the following definition of the scalar ordering relation:

Definition 7.3.29 Scales. For all individuals or properties $A_1, A_2$, and syntactic expressions $B$,

- $A_1 >_B A_2$ iff there is some appropriate CC $X$ such that $A_1 \in \llbracket B \rrbracket_X$ and $A_2 \notin \llbracket B \rrbracket_X$.

Since we adopted van Benthem’s axioms constraining the interpretation of singular and complex collective predicates, it is easy to see that the scales associated with the $>_P$s and $>_Q$s will be both non-trivial (i.e. there are models in which these relations distinguish more than two individuals) and open (i.e. they have no non-arbitrary endpoints). Furthermore, we can verify that these properties hold of plural distributive predicates just in case they hold of their singular counterparts.

Theorem 7.3.30 Plural open scales. Let $M$ be a c-model and let $P$ be a singular predicate such that $>_P$ is a non-trivial open scale in $M$. Then:
1. $>_{P^*}$ is also non-trivial (i.e. for some $a_1, a_2, a_3 \in D$, $a_1 >_{P^*} a_2 >_{P^*} a_3$.)

2. $>_{P^*}$ is also an open scale.

Since the definition of $^*$ is identical to the definition of $^*$, a corollary of Theorem 7.3.30 is that scales associated with plural NPs are also possibly non-trivial and open.

**Corollary 7.3.31 Open NP scales.** Let $M$ be a c-model and let $P$ be a singular predicate such that $>_P$ is a non-trivial open scale in $M$. Then:

1. $>_{P^*}$ is also non-trivial.
2. $>_{P^*}$ is an open scale.

**Proof.** Immediate from the definition of $^*$ and Theorem 7.3.30.

Finally, based on the fact that the interpretation of existential DPs satisfies the Absolute Existential Theorem (Theorem 7.3.26), we can show that, like AAs and non-scalar adjectives, the scales associated with existential DPs are trivial, i.e. do not distinguish between more than two properties. Observe that since existential DPs contain a context sensitive predicate ($P^*$), the scales associated with the whole DP must be parametrized to a single interpretation of $P^*$ at a CC $X$. Therefore, we write $>_{[\exists P^*]_X}$ to reflect this.

**Theorem 7.3.32** For all singular predicates $P$ and $X \subseteq AT(D)$, there is no model $M$ such that there are distinct properties $P_1, P_2, P_3 \subseteq D$ such that $P_1 >_{[\exists P^*]_X} P_2 >_{[\exists P^*]_X} P_3$.

**Proof.** Immediate from the definition of $>_{[\exists P^*]_X}$ and Theorem 7.3.26.

In sum, in the system so far, pluralized distributive predicates are associated with open classical scales just in case the classical scale associated with their singular counterpart is open, and existential bare plural DPs are associated with trivial classical scales, just like $AAP$s and $AA^T$s.

In Section 7.3.3, we define the tolerant and strict semantics for the expressions in the M-DelTCS language.

### 7.3.3 Tolerant/strict semantics

As we did in Chapters 2–5, we will extend our c-models to t-models through adding the $\sim$ function, as shown in Definition 7.3.33. In the adjectival domain, $\sim$ mapped singular predicates and atomic CCs to binary predicate-relative indifference relations on the members of the
classes. Now that we have more context sensitive expressions in the language (collective predicates and existential DPs), indifference relations will hold between members of the associated CCs.

**Definition 7.3.33 T-model.** A t(olerant) model is a tuple $M = \langle D, \preceq, \llbracket \cdot \rrbracket, \sim \rangle$, where $\langle D, \preceq, \llbracket \cdot \rrbracket \rangle$ is a c-model and $\sim$ function from pairs of predicates/DPs and CCs of the appropriate type such that:

1. If $P$ is a singular predicate, and $X \subseteq AT(D)$, then $\sim(P, X)$ is a reflexive binary relation (written $\sim_X^P$) on $X$.
2. If $Q$ is a collective predicate, and $Y \subseteq D - AT(D)$, then $\sim(Q, Y)$ is a reflexive binary relation ($\sim_Y^Q$) on $Y$.
3. If $\exists P^*$ is an existential DP and $Z \subseteq P(D)$, then $\sim(\exists P^*, Z)$ is a reflexive binary relation ($\sim_Z^{\exists P^*}$) on $Z$.

As in Chapters 2–5, we will impose other constraints on the definition of the indifference relations across CCs. A natural thing to do would be to carry through the constraints associated with different classes of singular predicates, summarized in Table 7.4, into the new system. However, to keep things as simple as possible, we now have only one set of singular predicates, and I have been assuming that these predicates come from the relative class. Thus, in what follows, I will assume that the $\sim_X^P$ relations are all reflexive and symmetric, and satisfy the Tolerant/Strict Convexity (TC/SC), Minimal Difference (MD), Granularity, and CP axioms. Furthermore, I assume that the same axioms characterize the indifference relations associated with collective predicates (the $\sim_Y^Q$ relations), even though, again, the English predicates in this plural predication class can be relative, absolute, or non-scalar.

<table>
<thead>
<tr>
<th>Axiom</th>
<th>RA</th>
<th>AA$^T$</th>
<th>AA$^P$</th>
<th>NS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflexivity (R)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tolerant Convexity (TC)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Strict Convexity (SC)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Granularity (G)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Minimal Difference (MD)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Contrast Preservation (CP)</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Symmetry (S)</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Total Axiom (TA)</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Partial Axiom (PA)</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Be Precise (BP)</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>
Exactly as before, we use the predicate’s basic semantic extension and the \(\sim\) relations to define tolerant and strict extensions of predicates as follows:

**Definition 7.3.34** Tolerant/strict extensions of predicates. For all singular predicates \(P\), collective predicates \(Q\), \(X \subseteq AT(D)\) and \(Y \subseteq D - AT(D)\),

1. \(\llbracket P \rrbracket^t_X = \{a_1 : a_1 \in X \text{ and } \exists a_2 \sim_X a_1 \text{ and } a_2 \in \llbracket P \rrbracket_X\}\).
2. \(\llbracket Q \rrbracket^t_Y = \{a_1 : a_1 \in Y \text{ and } \exists a_2 \sim_X a_1 \text{ and } a_2 \in \llbracket Q \rrbracket_Y\}\).
3. \(\llbracket P \rrbracket^s_X = \{a_1 : a_1 \in X \text{ and } \forall a_2 \sim_X a_1, a_2 \in \llbracket P \rrbracket_X\}\).
4. \(\llbracket Q \rrbracket^s_Y = \{a_1 : a_1 \in Y \text{ and } \forall a_2 \sim_X a_1, a_2 \in \llbracket Q \rrbracket_Y\}\).

Furthermore, we will interpret constants \((a_1, a_2, a_3 \text{ etc.})\) “crisply”; that is, their classical, tolerant, and strict denotations are identical.

**Definition 7.3.35** Tolerant/strict interpretation of constants. For all \(a_1 \in D\),

\(\llbracket a_1 \rrbracket^t = \llbracket a_1 \rrbracket = \llbracket a_1 \rrbracket^s\).

I now consider the tolerant and strict interpretations of distributive pluralized predicates: \(P^\ast\) and \(P^\star\). In Section 7.2, it was suggested that we should distinguish vagueness that arises through the presence of a vague adjectival verbal or nominal predicate (such as *tall*, *gather*, or *heap*) and vagueness that arises through the act of predication (i.e. the combination of an appropriately headed DP with an appropriately pluralized predicate). The idea is that, even if we have a precisely interpreted DP subject (i.e. we have removed the possibility of vague predication in examples such as (28)), the vagueness of *tall* and *heap* remains.

(28) a. Exactly three heaps are tall.
    b. All the heaps are tall.

Based on examples such as (28), we can conclude, then, that the tolerant truth of sentences with vague distributive predicates and non-vague plural subjects should be calculated on the basis of whether the predicate tolerantly applies to the subjects' atoms. In order to reflect the relationship between plural tolerant distributive predication and singular tolerant predication, I propose that indifference relations associated with pluralized distributive predicates are constructed out of the indifference relations associated with singular ones through closure under pointwise join, the binary operation over pairs defined in Definition 7.3.24.
Definition 7.3.36  Pointwise join. (→∨) For ⟨w, x⟩ and ⟨y, z⟩, ⟨w, x⟩→∨⟨y, z⟩ = ⟨w ∨ y, x ∨ z⟩.

Definition 7.3.37  ∼∗ / ∼⋆. For all P∗ and all X, ∼X P∗ is the closure of ∼X under →∨. Likewise, ∼X P⋆ is the closure of ∼X under →∨.

With this definition, ∼X P∗ inherits a variety of important properties from the singular indifference relation. For example, it is reflexive and symmetric if ∼X P is.

Theorem 7.3.38  For all P∗, if ∼X P is reflexive and symmetric, then so is ∼X P⋆.

Finally, I consider the tolerant and strict denotation of the existential DPs. Since these constituents are analyzed as denoting generalized quantifiers (i.e. properties of properties), again in parallel to adjectival properties, we will suppose that the elements that constitute the CCs with respect to which DPs (of all kinds) are evaluated. In this case, properties of singular and plural individuals are related by binary relations, which encode indifference with respect to how the property satisfies the DP. But, what exactly does it mean for two (pluralized) properties to be considered indifferent with respect to a DP? In what follows, I will adopt the following hypothesis (29) concerning the source of vague predication, which I call the **Mereological Structure Hypothesis** (MSH).

(29) Mereological Structure Hypothesis (MSH)

Vague predication is the result of indifference with respect to how a predicate affects the **subparts** of a plural individual.

I will return to the MSH and how it can be empirically tested again in Section 7.4; however, in the case of an existential bare plural like townspeople, the idea is that ∼Z townspeople relates properties that are consider roughly equivalent in the context for holding of at least one townsperson.

