

UCLA

UCLA Electronic Theses and Dissertations

Title

The Grammar of Tolerance: On Vagueness, Context-Sensitivity, and the Origin of Scale Structure

Permalink

<https://escholarship.org/uc/item/5g7957z0>

Author

Burnett, Heather Susan

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

**The Grammar of Tolerance
On Vagueness, Context-Sensitivity, and the Origin of Scale
Structure**

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Linguistics

by

Heather Susan Burnett

2012

© Copyright by
Heather Susan Burnett
2012

ABSTRACT OF THE DISSERTATION

The Grammar of Tolerance
On Vagueness, Context-Sensitivity, and the Origin of Scale
Structure

by

Heather Susan Burnett

Doctor of Philosophy in Linguistics

University of California, Los Angeles, 2012

Professor Dominique Sportiche, Co-chair

Professor Edward L. Keenan, Co-chair

This dissertation presents a new theory of the relationship between vagueness, context-sensitivity, and adjectival scale structure. Based on both new and well-known data, I show that the following subclasses of adjectival predicates are empirically distinguished in languages like English and French based on these three phenomena: **Relative Adjectives** (ex. *tall, short, expensive, cheap, intelligent, stupid, narrow, wide...*), **Total Absolute Adjectives** (ex. *bald, empty, full, clean, smooth, dry, straight, flat...*), **Partial Absolute Adjectives** (ex. *dirty, bent, wet, curved, crooked, dangerous, awake...*), and **Non-Scalar Adjectives** (ex. *atomic, geographical, pregnant, illegal, dead, hexagonal...*). The main goal of this work is to develop an formal account of both the semantic and pragmatic similarities and differences between the four subclasses of adjectives. I propose that the patterns concerning the behaviour of relative adjectives like *tall* vs absolute adjectives like *straight* and *bent* are reflexes of a single underlying difference in the semantics of these lexical items involving (a certain kind of) context-sensitivity. Moreover, I show 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 rela-

tions and the Sorites paradox from the work of Cobreros et al. (2011a) (among others), I propose a new logical framework that captures the intimate and complex relationship between these three aspects of adjectival meaning.

The dissertation of Heather Susan Burnett is approved.

Paul Égré

Jessica Rett

Edward P. Stabler

Edward L. Keenan, Committee Co-chair

Dominique Sportiche, Committee Co-chair

University of California, Los Angeles

2012

To my parents

TABLE OF CONTENTS

1	Introduction	1
1.1	Vagueness, Context-Sensitivity, and Scalarity	1
1.2	Organization of the Dissertation	5
2	Context-Sensitivity in the Adjectival Domain	9
2.1	Introduction	9
2.2	Adjectival Context-Sensitivity Patterns	10
2.2.1	The Context-Independence Absolute Adjectives	12
2.2.2	The Context-Sensitivity of Absolute Adjectives	14
2.2.3	The Context-Independence of Non-Scalar Adjectives	18
2.2.4	Summary of Context-Sensitivity Data	21
2.3	Type 1 vs Type 2 Context-Sensitivity	21
2.3.1	An Alternative Style of Analysis	24
2.4	Analysis of Type 1 Context-Sensitivity	27
2.4.1	Relative Adjectives	28
2.4.2	Absolute/Non-Scalar Adjectives	45
2.5	Conclusion	48
2.5.1	The Puzzle of Absolute Adjectives	49
3	Vagueness in Logic and Linguistics	54
3.1	Introduction	54
3.2	Our Classical Semantic Theory	56
3.2.1	Classical FOL	56
3.2.2	Extensions in Linguistics	61