Proof. Let P be a singular predicate and X ⊆ AT(D). Reflexivity. Let a1 ∈ X* to show a1 ∼X P a1. Let A be the set of atoms under a1. By the reflexivity of ∼X P, for all a2 ∈ A, a2 ∼X P a2. By Definition 7.3.37, the pointwise join of all the pairs ⟨a2, a2⟩, for a2 ∈ A, is in ∼X P, i.e. ⟨A, A⟩ ∈ ∼X P. Since, by assumption and the atomicity of the domain, √A = a1, (a1, a1) ∈ ∼X P. Symmetry. Let ⟨a1, a2⟩ ∈ ∼X P to show ⟨a2, a1⟩ ∈ ∼X P. Call the set of atoms under a1, A and the set of atoms under a2, B. Because the domain is atomic, ⟨a1, a2⟩ = ⟨A, B⟩. Since ⟨a1, a2⟩ ∈ ∼X P, it is the pointwise join of some subset R of ∼X P. Since ∼X P is symmetric, the inverse of R, R⁻¹ is also a subset of ∼X P. Consider the pointwise join of R⁻¹: ⟨A, A⟩, a.k.a ⟨a2, a1⟩. By Definition 7.3.37, ⟨a2, a1⟩ ∈ ∼X P. □
We therefore define the tolerant and strict definitions of existential DPs in an exactly parallel manner to the tolerant/strict denotations of adjectival predicates in Definition 7.3.39, and we give the tolerant and strict interpretation of formulas as in Definition 7.3.40 and 7.3.41.

**Definition 7.3.39**  
For all singular predicates $P$, $X \subseteq AT(D)$ and $Z \subseteq P(D)$,

1. $[\exists P^\star_1 P^\star_2]_{Z,X} = \{ P_1 : P_1 \in Z \text{ and there is some } P_2 \sim_{Z,P^\star_1} P_1 \text{ and } P_2 \in [\exists P^\star_2]_{Z,X} \}.$
2. $[\exists P^\star_1 P^\star_2]_{Z,X} = \{ P_1 : P_1 \in Z \text{ and for all } P_2 \sim_{Z,P^\star_1} P_1, P_2 \in [\exists P^\star_2]_{Z,X} \}.$

**Definition 7.3.40**  
Tolerant interpretation of formulas  
For all constants, predicates, $X, X_1, X_2 \subseteq AT(D)$, $Y \subseteq D - AT(D)$, and $Z \subseteq P(D)$,

1. $[[P(a)]^t_X] = \begin{cases} 1 & \text{if } [a]^t \in [P]^t_X \\ 0 & \text{if } [a]^t \in X - [P]^t_X \\ 1 & \text{otherwise} \end{cases}$
2. $[[P^\star(a)]^t_X] = \begin{cases} 1 & \text{if } [a]^t \in [P^\star]^t_X \\ 0 & \text{if } [a]^t \in (X - [P]^t_X)^* \\ 1 & \text{otherwise} \end{cases}$
3. $[[Q(a)]^t_Y] = \begin{cases} 1 & \text{if } [a]^t \in [Q]^t_Y \\ 0 & \text{if } [a]^t \in Y - [Q]^t_Y \\ 1 & \text{otherwise} \end{cases}$
4. $[[\exists P^\star_1 P^\star_2]_{Z,X_1,X_2}] = \begin{cases} 1 & \text{if } [P^\star_2]^t_{X_1} \in [\exists P^\star_1]_{Z,X_1} \\ 0 & \text{if } [P^\star_2]^t_{X_1} \in Z - [\exists P^\star_1]_{Z,X_1} \\ 1 & \text{otherwise} \end{cases}$
5. $[[\exists P^\star_1 Q]_{Z,X_1,Y}] = \begin{cases} 1 & \text{if } [Q]^t_Y \in [\exists P^\star_1]_{Z,X_1} \\ 0 & \text{if } [Q]^t_Y \in Z - [\exists P^\star_1]_{Z,X_1} \\ 1 & \text{otherwise} \end{cases}$

**Definition 7.3.41**  
Strict interpretation of formulas  
For all constants, predicates, $X, X_1, X_2 \subseteq AT(D)$, $Y \subseteq D - AT(D)$, and $Z \subseteq P(D)$,

1. $[[P(a)]^s_X] = \begin{cases} 1 & \text{if } [a]^t \in [P]^t_X \\ 0 & \text{if } [a]^t \in X - [P]^t_X \\ 1 & \text{otherwise} \end{cases}$
2. \( \llbracket P^s(a) \rrbracket^*_X = \begin{cases} 
1 & \text{if } \llbracket a \rrbracket^t \in \llbracket P^s \rrbracket^*_X \\
o & \text{if } \llbracket a \rrbracket^t \in (X - \llbracket P \rrbracket^*_X)^* \\
i & \text{otherwise} 
\end{cases} \)

3. \( \llbracket Q(a) \rrbracket^*_Y = \begin{cases} 
1 & \text{if } \llbracket a \rrbracket^t \in \llbracket Q \rrbracket^*_Y \\
o & \text{if } \llbracket a \rrbracket^t \in Y - \llbracket Q \rrbracket^*_Y \\
i & \text{otherwise} 
\end{cases} \)

4. \( \llbracket \exists P^s_1 (P^s) \rrbracket^t_{Z,X_1,X_2} = \begin{cases} 
1 & \text{if } \llbracket P^s_2 \rrbracket^t_{X_3} \in \llbracket \exists P^s_1 \rrbracket^t_{Z,X_1} \\
o & \text{if } \llbracket P^s_2 \rrbracket^t_{X_3} \in Z - \llbracket \exists P^s_1 \rrbracket^t_{Z,X_1} \\
i & \text{otherwise} 
\end{cases} \)

5. \( \llbracket \exists P^s_1 (Q) \rrbracket^t_{Z,X_1,Y} = \begin{cases} 
1 & \text{if } \llbracket Q \rrbracket^t_Y \in \llbracket \exists P^s_1 \rrbracket^t_{Z,X_1} \\
o & \text{if } \llbracket Q \rrbracket^t_Y \in Z - \llbracket \exists P^s_1 \rrbracket^t_{Z,X_1} \\
i & \text{otherwise} 
\end{cases} \)

Again, Definitions 7.3.26 and 7.3.27 merit a quick note: I have defined a global notion of tolerant and strict truth for formulas that contain multiple context sensitive predicates. We will continue in this vein because it will make the system and the results much simpler. However, this is not necessary. In principle, we could allow \( \llbracket \cdot \rrbracket^{t,t,s} \) or \( \llbracket \cdot \rrbracket^{t,s,t} \) interpretations (or others); that is, interpretations where the predicates and DPs are interpreted strictly or tolerantly independently. So, for a formula like \( \exists P^s_1 (P^s), \exists P^s_1 (P^s) \rrbracket^{t,t,s}_{Z,X_1,X_2} \) would be (tolerantly, tolerantly, strictly) true just in case there is some property \( P_3 \in Z \) such that \( P_3 \) is indifferent from \( \llbracket P^s \rrbracket^*_X \) and \( P_3 \) affects least one tolerantly \( P_1 \) individual.

To give an English example, as shown in Figure 7.3, a \( t, t, s \) interpretation of the sentence *Some cups are empty* would be (tolerantly, tolerantly, strictly) true just in case there is some contextually relevant property (i.e. set of individuals) that is indifferent from the (contextually relevant) set of completely empty individuals (with respect to how many cup-ish things it affects), and that has a non-empty intersection with the set of (contextually relevant) tolerantly ‘cup’-ish individuals. In other words, this sentence is satisfied just in cases where there is at least one cup-ish object that is completely empty.

I will return to the question of the compositional tolerant and strict semantics of complex syntactic constituents in Section 7.4, where we look at the case of definite plurals in distributive contexts.
176 BEYOND THE ADJECTIVAL DOMAIN

Finally, we must put some constraints on how the \( \sim_{DP} \) relations can be established across CCs. Unless otherwise stated, these axioms will be general and apply to all DP constituents (including definite plural DPs). I will therefore simply write \( \sim_{DP}^Z \) for the appropriate axiom schemata.

The first group of axioms that we will adopt consists of straightforward generalizations of the constraints that I proposed characterize adjectival indifference relations across classes: granularity, MD, CP. These are general categorization and CP constraints, so it is reasonable to think that they would hold in a parallel manner in the DP domain as in the adjectival domain.

(30) **Granularity (G):**
For all \( Z \subseteq \mathcal{P}(D) \) and all \( P_1, P_2 \in Z \), If \( P_1 \sim_{DP}^X P_2 \), then for all \( Z' : Z \subseteq Z' \), \( P_1 \sim_{DP}^Z P_2 \).

(31) **Minimal difference (D):**
For all \( Z \subseteq \mathcal{P}(D) \) and all \( P_1, P_2 \subseteq D \), if \( P_1 >_{DP} P_2 \), then \( P_1 \not\sim_{DP}^{\{P_1,P_2\}} P_2 \).

(32) **Contrast preservation (CP):**
For all \( Z, Z' \subseteq \mathcal{P}(D) \), if \( Z \subset Z' \) and there are \( P_1, P_2 \in Z \) such that \( P_1 \not\sim_{DP}^Z P_2 \) and \( P_1 \sim_{DP}^Z P_2 \), then there is some \( P_3 \in Z' - Z \) such that \( P_1 \not\sim_{DP}^Z P_3 \).

The main differences between the constraint set that we will adopt for DPs and the one that we adopted for adjectives concerns the convexity axioms and the Total/Partial Axioms (TAs/PAs). In the adjectival domain, we had to state the TC/SC and TAs/PAs directly; however, now, given the MSH, we will prove properties of the tolerant and strict orderings associated with predicates from more basic principles that build off of the mereological relations between the individuals that make up the properties being related by \( \sim \). First of all, we need a way of referring to the maximal individuals that a property affects. We will do this
through using singleton witness for the denotations of the various DP
subjects.\footnote{Readers familiar with Generalized Quantifier Theory will see the parallels between the singleton witness defined here and Barwise and Cooper’s (1981) notion of witness set.}

**Definition 7.3.42 Singleton witness.** An individual $a_i$ is a singleton witness for a quantifier $D(P^*)$ (at CCs $X$ and $Z$) just in case $a_i \in [\llbracket P^* \rrbracket_X]$ and $\{a_i\} \in [\llbracket D(P^*) \rrbracket_{X,Z} \cup \{a_i\}]$.