3.3	The Phenomenon of Vagueness	62
3.3.1	Borderline Cases	63
3.3.2	Fuzzy Boundaries	66
3.3.3	The Sorites Paradox	67
3.3.4	Summary	68
3.4	<i>Tolerant, Classical, Strict</i>	69
3.4.1	Definition	69
3.4.2	Account of the Puzzling Properties	75
3.5	Lasersohn (1999)'s <i>Pragmatic Halos</i>	77
3.5.1	Definition	77
3.5.2	Comparison with TCS	81
3.6	Conclusion	82
4	Potential Vagueness and Scalar Asymmetries	84
4.1	Introduction	84
4.2	Contextual Variation in Vagueness	86
4.2.1	Vagueness and Relative Adjectives	87
4.2.2	Vagueness and Absolute Adjectives	88
4.3	Contextual Variation with RAs	92
4.3.1	A New Generalization	94
4.3.2	Potential Vagueness	96
4.4	(A)symmetric Vagueness	96
4.5	A Vagueness-Based Characterization of Scale-Structure Distinction	100
4.5.1	Summary	101
4.6	Analysis of Vagueness Patterns	102

4.6.1	Semantics of (Non)Scalar Adjectives	102
4.6.2	Absolute Adjectives	103
4.6.3	Relative Adjectives	129
4.6.4	Non-Scalar Adjectives	139
4.6.5	Coerced Non-Scalar Adjectives	144
4.7	Conclusion	148
4.8	Appendix: Transitivity Countermodel	150
5	Adjectival Scale Structure	154
5.1	Introduction	154
5.2	Scale Structure Patterns	156
5.2.1	Tests for a Maximal Element	156
5.2.2	Tests for a Minimal Element	167
5.2.3	Summary of Scale Structure Patterns	170
5.2.4	Non-Scalar Adjectives	177
5.2.5	Summary of Scale Structure Data	180
5.3	Scale Structure in Delineation Semantics	180
5.3.1	Scale Structure Results	182
5.3.2	Summary	185
5.4	Consequences: Antonymy	186
5.5	Conclusion	194
6	Delineation TCS	195
6.1	Introduction	195
6.2	Vocabulary	195
6.3	Syntax	196

6.4	Semantics	196
6.5	Axiom Schemata	201
6.5.1	Some Definitions	201
6.5.2	Axioms governing $[[\cdot]]$	201
6.5.3	Axioms governing \sim	202
6.5.4	Summary	203
6.6	Empirical Patterns	203
7	Comparison with Other Approaches	205
7.1	Introduction	205
7.2	Review of the Core Data	206
7.3	Basic Features of Degree Semantics	208
7.3.1	Degrees in the Ontology	208
7.3.2	Scalarity as Argument Structure	209
7.3.3	Comparatives as Quantifiers	210
7.3.4	Basic Features of Delineation Semantics	212
7.3.5	Summary	222
7.4	Kennedy (2007)	223
7.4.1	Comparison with Delineation TCS Analysis	225
7.5	Other Analyses of the RA/AA/NS Distinctions	226
7.5.1	Kennedy and McNally (2005)	227
7.5.2	Rotstein and Winter (2004)	229
7.6	Conclusion	230
8	Conclusion	231

LIST OF FIGURES

2.1	Pass me the tall/# empty one	13
2.2	Type 1 context-sensitivity: “Give me the tall/# empty one”	17
2.3	Type 2 context sensitivity: “Give me the empty one”	17
2.4	Two-element (minimal) comparison class X	29
2.5	Application of <i>tall</i> in X	29
2.6	Four-element comparison class X'	30
2.7	Application of <i>tall</i> in X'	30
2.8	Weird application of <i>tall</i> in X'	31
2.9	Uranus (51 118 km diameter) vs. Venus (12 100 km diameter)	39
2.10	Uranus (51 118 km diameter) vs. Neptune (49 500 km diameter)	39
2.11	Application of <i>empty</i> in $\{a, b\}$	46
2.12	Application of <i>empty</i> in $\{a, b, c\}$	46
2.13	Degree scales associated with the semantic denotations of adjectives	49
3.1	Consequence relations in TCS	73
4.1	A picture of Yul Brynner’s head	90
4.2	A heap of penny candy	92
4.3	Tolerant No Skipping	110
4.4	Minimal difference	111
4.5	Contrast Preservation (CP)	112
4.6	Yul Brynner vs Homer Simpson	115
4.7	Yul Brynner vs Homer Simpson vs Marge Simpson	116
5.1	The Degree Scales ($\gg^{t/s}$) of <i>dry</i> and <i>wet</i>	184