So the singleton witnesses for an existential DP like *girls* will be any singular or plural individuals that are girls. Definite plural DPs like *the girls*, in contrast, will have only one singleton witness: the join of all the girls.

The main proposal underlying the MSH is that the indifference relations associated with DP constituents (rather than simple adjectives or NP constituents) “care” about part–structure relations. Thus, we might think that the $\sim$ relations associated with DPs that relate properties must be (at least minimally) determined by how those properties affect parts of the DP’s singleton witness(es). I therefore propose the following constraint, which I call *shared parts* (SP), that states that, if two properties are indifferent with respect to a DP, this is because they affect at least some of the same parts of a singleton witness for the DP.

\begin{equation}
\tag{33} \text{Shared parts (SP):}
\end{equation}

For all singleton witnesses for a DP, $a_i$, all $Z \subseteq \mathcal{P}(D)$ and all distinct $P_1, P_2 \in Z$,

If \( P_1 \sim_{DP}^Z P_2 \), then there is some $a_2 \preceq a_i$ such that $a_2 \in P_1$ and $a_2 \in P_2$.

Although the mereological constraint *shared parts* (33) is very basic, it is already enough to (correctly) predict that existential bare plurals, in contrast to AA\( P^* \)'s, are not potentially vague and are therefore not associated with non-trivial strict scales. Note that to make the relevant observation more perspicuous, we will make the simplifying assumption that the individual predicates in Theorem 7.3.43 are interpreted precisely, i.e. we will abstract away from predicate vagueness. The proof of Theorem 7.3.43 is given in the appendix to this chapter.

**Theorem 7.3.43 Precise existentials.** Let $P_1, P_2$ be predicates and let $X_1, X_2 \subseteq D$ be (atomic) CCs. Suppose, furthermore, that $\llbracket P_1 \rrbracket_{X_1}^X = \llbracket P_1 \rrbracket_{X_1}^X$ and $\llbracket P_2 \rrbracket_{X_2}^X = \llbracket P_2 \rrbracket_{X_2}^X$. Then, for all $Z \subseteq \mathcal{P}(D)$,

\[
\llbracket \exists P_1^*(P_2^*) \rrbracket_{Z,X_1,X_2} = \llbracket \exists P_1^*(P_2^*) \rrbracket_{Z,X_1,X_2} = \llbracket \exists P_1^*(P_2^*) \rrbracket_{Z,X_1,X_2}.
\]
7.3.4 Summary

This section presented a mereological extension of the framework developed in Chapters 2–5 and showed how it could be used to develop an analysis of the vagueness and gradability properties of existential bare plurals. Of course, since I argued that existential DPs were not P-vague, no great use was made of most of the structure of the logic, the DP-related CCs, and the “coherence” constraints applying to indifference relations that deal with both general properties of the categorization process and the MSH.

This will all change, however, when we examine the case of definite plural DPs in Section 7.4.

7.4 Definite plural DPs and maximality

This section proposes an analysis of the context sensitivity, vagueness, and scale structure properties of definite plurals. As discussed in Section 7.2, under certain linguistic and extra-linguistic conditions, these constituents can be vague, in the way that AA's like empty and straight can be vague. I will start by giving an analysis of sentences containing definites paired with distributive predicates (34-a) and collective predicates of the gather class (34-b).

(34)  a. The townspeople are asleep.
     b. The townspeople gathered in the hall.

I will then give an analysis of sentences with definite plurals paired with collective predicates of the numerous class (35), which, as we saw in Section 7.2, display a different vagueness and gradability pattern.

(35)  a. The townspeople are numerous.
     b. The townspeople are a group of 52.

7.4.1 Language and classical semantics

As for the analysis of existential subjects outlined in Section 7.3, we will analyze definite subject DPs as denoting generalized quantifiers (i.e. properties of properties). In the case of definite plural subjects, we assume that these expressions denote a particular subclass of quantifiers: Montagovian Individuals. We therefore add to the language of M-DeTCS a series of Montagovian Individuals: $A, A_1, A_2$. In
Table 7.5  M-DelTCS translations of English sentences with definite plural subjects

<table>
<thead>
<tr>
<th>English expression</th>
<th>M-DelTCS translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>The townspeople are asleep.</td>
<td>(A_5(P_6^*))</td>
</tr>
<tr>
<td>The townspeople met.</td>
<td>(A_3(Q_2))</td>
</tr>
</tbody>
</table>

Particular, for every singular or plural individual constant \(a_i\), there is a Montagovian Individual \(A_i\) in the language. Furthermore, Montagovian Individuals will combine directly with pluralized predicates \(P^*\) and collective predicates \(Q\) to form well-formed expressions of the form \(A(P^*)\) and \(A(Q)\) (see Table 7.5).

Like existential plural subjects, the interpretation of definite plural subjects will be relativized to CCs consisting of families of properties that may influence whether or not its argument satisfies the predicate. In the spirit of Montague’s (1974) analysis of proper names, then, we give the classical semantic interpretations of Montagovian Individuals as in Definition 7.4.1.

**Definition 7.4.1 Classical semantics for \(A\).** For all DP CCs \(Z \subseteq \mathcal{P}(D)\),

\[
\|A_1\|_Z = \{P : P \in Z \text{ and } a_1 \in P\}
\]

The Montagovian Individual \(A_1\) picks out the subset of its distinguished CC that contains \(a_i\). In other words, we are tying the interpretation of the expressions of the \(A\) series to the interpretation of the expressions of the \(a\) series.

The interpretations of well-formed formulas containing \(A\) expressions are given in an exactly parallel manner to those containing existential DP subjects, as shown in Definition 7.4.2.

**Definition 7.4.2 Interpretation of formulas with \(A\).** For all MIs \(A\), all predicates \(P, Q\), all \(X \subseteq AT(D)\), all \(Y \subseteq D - AT(D)\) and \(Z \subseteq \mathcal{P}(D)\),

1. \[
\|A(P^*)\|_{Z,X} = \begin{cases} 
1 & \text{if } \|P^*\|_X \in \|A\|_Z \\
0 & \text{if } \|P^*\|_X \in Z - \|A\|_Z \\
i & \text{otherwise}
\end{cases}
\]

2. \[
\|A(Q)\|_{Z,Y} = \begin{cases} 
1 & \text{if } \|Q\|_Y \in \|A\|_Z \\
0 & \text{if } \|Q\|_Y \in Z - \|A\|_Z \\
i & \text{otherwise}
\end{cases}
\]
7.4.1.1 Classical semantics results With these definitions, we can already show a couple of results concerning both the distributivity properties of definite plural subjects and the behaviour of their interpretations across comparative classes. Firstly, we can show that, like existential plural subjects, the classical denotations of definite plural subjects are invariant across CCS; that is, we can prove that they satisfy a version of the AAA, which is given as theorem 7.4.3. Furthermore, by reasoning parallel to that associated with the classical denotations of AAs and existential plural DPs, we can conclude that the scales associated with Montagovian Individuals (≥A) will be trivial.

Theorem 7.4.3 Absolute Definite Theorem. For all singular predicates \( P_2 \), all Montagovian Individuals \( A_i \), all \( X_z \subseteq AT(D) \), and all \( Z \subseteq D \),

1. If \( \mathcal{A}(P_2^*) \mid_{Z,X_z} = 1 \), then \( \mathcal{A}(P_2^*) \mid_{P(D),X_z} = 1 \).
2. If \( \mathcal{A}(P_2^*) \mid_{P(D),X_z} = 1 \) and \( \mathcal{A}(P_2^*) \mid_{Z,X_z} \neq 1 \), then
   \( \mathcal{A}(P_2^*) \mid_{Z,X_z} = 1 \). \(^{19}\)

Secondly, we can show that formulas containing an MI and a starred predicate are distributive, i.e. if the classical interpretation of an MI \( A_i \) holds of the classical interpretation of a pluralized predicate \( P^* \), then the corresponding classical interpretation of \( P \) holds of all the atoms underneath \( a_i \). This is shown by Theorem 7.4.4, which follows almost immediately from the distributivity of classical \( P^* \) (Theorem 7.3.27) and the definition of the classical interpretation of \( A \).

Theorem 7.4.4 Classical distributivity with MIs. \(^{20}\) For all singular predicates \( P \), Montagovian Individuals \( A_i \), \( X \subseteq AT(D) \) and \( Z \subseteq D \),

If \( \mathcal{A}_i(P^*) \mid_{Z,X} = 1 \), then for all atoms \( a_2 \preceq a_1 \), \( \mathcal{A}(P(a_2)) \mid_X = 1 \).

Thus, we predict that, on their classical interpretations, (36-a) entails (36-b).

(36) a. The townspeople are asleep.
   b. For all individuals \( x \), if \( x \) is one of the townspeople, then \( x \) is asleep.

\(^{19}\) Proof. 1. Let \( X_z \subseteq AT(D) \) and let \( Z \subseteq D \). Suppose \( \mathcal{A}(P_2^*) \mid_{Z,X_z} = 1 \). Then \( P_2^* \mid_{X_z} \in Z \) and \( a \in P_2^* \mid_{X_z} \). Since \( Z \subseteq D \), \( P_2^* \mid_{X_z} \in D \). Since \( a \in P_2^* \mid_{X_z} \), by Definition 7.4.1, \( \mathcal{A}(P_2^*) \mid_{P(D),X_z} = 1 \).

2. Suppose \( \mathcal{A}(P_2^*) \mid_{P(D),X_z} = 1 \) and \( \mathcal{A}(P_2^*) \mid_{Z,X_z} \neq 1 \). Since \( \mathcal{A}(P_2^*) \mid_{P(D),X_z} = 1 \), \( a \in P_2^* \mid_{X_z} \). Since \( \mathcal{A}(P_2^*) \mid_{Z,X_z} \neq 1 \), \( P_2^* \mid_{X_z} \in Z \). So, by Definition 7.4.1, \( \mathcal{A}(P_2^*) \mid_{Z,X_z} = 1 \).