7.1	Adjectival Scale Structure Typology	207
7.2	Adjectival Scale Structure Typology	222

LIST OF TABLES

2.1	Context-Sensitivity Patterns	10
2.2	Context-Sensitivity Patterns of Scalar Adjectives	18
2.3	Adjectival Context-Sensitivity Patterns	21
2.4	Two Approaches to the CS of AAs	26
2.5	Axioms governing the semantic denotation of adjectives	49
3.1	Non-Mixed Consequence Relations	74
3.2	Mixed Consequence Relations Validating Tolerance	75
4.1	Potential Vagueness Typology of Adjectives	86
4.2	(Potential) Vagueness Patterns	101
4.3	Potential Vagueness Patterns	101
4.4	Context-Sensitivity Patterns	102
4.5	Axioms governing the semantic denotation of adjectives (repeated)	103
4.6	Pragmatic Axioms for Absolute Adjectives	112
4.7	Context-Sensitivity Patterns	115
4.8	Scalarity Patterns with Absolute Adjectives	123
4.9	Potential Vagueness Patterns (AAs)	126
4.10	Pragmatic Axioms for Scalar Adjectives	131
4.11	Scalarity Patterns with Scalar Adjectives	136
4.12	Pragmatic Axioms for (Non)Scalar Adjectives	140
4.13	Context-Sensitivity Patterns	143
4.14	Potential Vagueness Patterns	143
4.15	Scalarity Patterns	143

4.16	Scalarity Patterns	148
4.17	Pragmatic Axioms for (Non)Scalar Adjectives	149
5.1	Scale Structure Patterns	155
5.2	The fine-grained nature of modification patterns	165
5.3	Tests for a Maximal Endpoint	166
5.4	Scale Structure of Scalar Adjectives	167
5.5	Tests for a Minimal Endpoint	170
5.6	Scale Structure of Scalar Adjectives	170
5.7	Tests for a Maximal Endpoint with <i>open/closed</i>	174
5.8	Tests for a Minimal Endpoint	174
5.9	Tests for a Minimal Endpoint	176
5.10	(Some) Tests for a Maximal Endpoint	177
5.11	Scale Structure of Scalar Adjectives	177
5.12	Scale Structure Patterns	180
5.13	Scalarity Patterns	182
5.14	Absolute Adjective Scale Structure Patterns	184
5.15	Absolute/Non-Scalar Scale Structure Patterns	184
5.16	Absolute/Non-Scalar Scale Structure Patterns	185
6.1	Axioms governing the semantic denotation of adjectives	203
6.2	Axioms governing \sim with (Non)Scalar Adjectives	204
6.3	Scale Structure Distinctions	204
7.1	Degree Semantics vs Delineation Semantics	213
8.1	Context-Sensitivity Patterns	231

8.2	Potential Vagueness Patterns	231
8.3	Scalarity Patterns	232
8.4	Absolute/Non-Scalar Scale Structure Patterns	232

ACKNOWLEDGMENTS

I owe a debt of gratitude to very many people, many more than I will remember to name here. Firstly, I thank my amazing co-chairs, Ed Keenan and Dominique Sportiche. Working with them has enriched me enormously both intellectually and personally, and I would have been completely lost without their academic and professional guidance over the past five years.

Secondly, I thank my other committee members: Jessica Rett, Ed Stabler, and Paul Égré, my external advisor. Every single member of my committee has made contributions to my dissertation project whose scope go well above and beyond those of regular committee members, and the thesis would not have been possible without their detailed comments and help.

Thirdly, I thank the members of the ‘Los Angeles-Paris connection’ (Natasha Abner, Mel Bervoets, Isabelle Charneval, Vincent Homer, Hilda Koopman, Florian Lionnet, and Jérémy Zehr) for their guidance, intellectual support, and friendship across two continents.