\(^{20}\) Proof. Suppose \( \mathcal{A}_i(P^*) \mid_{Z,X} = 1 \). Then, by Definition 7.4.1, \( a \in \mathcal{A}(P^*) \mid_{X} \). Since \( P^* \mid_{X} \) is distributive (Theorem 7.3.27), for all atoms \( a_2 \preceq a_1 \), \( \mathcal{A}(P(a_2)) \mid_{X} = 1 \).
We are now ready to give the tolerant and strict semantic interpretations for formulas containing definite plural subjects.

7.4.2 Tolerant/strict semantics

In Section 7.2, I argued that definite plural DPs in distributive and gather-type contexts are potentially vague: while, in some situations, it may be necessary to be precise and have the predicate apply to the entire group indicated by the subject DP (as in Lasersohn’s sleep study example: *The subjects are asleep*), we can also find many contexts in which such sentences are used more loosely and in which we can construct Sorites arguments. Furthermore, there are reasons to think that, again as with AA’s, CCs, or at least some kind of comparison of alternatives, are important for determining what the tolerant interpretations of definite plural subjects pick out. For example, it has been suggested (by Brisson, 1998; Brogaard, 2007) that “the smaller the number of individuals in the domain the more likely it is that all the individuals are taken to satisfy the predicate (individually or collectively)” (Brogaard, 2007, p. 419). Additionally, in a recent experimental study on the “loose” use of definite DPs, Schwarz (2013) shows that both the proportion of a group satisfying the predicate and the way in which its members are presented (i.e. whether they are displayed contiguously or non-contiguously) are significant factors in determining whether or not participants will accept a loose use of a definite plural. I therefore suggest that CCs (or something similar) have important uses for the calculation of tolerant interpretations and scales associated with definite plurals in the same way that they are useful for the formal analysis of vagueness and gradability of AAs.

Again, in parallel to the existential plural subjects discussed in Section 7.3, we will assume that the \( \sim \) function maps every \( \mathcal{A} \) to a binary indifference relation between properties at every CC \( Z \), which we will notate \( \sim^Z_A \). Using the classical semantic denotations already defined and these indifference relations, we define the tolerant and strict interpretations of Montagovian Individuals as follows.

**Definition 7.4.5 Tolerant/strict interpretations of \( \mathcal{A} \).** Let \( \mathcal{A} \) be a Montagovian Individual and let \( Z \subseteq \mathcal{P}(D) \). Then,

1. \( \| \mathcal{A} \|_Z^t = \{ P_1: \text{there is some } P_2 \sim^Z_A P_1 \text{ and } P_2 \in \| \mathcal{A} \|_Z \} \).
2. \( \| \mathcal{A} \|_Z^s = \{ P_1: \text{for all } P_2 \sim^Z_A P_1, P_2 \in \| \mathcal{A} \|_Z \} \).
Finally, the tolerant and strict truth of formulas containing definite plural subjects is given in the predictable way:

**Definition 7.4.6** Tolerant/strict interpretation of formulas with $A$.
For all MIs $A$, all predicates $P$, $Q$, all $X \subseteq AT(D)$, all $Y \subseteq D - AT(D)$ and $Z \subseteq \mathcal{P}(D)$,

\[
\begin{align*}
1. \llbracket A(P^*) \rrbracket^l_{Z,X} &= \begin{cases} 
1 & \text{if } \llbracket P^* \rrbracket^l_X \in \llbracket A \rrbracket^l_Z \\
0 & \text{if } \llbracket P^* \rrbracket^l_X \in Z - \llbracket A \rrbracket^l_Z \\
i & \text{otherwise}
\end{cases} \\
2. \llbracket A(Q) \rrbracket^l_{Z,Y} &= \begin{cases} 
1 & \text{if } \llbracket Q \rrbracket^l_Y \in \llbracket A \rrbracket^l_Z \\
0 & \text{if } \llbracket Q \rrbracket^l_Y \in Z - \llbracket A \rrbracket^l_Z \\
i & \text{otherwise}
\end{cases}
\end{align*}
\]

\[
\begin{align*}
3. \llbracket A(P^*) \rrbracket^s_{Z,X} &= \begin{cases} 
1 & \text{if } \llbracket P^* \rrbracket^s_X \in \llbracket A \rrbracket^s_Z \\
0 & \text{if } \llbracket P^* \rrbracket^s_X \in Z - \llbracket A \rrbracket^s_Z \\
i & \text{otherwise}
\end{cases} \\
4. \llbracket A(Q) \rrbracket^s_{Z,Y} &= \begin{cases} 
1 & \text{if } \llbracket Q \rrbracket^s_Y \in \llbracket A \rrbracket^s_Z \\
0 & \text{if } \llbracket Q \rrbracket^s_Y \in Z - \llbracket A \rrbracket^s_Z \\
i & \text{otherwise}
\end{cases}
\end{align*}
\]

### **7.4.2.1 Constraints on $\sim_A$**

The axioms governing $\sim_{DP}$ relations that were proposed in Section 7.3 (*Minimal Difference, CP, Granularity, and Shared Parts*) are meant to be general constraints that either characterize similarity judgments associated with concepts of all syntactic categories (MD, CP, and G), or constraints that cash out the MSH, i.e. *Shared Parts*. Therefore, I assume that the $\sim_A$ relations are also constrained by these conditions. However, it is clear that the part structure of an individual puts more limitations on the tolerant/strict denotations of a definite plural subject across CCs than simply $SP$. In particular, it seems reasonable to think that there might also be some convexity principle at work in the DP domain, as in the adjectival domain, but one that is stated over part-structures.

In stating the appropriate convexity axiom, we will first need a notion of upper bounds of a property with respect to an individual:

**Definition 7.4.7** Upper bounds ($\uparrow_{a_i}$). For all properties $P_i$ and $a_i \in D$, 
$\uparrow_{a_i} (P_i) = \{ a_2 : a_2 \leq a_i \text{ and there is no } a_3 \in P_i \text{ such that } a_2 < a_3 < a_i \}$.
We can now state the following axiom that constrains the establishment of DP indifference relations across CCs.

(37) Mereological Convexity (MC):
Let $DP$ be a quantifier. Then, for all singleton witnesses of $DP$ $a$, all $Z \subseteq \mathcal{P}(D)$ and $P_1, P_2 \in Z$,
If $P_1 \sim_{DP}^Z P_2$ and there is some $P_3 \in Z$ such that $a_2 \leq a_3 \leq a_1$, for some $a_2 \in \uparrow_a (P_2), a_3 \in \uparrow_a (P_3)$ and $a_1 \in \uparrow_a (P_1)$, then $P_1 \sim_{DP}^Z P_3$.

The axiom in (37) says that if two properties are indifferent and they affect different subparts ($a_1 \prec a_2$) of a singleton witness for a quantifier, and there is some other property that affects a subpart of the singleton witness that lies in between $a_1$ and $a_2$, then that property is indifferent from the property affect $a_2$. Again, this constraint is importantly different from the TC and SC constraints that were proposed in Chapter 4 because the ordering relation according to which the indifference relations are proposed to be convex is the part–structure relation in the model ($\leq$), not the constructed tolerant/strict scales.

Additionally, we make a further very strong assumption that the contents of CCs (i.e. the properties that make them up) have to all contain properties that are comparable with respect to the singleton witness of a particular DP.\(^{21}\)

(38) Incomparability (I):
Let $DP$ be a quantifier. For all singleton witnesses of $DP$, $a_1$, and all $Z \subseteq \mathcal{P}(D)$ and $P_1 \in Z$, then
For all $P_2 \subseteq D$, if there is some $a_2 \in \uparrow_{a_1} (P_2)$ such that $a_2$ is incomparable to $a_3$ (for some $a_3 \in \uparrow_{a_1}$), then $P_2 \not\in Z$.

Our final axiom set governing indifference relations associated with DP constituents (of all classes) therefore consists of:

1. Reflexivity (R)
2. Minimal Difference (MD).
3. Contrast Preservation (CP).
4. Granularity (G).
5. Shared Parts (SP).

\(^{21}\) This condition is very strong but is used in the proofs of asymmetry and transitivity of the $\succ_A$ relations. I therefore take it to be an open question whether this axiom can be replaced with a better motivated one, while still preserving the ordering results presented in Section 7.5.
184  BEYOND THE ADJECTIVAL DOMAIN

Figure 7.4  Strict, tolerant, strict interpretation of The cups are empty

7. Incomparability (I).

Again, we can go through some English examples to show how the tolerant/strict compositional semantics would work. Note that we have been treating definite plurals as proper names in the logic, so we do not have a denotation for the definite determiner, but, for this example, we can simply assume that the definite determiner is restricted so that it must pick from the tolerant/classical/strict denotation of its NP complement. For example, English paraphrases of the corresponding interpretations of The cups are empty are shown in (39).

(39)  a. \([\text{The cups are empty}]:^s,t,s \approx \text{‘All the cups are clearly cups and are (at least) approximately empty.’}\)
    b. \([\text{The cups are empty}]:^t,s,t \approx \text{‘The relevant cups (which are clearly cups) are completely empty.’}\)
    c. \([\text{The cups are empty}]:^t,t,s \approx \text{‘All the cups are completely empty, but some are borderline cups.’}\)

The composition of (39-c) proceeds as shown in Figure 7.4.

Based on our definitions, this sentence is \((s,t,s)\) satisfied just in case the set of (relevant) completely empty individuals contains all the members of the group denoted by the cups, which may include individuals that are only tolerantly cups.

Note that the number of calculations and CCs grows incrementally with the number of determiners and context sensitive predicates. However, as natural language speakers, we may have heuristics that help us cut down on having to do all these calculations in real time when we speak. For example, Yoon (1996) observes that using an AA\(^P\) with a definite plural favors a “weak” (or what we would call a tolerant reading) of the definite description (40-a); whereas, using an AA\(^T\) with a definite description favors a “strong” (here, strict) reading (40-b).
definite plural DPs and maximality

(40) a. Are the toys dirty?
   b. Are the toys clean?

However, I leave the constraints/interactions between tolerant and strict readings of the subject and the predicate to future research.

7.4.2.2 Tolerant/strict results  With these constraints, we can now show that, contrary to the trivial tolerant scales associated with existential plural DPs, the tolerant scales associated with definite plural DPs, when the \( >^t_A \) relations are restricted to distributive or gather-collective predicates, can distinguish more than two individuals and, furthermore, have certain ordering properties. In particular, the tolerant scales associated with definite plurals (restricted to the appropriate classes of predicates)\(^{22}\) are (at least) asymmetric and transitive. These facts are stated as Theorems 7.4.8 and 7.4.9, and they are proved in the appendix for \( >^t_A \) restricted to distributive predicates, but the proof proceeds the same way with gather-collective predicates.

**Theorem 7.4.8** Asymmetry of \( >^t_A \): If \( A \) be a Montagovian Individual and \( >^t_A \) is restricted to distributive properties, then \( >^t_A \) is asymmetric.

**Theorem 7.4.9** Transitivity of \( >^t_A \): If \( A \) be a Montagovian Individual and \( >^t_A \) is restricted to distributive properties, then \( >^t_A \) is transitive.

This being said, Shared Parts will work the same way with definite plurals as it worked with existential plurals: although \( >^t_A \) can be articulated, the strict scale associated with a definite plural \( (>^s_A) \) is trivial, and the strict denotations of a definite plural are always identical to their classical ones, as shown by Theorem 7.4.10. Thus, we correctly predict that definite plurals are associated with a single scale: the one that is derived through looking at how their tolerant denotations change across CCs.

**Theorem 7.4.10** For all Montagovian Individuals \( A \) and \( Z \subseteq \mathcal{P}(D) \),\(^{23}\)

\[
\llbracket A \rrbracket^t_Z = \llbracket A \rrbracket_Z.
\]

Finally, now that we have an articulated ordering associated with a DP constituent, we can wonder about what its properties are. As shown in

\(^{22}\) Recall that, by virtue of the proposed interpretations of predicates in Section 7.3, the properties denoted by the predicates of difference classes have different structures: distributive properties are those that have a join-semilattice structure, while gather-collective properties can contain individuals that are proper subparts of other individuals and do not contain atoms.

\(^{23}\) The proof of Theorem 7.4.10 proceeds in parallel to the relevant portion of the proof of Theorem 7.3.43.
Theorem 7.4.11, it turns out that definite plurals DPs like *the townspeople* are predicted to be associated with top-closed tolerant scales, just like total $AA^T$s.²⁴

**Theorem 7.4.11** Let $A$ be a Montagovian Individual. Then $>^A$ is a top-closed scale.

I suggest that this is a welcome result, since there are good reasons to think that predicates like *empty* and *straight* have the same scale structure properties as definite plurals in distributive contexts. For example, if we consider a scalar modifier like French *tou(te)* ‘all’ which appears in both the DP and adjectival domains, we see that it combines with $AA^T$s to yield a maximal or completive interpretation (41).

(41) a. La salle est toute vide.
   The room is all empty
   ‘The room is completely empty.’
   b. Cette ligne est toute droite.
   This line is all straight
   ‘This line is completely straight.’

Furthermore, this element also combines with definite plural DPs to yield the exact same kind of interpretation.

(42) a. Tous les villageois sont endormis.
   All the townspeople are asleep
   ‘All the townspeople are asleep.’
   b. Tous les villageois se sont rassemblés.
   All the townspeople refl are gathered
   ‘All the townspeople gathered.’

We also find this basic pattern in English, but the use of *all* in the adjectival domain is much more restricted in this language (see Bolinger, 1972).

(43) This room is all empty.
(44) a. All the townspeople are asleep.
   b. All the townspeople gathered in the park.

²⁴ The proof precedes in an exactly parallel manner to the adjectival domain, since $A$ satisfies the Absolute Definite Theorem.
We might note at this point that *tout* also applies to $AA^P$'s in French, in which case, it creates an intensive, not a completive interpretation.\(^{25}\)

\[\begin{align*}
(45) & \quad \text{a. } \text{Le chat est } \textit{tout} mouillé. \\
& \quad \text{The cat is all wet} \\
& \quad \text{‘The cat is really wet.’} \\
& \quad \text{b. } \text{Ta robe est } \textit{toute} sale. \\
& \quad \text{You dress is all dirty} \\
& \quad \text{‘Your dress is really dirty.’}
\end{align*}\]

However, in contrast to the parallels that we find between $AA^T$'s and definite plural DPs, *tou(te)s* does not create an intensive interpretation with existential plurals; rather, applying *tou(te)s* to an existential plural DP creates ungrammaticality (46).

\[\begin{align*}
(46) & \quad \text{*Il y a tous des chiens dans la cour.} \\
& \quad \text{It there has all some dogs in the yard} \\
& \quad \text{Intended: ‘There are a lot of dogs in the yard.’}
\end{align*}\]

I suggest that the distributional and interpretative behavior of French *tou(te)s* gives us another empirical argument in favor of the analysis of existential bare plurals proposed above: unlike $AA^P$'s, which are associated with articulated scales that have no top endpoint, existential DPs are non-scalar predicates. As shown by Theorem 7.3.43, this result is directly predicted by the system developed in this chapter.

### 7.4.3 Gather predicates vs numerous predicates

The first part of this section gave an analysis of the vagueness and gradability properties of definite plural DPs paired with distributive and collective predicates of the *gather* class. However, there is another lexical class of plural predicates that give rise to different patterns. For example, consider the sentences in (47).

\[\begin{align*}
(47) & \quad \text{a. } \text{The girls are numerous.} \\
& \quad \text{b. } \text{The girls are a group of four.}
\end{align*}\]

\(^{25}\) The examples in (45) also have a “mereological” interpretation in which the modifier applies to the parts of the subject DP (i.e. ‘All of the cat is wet’ and ‘All of your dress is dirty’), but these are not the interpretations that are of interest to us here.
As observed by Dowty (1979), and discussed by many other authors such as Taub (1989), Brisson (1998), Winter (2001), Corblin (2008), Champollion (2011), the plural predicates *be numerous* and *be a group of four* must apply to the entire group referred to by the subject DP: there is no tolerance for exceptions (i.e. ?*The girls are a group of four... except Mary*).

In other words, vague predication is impossible with predicates of the same plural predication class as those in (47). Furthermore, unlike definite plurals in distributive and *gather*-collective contexts, which, I argued, are associated with articulated top-closed scales (as witnessed, among other things, by their ability to be modified by *tou(te)s/all*), these scales seem to disappear when the definite plurals are combined with *numerous*-collective predicates, as shown in (48) and (49).

(48)  a. ?All the girls are numerous.
     b. ?All the girls are a group of four.

(49)  a. Les filles sont quatres.
     The girls are four
     ‘The girls are four.’
     b. *Toutes les filles sont quatre.*
     All the girls are four
     ‘All the girls are four.’

We therefore have arrived at a puzzle, stated in (50):

(50)  **Linguistically conditioned potential vagueness and gradability puzzle:**

How can definite plural DPs be P-vague and associated with top-closed scales in some plural predication contexts, while being precise and non-scalar in other predication contexts?

In other words, what makes predicates like *numerous* and *be a group of four* different from *gather* and *be asleep?*

7.4.3.1 The Landman-Link proposal  Although it is most often exemplified by the predicate *numerous*, the class of what I will call, following Corblin (2008), holistic predicates is, in fact, quite large and varied.

---

26 Of course (47) is still a P-vague expression (i.e. ‘how big does a group have to be to be called *numerous*’); however, this is a straightforward case of predicate vagueness associated with the lexical item *numerous*, not vague predication.
(51) Examples of holistic predicates:
   a. Kroch (1974): be politically homogeneous, be a motley crew, suffice to defeat the king
   b. Dowty (1979): be numerous, be a large group, be a group of four, be few in number, be a couple, be denser in the middle of the forest.
   c. Taub (1989): pass the pay raise, elect Bush, return a verdict of “not guilty,” decide unanimously to skip class, eat up the cake, finish building the boat
   d. Brisson (1998): be too heavy to carry
   e. Winter (2001): be a good team, form a pyramid, constitute a majority, outnumber

Furthermore, there are many different (possibly non-mutually exclusive) analyses in the literature that aim to account for the differences between the predicates in (51) and collective predicates like *gather* and *meet*. For example, it has been suggested that, although they do not directly apply to atoms, *gather*-predicates make particular claims about the participation of atomic individuals in a collective event that holistic predicates do not make. In particular, Dowty (1979) proposes that *gather* (unlike *numerous*) has distributive subentailments. Similarly, Landman (2000) and Champollion (2011) propose that *gather* has thematic entailments, rather than non-thematic entailments (like *numerous*). Another idea, due to Taub (1989) and developed by Brisson (2003), is that the *gather*-collective/holistic predicate difference is due to an aspectual difference: *gather*-predicates are proposed activities or accomplishments, while *numerous*-predicates are states or achievements. Finally, recently Križ (2014) proposes that the heart of the difference between plural predicates lies in semantic underspecification. He proposes that *gather*-predicates have extension gaps that create homogeneity effects (which will be exemplified and further discussed in Section 7.4.4), while holistic predicates have no such extension gaps and correspondingly are not homogeneous.

Although all these ideas may have some merit to them (and may not, in fact, be mutually incompatible), in this work, I will adopt an analysis of the differences between *gather* and *numerous* that Winter

---

27 However, one difficulty with these approaches is that it is not so clear how to get an air-tight characterization of these subentailments/thematic entailments. As (Champollion, 2011, p. 344) says, “Thematic entailments are a slippery and ill-defined concept.”

28 Although, see Champollion (2011) for some potential counter-examples to this generalization.
(2001) calls the Landman–Link proposal. This idea is summarized in (52), and versions of it have been adopted by Link (1984), Landman (1989, 2000), Winter (2001, 2002), Champollion (2011), among others.

(52) **The Landman–Link proposal:**
Holistic predicates are composed of atoms; that is, the denotations of holistic predicates contain only individuals that, while plural, have no subpart structure.