Fourthly, I thank the many people with whom I have corresponded about this material. My interactions with these people have greatly improved both the text and the research programme behind it. The long list includes, but is by no means limited to: Alan Bale, Rajesh Bhatt, Denis Bonnay, David-Étienne Bouchard, Lucas Champollion, Francis Corblin, Rose-Marie Déchaine, Michael Friesner, Jenny Doetjes, Brendan Gillon, Thomas Graf, Irene Heim, Chris Kennedy, Dan Lassiter, Eliot Michaelson, Friederike Moltmann, David Nicolas, Walter Pedersen, François Récanati, Uli Sauerland, Yael Sharvit, Stephanie Solt, Benjamin Spector, Anna Szabolsci, Mireille Tremblay, Daniel Valois, and Yoad Winter.

Fifthly, I thank the Social Sciences and Humanities Research Council of Canada (Doctoral fellowship (#752-2007-2382)), the USA/France Partner University Fund grant (*Theoretical/Experimental Linguistic Cognition Advanced Studies* grant (UCLA/ENS, Paris)), and the UCLA Linguistics department for supporting my research.

Finally, and most importantly, I thank my family: my parents, Dawn and Rick, and my sister, Kate, for their love and support over these past 27 years.

VITA

- 2002 Graduated high school at Lisgar Collegiate Institute, Ottawa, Canada.
- 2006 B.A. (Honours)
Special field concentration in Linguistics, French Studies, and Philosophy
Queen's University, Kingston, Canada.
- 2007–present PhD. (Linguistics)
University of California, Los Angeles
Los Angeles, California, USA

PUBLICATIONS

Burnett, Heather. (in press). “Vague Determiner Phrases and Distributive Predication” in Dan Lasnik and Marija Slavkovic (eds). *New Directions in Logic, Language, and Computation*. Springer FoLLI Lecture Notes in Computer Science.

Burnett, Heather. (in press). “Structure événementielle et modification pragmatique: On connaît-tu tout sur /tUt/?” in Davy Bigot, Michael Friesner & Mireille Tremblay (eds). *Les français de l'Amérique du Nord: Théories et idéologies*. Presses universitaires de Laval.

Burnett, Heather. (2012). “The Role of Microvariation in the Study of Semantic Universals: Adverbial Quantifiers in European and Québec French.” *Journal of Semantics*. Vol 28(1): 1-38.

Bouchard, David-Étienne, Heather Burnett & Daniel Valois. (2012). “Degree Fronting in Québec French and the Syntactic Structure of Gradation Quantifier DPs.” in Haïke Jacobs et al. (eds) *Romance Languages and Linguistic Theory 2009*

Burnett, Heather & Mireille Tremblay. (2012). "Change in the Encoding of Direction in the History of French: A Quantitative Approach to Argument Structure Change." in Nynke de Haas & Ans van Kemenade (eds.) *Historical Linguistics 2009*. Amsterdam: John Benjamins. pp. 333-353.

Burnett, Heather. (2010). "Pitch Accent, Focus, and the Interpretation of Non-WH Exclamatives in French." in Sonia Colina et al. (eds) *Romance Linguistics 2009*. Amsterdam: John Benjamins. pp. 369-386.

Bouchard, David-Étienne & Heather Burnett. (2010). "La quantification et l'intensification à distance en français québécois." in Carmen Leblanc et al. (eds). *Vues sur les français d'ici*. Québec: Presses Universitaires de Laval. pp. 31-48.

Burnett, Heather & Mireille Tremblay. (2010). "Directionnalité et aspect en ancien français: l'apport du système prépositionnel." in Mario Barra-Jover et al. (eds). *Approches de la variation linguistique gallo-romane*. Presses de l'Université de Vincennes.

Burnett, Heather & Mireille Tremblay. (2009). "Variable Behaviour Ps and the Location of PATH in Old French." in Enoch Aboh et al. (eds) *Romance Languages and Linguistic Theory 2007*. Amsterdam: John Benjamins. pp. 25-50.

CHAPTER 1

Introduction

1.1 Vagueness, Context-Sensitivity, and Scalarity

This dissertation presents a new theory of the relationship between vagueness, context-sensitivity, and adjectival scale structure. In particular, this work is devoted to the description and analysis of the distribution of these phenomena in 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 chapters 3 and 4; however, we can illustrate some of the puzzles that it raises with the following example: Suppose we take someone who is 1.9 metres tall, and suppose that we agree that, 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 one centimetre each. The 1.9m tall man is at the front of the line, and there is someone who is only 1.5m tall at the end. We can agree that the last person is not tall. Given this setup, there must be some point in this line at which we move from a tall person to his not tall follower, who is one centimetre 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 '± one centimetre' (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 one centimetre, 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 one centimetre, then y is also not tall.