More concretely, this involves proposing that there is a function that relates the denotation of a definite plural DP, which is based on an individual that has part-structure to another individual that has no structure. As Winter (2002, p. 99) says, "According to the Landman/Link proposal, the set that corresponds to the noun phrase **the students** can be mapped in the semantic analysis to the atom denoting the noun phrase **the school's basketball team**. Using this mapping, the sentence **the students are a good team** is interpreted as equivalent to the sentence **the school's basketball team** is a good team."

In this chapter, I will adopt a particular implementation of the Landman–Link proposal and show how, by doing so, we predict the correct vagueness and gradability properties of definite plurals paired with holistic predicates. This being said, I take it to be conceivable that other accounts incorporating some of the other ideas present in the literature may do just as well.

The first thing we do is add a new set of predicates to our language that will model holistic predicates in languages like English and French. These predicates will be notated using the $R$ class: $R, R_1, R_2, R_3$ etc. Furthermore, like other plural predicates, the $R$ predicates will combine with Montagovian Individuals to create formulas of the form: $\mathcal{A}(R)$ as shown in Table 7.6.

With respect to the semantics, we first add a separate domain of atomic pluralities to our model structure in which holistic predicates denote.

| Table 7.6 M-DeTCS translations of English sentences with definite plural subjects |
|--------------------------------------|------------------|
| **English expression**               | **M-DeTCS translation** |
| The townspeople are asleep.          | $\mathcal{A}_s(P_s^e)$ |
| The townspeople met.                 | $\mathcal{A}_s(Q_s)$  |
| The townspeople are numerous.        | $\mathcal{A}_s(R_s)$  |
Definition 7.4.12 T-model structure (revised). A t-model structure $\mathcal{M}$ is a tuple $\langle D, \preceq, A, \sim \rangle$, where $\langle D, \preceq \rangle$ is a classical extension mereology and $A$ is an unordered set of individuals whose cardinality is the same as $D$’s.

Holistic predicates are interpreted as subsets of $A$, relativized to CCs, which are themselves subsets of $A$, as shown in Definition 7.4.13. Furthermore, $\sim$ maps a singular CC $W$ and $R$ to a binary indifference relation on $W$ and a CC associated with a Montagovian Individual $U$ to an indifference relation on $U$.

Definition 7.4.13 Classical semantics for holistic predicates. For all holistic predicates $R$ and $W \subseteq A$, Montagovian Individuals $A$ and $U \subseteq \mathcal{P}(A)$,

1. $\llbracket R \rrbracket_W \subseteq W$.
2. $\sim^W_R$ is a binary relation the elements of $W$ that satisfies the relevant constraints, depending on the scale structure class of $R$.
3. $\sim^U_A$ is a binary relation on the elements of $U$ that satisfies the DP indifference relation constraints proposed above $(R, MD, CP, G, SP, MC, I)$.

Tolerant and strict denotations of holistic predicates are calculated as expected (i.e. $\llbracket R \rrbracket_W = \{ b_1 : \exists b_2 \sim^W_R b_1 \& b_2 \in \llbracket R \rrbracket_W \}$, and $\llbracket R \rrbracket_W = \{ b_1 : \forall b_2 \sim^W_R b_1, b_2 \in \llbracket R \rrbracket_W \}$). Since holistic predicates denote in a different domain than the Montagovian Individuals, we have to have some special way of calculating the interpretation of formulas of the form $\mathcal{A}_5(R_4)$. In order to do this, we take advantage of a proposal by Link (1984), namely that, as he says (in his terminology), “For every usual i-sum $a$, there is an atom $\gamma(a)$ uniquely representing the group consisting of $a$” (Link, 1984, p. 249). Within the framework developed here, if $a$ is a member of $D$, then $\gamma(a) \in A$. Additionally, we stipulate that $\gamma$ is a bijection (which is possible since, by assumption, $D$ and $A$ have the same cardinality). Furthermore, the definition of $\gamma$ can be extended pointwise to properties of individuals and generalized quantifiers as shown in (53).²⁹

(53) a. For all $X \subseteq D$, $\gamma(X) = \{ b : \gamma(a) = b, \text{for all } a \in X \}$
   b. For all $\mathcal{A} \subseteq \mathcal{P}(D)$, $\gamma(\mathcal{A}) = \{ P_1 : \gamma(P_2) = P_1, \text{for all } P_2 \in \mathcal{A} \}$

The switching of domains that is necessary for the interpretation of sentences with holistic predicates makes the calculation of the truth of

²⁹ I thank Viola Schmitt for discussion of the importance of using and extending $\gamma$. 
formulas containing Montagovian Individuals and holistic predicates a bit more complicated. Namely, classical, tolerant, or strict truth of a formula consisting of $A$ paired with $R$ will be evaluated through looking at the image of $A$ in $A$, as shown in Definition 7.4.14.

**Definition 7.4.14** Tolerant, Classical, Strict interpretation of formulas. Let $A$ be a Montagovian Individual, let $R$ be a holistic predicate, let $W \subseteq A$ and let $Z \subseteq \mathcal{P}(D)$.

1. $\llbracket A(R) \rrbracket_{Z,W}^t = \begin{cases} 1 & \text{if } \llbracket R \rrbracket_{W} \in \gamma(\llbracket A \rrbracket_{Z}) \\ 0 & \text{if } \llbracket R \rrbracket_{W} \in \gamma(Z) - \gamma(\llbracket A \rrbracket_{Z}) \\ i & \text{otherwise} \end{cases}$

2. $\llbracket A(R) \rrbracket_{Z,W}^t = \begin{cases} 1 & \text{if } \exists R_1 \sim_{\mathcal{A}} \gamma(Z) \llbracket R \rrbracket_{W}^{i}, R_1 \in \gamma(\llbracket A \rrbracket_{Z}) \\ 0 & \text{if } \llbracket R \rrbracket_{W} \in \gamma(Z) \& \neg \exists R_1 \sim_{\mathcal{A}} \gamma(Z) \llbracket R \rrbracket_{W}, R_1 \in \gamma(\llbracket A \rrbracket_{Z}) \\ i & \text{otherwise} \end{cases}$

3. $\llbracket A(R) \rrbracket_{Z,W}^t = \begin{cases} 1 & \text{if } \forall R_1 \sim_{\mathcal{A}} \gamma(Z) \llbracket R \rrbracket_{W}^{i}, R_1 \in \gamma(\llbracket A \rrbracket_{Z}) \\ 0 & \text{if } \llbracket R \rrbracket_{W} \in \gamma(Z) \& \neg \forall R_1 \sim_{\mathcal{A}} \gamma(Z) \llbracket R \rrbracket_{W}, R_1 \in \gamma(\llbracket A \rrbracket_{Z}) \\ i & \text{otherwise} \end{cases}$

With these definitions, in addition to having an analysis of holistic predication with definite plural subjects in M-DelTCS, we can show that without adding any axioms beyond what we already proposed, the interpretation of formulas containing holistic predicates is invariant across CCs (Theorem 7.4.15, proved in the appendix), and, therefore, the scales associated with definite plural subjects (restricted to holistic properties) are trivial. In particular, the proof goes through because *Shared Parts* prohibits the establishment of indifference relations between distinct properties that affect only atomic individuals.

**Theorem 7.4.15** Precise holistic predication. Let $A$ be a Montagovian Individual and let $R$ be a holistic predicate. Then, for all $W \subseteq A$ and $Z \subseteq \mathcal{P}(D)$,

$$\llbracket A(R) \rrbracket_{Z,W} = \llbracket A(R) \rrbracket_{Z,W}^t = \llbracket A(R) \rrbracket_{Z,W}^s.$$
I therefore conclude that the variability in vagueness and gradability puzzle can be given a simple and elegant solution within the M-DelTCS framework developed in this chapter.

7.4.4 Negation and homogeneity with definite plurals

This chapter concludes with a small discussion of the interaction between sentential negation and definite plural DPs within the framework here. As it stands, we have not yet included a syntactic negation in the language of M-DelTCS. A main reason for this omission is that, as observed since Fodor (1970), negation behaves differently in the DP domain than in the adjectival domain. In particular, negation with definite plurals and distributive predicates gives rise to what is known as a homogeneity effect, which can be roughly characterized as the appearance of an extension gap. For example, (54) entails (or at least strongly implies) that for all $x \in \{\text{the townspeople}\}$, $x$ is not asleep. This sentence does not seem to mean that it is not the case that all the townspeople are asleep, as would be expected if we adopted an analysis of sentential negation with definite plural DPs along the lines proposed in Chapter 6 for adjectival negation, i.e. as a complement operator.

(54) The townspeople are not asleep.

Additionally, homogeneity has an important effect on how definite plurals display the characteristic properties of potential vagueness. In contrast to P-vague adjectives, which can occasionally allow borderline contradictions (55) (Ripley, 2011, among others), we cannot construct such acceptable contradictions with instances of vague predication with definites (56).

(55) Mary is tall and not tall. (OK when Mary is borderline tall.)
(56) #The townspeople are asleep and the townspeople are not asleep.
    (# even if we allow non-maximality.)

One way to account for the differences between DPs and adjectives in this way is to simply encode homogeneity into the classical semantics of negative sentences through the definition leaving an extension gap, as shown in Definition 7.4.9.

---

30 I thank Manuel Kriš for drawing my attention to this point. See also Kriš (2014).
Definition 7.4.9 Homogeneity of definite plurals:

\[
\lll -A(P^n) \rrl_{Z,X} = \begin{cases} 
1 & \text{if } [P^n]_X \in Z \text{ and there is no atom} \\
& a_i \preceq a : [P(a_i)]_X = 1 \\
0 & \text{if } [P^n]_X \in Z \text{ and there is some atom} \\
& a_i \preceq a : [P(a_i)]_X = 1 \\
i & \text{otherwise}
\end{cases}
\]

However, I suggest that it would be desirable to derive the homogeneity effects with definite plurals and their consequences for borderline contradictions with these constituents from more general principles that distinguish the DP and adjectival domains. I will therefore leave integrating negation and homogeneity effects more fully into this system to future work.

7.5 Conclusion

This chapter presented a mereological extension of the DelTCS system for the analysis of vagueness, context sensitivity, and gradability in the DP domain. I argued that we could arrive at elegant and empirically accurate analyses of the properties of existential bare plural subjects and definite DP subjects using this framework and appropriately modified versions of the axioms that we assumed characterized the establishment of

<table>
<thead>
<tr>
<th>Axiom</th>
<th>DP</th>
<th>$AA^T$</th>
<th>$AA^P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflexivity (R)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Granularity (G)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Minimal Difference (MD)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Contrast Preservation (CP)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Total Axiom (TA)</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Partial Axiom (PA)</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Tolerant Convexity (TC)</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Strict Convexity (SC)</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mereological Convexity</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Shared Parts</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Incomparability</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>
of indifference relations across CCs in the adjectival domain. A comparison of the axioms proposed for adjectives with those proposed for DPs is shown in Table 7.7.

I argued that this analysis captures both certain important similarities between the scale structure of AA\(\uparrow\)s and definite plural DPs in distributive contexts, as well as differences between AA\(\uparrow\)s and existential plurals. I therefore conclude that the DeITCS framework has useful applications for the analysis of vagueness and gradability across syntactic domains.

7.6 Appendix: Longer proofs

Theorem 7.6.1 Precise existentials. Let \(P_1, P_2\) be predicates and let \(X_1, X_2 \subseteq D\) be (atomic) CCs. Suppose furthermore that \(\llbracket P_1 \rrbracket_{X_1}^I = \llbracket P_1 \rrbracket_{X_1}^I\) and \(\llbracket P_2 \rrbracket_{X_2}^I = \llbracket P_2 \rrbracket_{X_2}^I\). Then, for all \(Z \subseteq \mathcal{P}(D)\),

\[
\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I = \llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^{i}
\]

Proof. By virtue of Cobreros et al. (2012b)'s Lemma 1, we only need to show 1) that if \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I = 1\) and \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^{i} = 1\), and 2) that if \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^{i} = 0\), then \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I = 0\).

1. Suppose \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I = 1\) and \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^{i} = 0\). Since \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I = 1\), there is some \(P_3 \sim_{[\exists P_1^*(P_2^i)]_{X_1}^I} \llbracket P_2^i \rrbracket_{X_2}^I\) : \(P_3 \in \llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I\). So, by Definition 7.3.17, there is some \(a_1 \in P_3 : a_1 \in \llbracket P_1^i \rrbracket_{X_1}^I\). By the definition of * (Definition 7.3.15), all subparts of \(a_1\) are also included in \(\llbracket P_1^i \rrbracket_{X_1}^I\). Since \(P_3 \sim_{[\exists P_1^*(P_2^i)]_{X_1}^I} \llbracket P_2^i \rrbracket_{X_2}^I\), by SP, there are two cases:

Case 1: \(P_3 = \llbracket P_2^i \rrbracket_{X_2}^I\). Then \(a_1 \in \llbracket P_2^i \rrbracket_{X_2}^I\). Since \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I = 1\), there is some \(P_4 \in Z\) such that \(P_4 \sim_{[\exists P_1^*(P_2^i)]_{X_1}^I} \llbracket P_2^i \rrbracket_{X_2}^I\) and \(P_4 \notin \llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I\). So by Definition 7.3.17, there is no \(a_2 \in P_4\) such that \(a_2 \in \llbracket P_1^i \rrbracket_{X_1}^I\). By Shared Parts, there is also no \(a_2 \in \llbracket P_2^i \rrbracket_{X_2}^I\) such that \(a_2 \in \llbracket P_1^i \rrbracket_{X_1}^I\).

Case 2: \(P_3 \neq \llbracket P_2^i \rrbracket_{X_2}^I\). Then, by Shared Parts, there is some subpart \(a_2\) of a singleton witness such that \(a_2 \in P_3\) and \(a_2 \notin \llbracket P_2^i \rrbracket_{X_2}^I\). By Definition 7.3.17, subparts of singleton witnesses are also singleton witnesses, so \(a_2 \in \llbracket P_1^i \rrbracket_{X_1}^I\). Since \(\llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I = 0\), there is some \(P_4 \in Z\) such that \(P_4 \sim_{[\exists P_1^*(P_2^i)]_{X_1}^I} \llbracket P_2^i \rrbracket_{X_2}^I\) and \(P_4 \notin \llbracket \exists P_1^*(P_2^i) \rrbracket_{Z,X_1,X_2}^I\). So by Definition 7.3.17, there is no \(a_2 \in P_4\) such that \(a_2 \in \llbracket P_1^i \rrbracket_{X_1}^I\). By Shared Parts, there is also no \(a_3 \in \llbracket P_2^i \rrbracket_{X_2}^I\) such that \(a_3 \in \llbracket P_1^i \rrbracket_{X_1}^I\).
2. Suppose \( \exists P^*(P_2^*) \subseteq Z, X, X_2 = 0 \) and \( \exists P^*(P_2^*) \subseteq Z, X, X_2 = 1 \). Since
\( \exists P^*(P_2^*) \subseteq Z, X, X_2 = 0 \), there is some \( P_3 \sim \exists P^* \subseteq Z, X, X_i \) such that
\( P_3 \notin \exists P^* \subseteq Z, X \). By Definition 7.3.17, there is no \( a \in P_3 \) such that
\( a \in \exists P^* \subseteq X_i \). Since \( P_3 \sim \exists P^* \subseteq Z, X, X_i \) by Shared Parts, \( \exists P^* \subseteq X_i \) is also empty. Since \( \exists P^*(P_2^*) \subseteq Z, X, X_2 = 1 \), there is some \( P_4 \sim \exists P^* \subseteq Z, X, X_i \) such that \( P_4 \cap \exists P^* \subseteq X_i \) is non-empty. Since \( P_4 \sim \exists P^* \subseteq X_1, \exists P^* \subseteq X_i \), by Shared Parts, \( \exists P^* \subseteq X_i \) is also non-empty. \( \square \)

**Theorem 7.6.2** Asymmetry of \( ^{t} A \). If \( A \) be a Montagovian Individual
and \( ^{t} A \) is restricted to distributive properties, then \( ^{t} A \) is asymmetric.

**Proof.** Let \( P_1, P_2 \) be distributive properties such that \( P_1 >^{t} A P_2 \). Show
\( P_2 \not>^{t} A P_1 \). Suppose for a contradiction that \( P_2 >^{t} A P_1 \). Since \( P_1 >^{t} A P_2 \), there is some CC \( Z \) such that \( P_1 \in \exists A \subseteq Z \) and \( P_2 \notin \exists A \subseteq Z \). Since \( P_1 \in \exists A \subseteq Z \), there is some \( P_3 \sim \exists A \subseteq P_1 \) such that \( P_3 \notin \exists A \subseteq Z \) and therefore that \( a \in P_3 \). Since \( P_1 >^{t} A P_2 \), \( P_3 \neq P_1 \) and therefore, by SP, there is some \( a_1 \leq a : a_1 \in P_3 \) and \( a_1 \in P_1 \), since \( P_2 >^{t} A P_1 \), there is some
\( Z' : P_2 \in \exists A \subseteq Z' \), and \( P_2 \notin \exists A \subseteq Z' \). So there is some \( P_4 \sim \exists A \subseteq P_1 \) such that
\( P_4 \notin \exists A \subseteq Z' \). So \( a \in P_4 \) and (by SP) there is some \( a_2 \leq a : a_2 \in P_4 \) and
\( a_2 \in P_2 \). Since \( P_2, P_1 \subseteq Z \), by Incomparability, \( a_1 \leq a_2 \) or \( a_2 \leq a_1 \).

Case 1: \( a_1 \leq a_2 \). Then, since \( P_3 \sim \exists A \subseteq P_1 \), by MC, \( P_3 \sim \exists A \subseteq P_2 \), and \( P_2 \in \exists A \subseteq Z \). \( \square \)

Case 2: \( a_2 \leq a_1 \). Then, since \( P_4 \sim \exists A \subseteq P_2 \), by MC, \( P_4 \sim \exists A \subseteq P_1 \), so \( P_1 \in \exists A \subseteq Z \). \( \square \)

**Theorem 7.6.3** Transitivity of \( ^{t} A \). If \( A \) be a Montagovian Individual
and \( ^{t} A \) is restricted to distributive properties, then \( ^{t} A \) is transitive.

**Proof.** Suppose \( P_1 >^{t} A P_2 \) and \( P_2 >^{t} A P_3 \). Since \( P_1 >^{t} A P_2 \), there is some
\( Z : P_1 \in \exists A \subseteq Z \) and \( P_2 \notin \exists A \subseteq Z \). Since \( P_1 \in \exists A \subseteq Z \), there is some
\( P_4 \sim \exists A \subseteq P_1 \) such that \( P_4 \notin \exists A \subseteq Z \). So \( a \in P_4 \) and there is some \( a_1 \in P_4, P_1 \)
such that \( a_1 \leq a \). Also, since \( P_3 \notin \exists A \subseteq P_2 \), \( P_2 \sim \exists A \subseteq P_2 \). Now consider the CC
\( Z \cup \{P_3\} \). Since \( P_2 >^{t} A P_3 \), the upper bounds of \( P_1, P_2, P_3 \) with respect to \( a \) are comparable, so \( Z \cup \{P_3\} \) is an acceptable CC. By Granularity, \( P_4 \sim \exists A \subseteq P_1 \).

Case 1: \( P_2 \in \exists A \subseteq Z \cup \{P_3\} \). So, by MC, \( P_4 \sim \exists A \subseteq P_2 \). Since \( P_4 \not>^{t} A P_2 \), by CP, \( P_4 \not>^{t} A P_3 \), and, by MC, there is no \( P_5 \in Z \cup \{P_3\} : a \in P_5 \) and
\( P_5 \sim \exists A \subseteq P_3 \). So \( P_3 \notin \exists A \subseteq Z \cup \{P_3\} \) and \( P_1 >^{t} A P_3 \). \( \sqrt{ } \)
Case 2: $P_2 \notin \llbracket A \rrbracket[Z_\cup\{P_1\}]$. Since $P_2 >_A P_3$, by Theorem 7.4.8, $P_3 \notin \llbracket A \rrbracket[Z_\cup\{P_1\}]$. So $P_1 >_A P_3$. ✓

Theorem 7.6.4 Precise holistic predication. Let $A$ be a Montagovian Individual and let $R$ be a holistic predicate. Then, for all $W \subseteq A$ and $Z \subseteq \mathcal{P}(D)$,

$$\llbracket A(R) \rrbracket_{Z,W} = \llbracket A(R) \rrbracket^l_{Z,W} = \llbracket A(R) \rrbracket^s_{Z,W}.$$

Proof. Suppose $\llbracket A(R) \rrbracket^l_{Z,W} = 0$. Then, by Cobreros et al.’s Lemma 1, $\llbracket A(R) \rrbracket_{Z,W} = \llbracket A(R) \rrbracket^s_{Z,W} = 0$. Likewise, if $\llbracket A(R) \rrbracket^s_{Z,W} = 1$, $\llbracket A(R) \rrbracket_{Z,W} = \llbracket A(R) \rrbracket^l_{Z,W} = 1$ and suppose for a contradiction that $\llbracket A(R) \rrbracket_{Z,W} = 0$. Then there is some $P_1 \in \gamma(Z)$ such that $P_1 \sim^\gamma(Z) R$ and $P_1 \notin \gamma(\llbracket A \rrbracket_Z)$. Since $P_1 \sim^\gamma(Z) [R]_W$, there is some $b \preceq \gamma(a)$ such that $b \in P_1$ and $b \in [R]_W$. By SP, $b = \gamma(a)$, but since $P_1 \notin \gamma(\llbracket A \rrbracket_Z)$, $\gamma(a) \notin P_1$. ⊥ And the proof proceeds in a parallel fashion to show that if $\llbracket A(R) \rrbracket^s_{Z,W} = 0$, $\llbracket A(R) \rrbracket_{Z,W} = \llbracket A(R) \rrbracket^l_{Z,W} = 0$. ✓
Conclusion

In this work, I have presented a new theory of the interaction between context sensitivity, vagueness, and scalarity in the adjectival and determiner phrase (DP) domains. In particular, I have argued that, from an empirical point of view, the three phenomena are intimately linked. I have proposed a new logical framework for capturing these observations (*Delineation Tolerant, Classical, Strict* (DelTCS)). In this system, general cognitive indifference relations create not only Sorites-style paradoxes with absolute adjectives, but also the very orderings upon which their tolerance premises are based. Thus, using the system that I have developed, we can arrive at a better understanding of the cognitive and linguistic underpinnings of the vagueness/context sensitivity/scalarity clustering effect that was exemplified throughout the monograph.

More concretely, I have argued for a number of proposals concerning vagueness, context sensitivity, and the semantics and pragmatics of (non-)scalar adjectives. From an empirical point of view, I proposed that the various subclasses of adjectives that were studied in this work show the following context sensitivity and potential vagueness patterns (Table 8.1).

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Relative</th>
<th>Total</th>
<th>Partial</th>
<th>Non-Scalar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context Sensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal context sensitivity</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Existential context sensitivity</td>
<td>(✓)</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Potential Vagueness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-vague ¬P</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>P-vague P</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

Table 8.2  Scalarity patterns

<table>
<thead>
<tr>
<th>Adjective</th>
<th>&gt;p: non-trivial SWO?</th>
<th>&gt;(p_a): non-trivial SWO?</th>
<th>&gt;(p_t): non-trivial SWO?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Total absolute</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Partial absolute</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>Non-scalar</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

Table 8.3  Absolute/non-scalar scale structure patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Relative</th>
<th>Total</th>
<th>Partial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximal element?</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Minimal element?</td>
<td>×</td>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>

I gave an analysis of these patterns within the DelTCS framework, and then I showed that, from this analysis, we correctly predict the scalarity and scale structure patterns associated with the different classes of adjectives (Tables 8.2 and 8.3).

Finally, I have shown that the puzzles raised by absolute adjectives for a theory of vagueness and comparison can be solved within a Delineation framework, provided that we have an appropriate account of the features of vague language. Furthermore, I have shown that the scale-structure properties that have been the exclusive domain of Degree semantics can arise naturally from certain intuitive statements about how individuals can and cannot be indifferent across comparison classes, and that this way of deriving the scalar/non-scalar, relative/absolute, and total/partial distinctions results in a more restrictive and empirically adequate theory of adjectival typology. I therefore conclude that the success of my proposal provides an argument in favor of viewing context sensitivity and general cognitive indifference relations as driving forces behind scalarity in natural language.

In addition, I applied the DelTCS analysis of adjectival context sensitivity, vagueness, and scale structure to account for similar patterns displayed by determiner phrases within a mereological extension of DelTCS: M-DelTCS. I showed that, using this new framework, we can capture certain cross-categorical patterns in the distribution of vagueness and scale structure properties, while attributing observed differences between adjectival and DP scale structure to differences in the part–structure relations that exist between members of DP.
semantic denotations, which are absent from the denotations of adjectives. I therefore conclude that (M-)DelTCS constitutes a well-defined and general framework for analyzing multiple aspects of the meaning of syntactically simple and complex linguistic expressions in the adjectival domain and beyond.
References


REFERENCES


Križ, M. 2014. Homogeneity, non-maximality and all. [online] Available at <http://jos.oxfordjournals.org/content/early/2015/05/21/jos.ffv006.abstract>.


References


Stanford Encyclopedia of Philosophy, [online] Available at <http://plato.stanford.edu/>


Index

absolute adjective axiom 67, 167
absolute scalar adjectives 38–42, 51–2, 67–72, 89
paradox of 70, 89

Bayesian Pragmatics 132–7
adjective interpretation in 135–6
van Benthem, J. 60–1, 63
Be Precise 78
bivalence 13, 19
Boolean algebra 13
borderline cases 17–9, 48–9

classical semantic theory 10–12
Cobreros, P. 9, 25–7
comparison classes 57
context-sensitivity
existential 40–2, 45–8
universal 38–9, 42, 45–8
consequence relations 12, 25–7
Contradiction 10
with explosion 13, 26
contrast preservation 81, 176
convexity
Tolerant 78
Strict 79
Cresswell, P. 15, 122
deduction theorem 14, 26
definite description test 38–9, 84
Degree modifiers 106–8, 111
Degree Semantics 15, 124–9
Degree Tolerant, Classical, Strict 139–44
Delineation Tolerant, Classical,
Strict 56–82, 122–3
language 56
semantics 57–82
determiner phrases
definite plurals 151–4, 178–94
existential plurals 154, 163–77
dimensions 125
distributivity
gather versus numerous 157, 187–91
versus collectivity 151–2, 168
Downward Difference 61, 166
Dowty, D. 151
Égré, P. 35, 53, 73, 75, 77
entailment see consequence relations
evaluativity 61–2, 93
excluded middle 14
existential DPs 154, 158
Faltz, L. 15–16, 150
Fara, D. 20
first order logic see classical semantic
theory
semantics 12
syntax 11
for–phrases 39, 41, 85
Franke, M. 132, 138
Frege, G. 19
fuzzy boundaries 19–20, 48–9
Generalized Quantifier Theory 162
Goodman, N. 132–8
Granularity 80, 176
hedges 29
Heim, I. 16, 125
homogeneity 153, 193
indefiniteness 59, 164
imprecision 40, 48
incomparability 183
indexical expressions 46–7
indifference relations 22, 51, 75–83
informativity 134, 138
Intensional DPs 150–1
Interpretive Economy 124, 128, 147
intolerance 53

Kamp, H. 18
Keenan, E. 15, 16, 150, 162
Kennedy, C. 40, 47–8, 50–1, 65, 93, 125–9
Kleene, S. 25
Klein, E. 58–60, 63
Kratzer, A. 16, 109
Krifka, M. 109

Landman, F. 188
Landman–Link Proposal 188, 190
Lasersohn, P. 29–33, 46, 131, 151
Lassiter, D. 132–8
Levin, B. 110
Lewis, D. 71, 81, 126, 133
Link, M. 15, 157, 161, 188, 190
Logic of Paradox 25

Mereological Delineation Tolerant, Classical Strict 172–94
Mereological Structure Hypothesis 173
mereology 159
McConnell–Ginet, S. 16
McNally, L. 40–1, 93, 108, 131
Minimal Difference 81, 176
Modus Ponens 14, 21, 26

natural precisifications 51
natural transitions 127, 129, 142
non-maximality 152
non-scalar adjectives 43–4, 53, 86, 95–8
No Reversal 61, 145, 165
numeral phrases 154–5

Peirce, C. S. 18
Pinkal, M. 50–1
plural predicates 157
POS 126, 143
pragmatic halos 29–33
pragmatic slack 29
precision 22, 70
Priest, G. 25
prototypicality 47, 76

question under discussion 133
radical pragmatics 89
Récanati, F. 40, 48, 70–1, 87, 131
reflexivity 75
relative scalar adjectives 38–9, 50, 57–66
resultative constructions 108
Rett, J. 62
Ripley, D. 9, 17
van Rooij, R. 9, 20, 38, 41, 53, 57, 60, 63, 65–7, 70, 81, 114, 143
salience 138
Sapir, E. 37, 71
scalar ”coercion” 37, 43–5, 95–9
scalar implicatures 132–5
scale structure 37, 106–14, 124
semantics–pragmatics boundary 89–90, 123, 131
Shared Parts 177
signaling games 132–5
Sorites paradox 20–2, 26, 28, 48–9
strict weak order 63, 68, 103
strong Kleene logic 25
tolerance 20, 22, 26, 52
Tolerant, Classical, Strict 22–8, 32
transitivity 14, 26, 28
Unger, P. 71, 81
Upward Difference 61, 166

vagueness
  definitions 16–22
  potential 50–4, 91–2
  vague predication 153, 172
Wechsler, S. 109
Williamson, T. 123, 127
Winter, Y. 16, 36, 93, 106, 156, 188–90
Wright, C. 20