Clearly the fact that both *tall* and *not tall* are tolerant creates a puzzle: why do we not conclude that both the 1.9m man and the 1.5m man are tall and not tall at the same time? Paradoxes of this type are known as *Sorites* paradoxes¹, and they will be discussed in much greater detail throughout the dissertation.

Another adjective that shows a similar pattern is *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 straight stick not straight. Thus, if we were to line up all 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 *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 *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 1mm bend, then y is also not straight.

¹The name of these puzzles comes from a puzzle attributed to Eubelides of Miletus known as 'the Heap' (*soros* being Greek for *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.)

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

The second phenomenon that will be treated in this dissertation is context-sensitivity. To be more specific, I will call a predicate *P* *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 *tall* and *straight* both have this property: someone who can be considered *tall* when we are considering jockeys might not be considered *tall* when we 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 in context 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 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 are straight.

The final phenomenon treated in this dissertation is scalarity. Again, *tall* and *straight* pattern alike on this dimension in that they can both appear in the comparative and many other degree construc-

²Consider, for example, the barrel of a rifle that must be perfectly straight for our shots to be accurate.

tions (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 the dissertation 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:

- (9) **Relative Adjectives (RAs):**
tall, short, expensive, cheap, nice, friendly, intelligent, stupid, narrow, wide...
- (10) **Total Absolute Adjectives (AA^Ts):**
bald, empty, full, clean, smooth, dry, straight, flat...
- (11) **Partial Absolute Adjectives (AA^Ps):**
dirty, bent, wet, curved, crooked, dangerous, awake...
- (12) **Non-Scalar Adjectives (NSs):**
atomic, geographical, polka-dotted, pregnant, illegal, dead, hexagonal...

I propose that the patterns concerning the behaviour 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. (2011a) (among others), I propose a new logical framework that captures the intimate and complex relationship between these three aspects of adjectival meaning.

1.2 Organization of the Dissertation

The dissertation is composed of three parts, each of which is devoted to the analysis of adjectival data associated with the three main topics of this work: context-sensitivity, vagueness, and scale structure.

Part 1: Context-Sensitivity

In chapter 2 (*Context-Sensitivity in the Adjectival Domain*), I present data concerning contextual variation in the meaning of adjectival predicates. Following previous work, I argue that relative adjectives like *tall* and *expensive*, absolute adjectives like *straight* and *empty*, and non-scalar adjectives like *prime* and *hexagonal* all display different context-sensitivity patterns. I give an analysis of these patterns within a *delineation* semantics for scalar terms and discuss the implications that this analysis has for the scales (relations with particular ordering properties) associated with absolute and non-scalar predicates.

Part 2: Vagueness

Chapter 3 (*Vagueness in Logic and Linguistics*) serves as ‘background’ introduction to the empirical phenomenon of vagueness. I outline why it is an important outstanding problem in

both the fields of philosophy and linguistics and then present the logical analysis of the puzzling properties of vague language that I will adopt in the dissertation (Cobrerros et al. (2011a)'s *Tolerant, Classical, Strict* (TCS)). Finally, I compare this approach to another influential system for modelling similar phenomena proposed in linguistics: Lasersohn (1999)'s *Pragmatic Halos* (PH). Readers who are already familiar with the puzzles posed by vague language and their proposed solution within TCS can skip this chapter.

In chapter 4 (*Potential Vagueness and Scalar Asymmetries*), I present new data concerning the distribution of the characterizing properties of vague language presented in chapter 3. I argue that relative, absolute, and non-scalar adjectives display different vagueness-based patterns. I extend the delineation system proposed in chapter 2 with the TCS system described in chapter 3 to give an analysis of these patterns.

Part 3: Scale Structure

In chapter 5 (*Adjectival Scale Structure*), following much previous work, I present data that shows that the adjectival predicates of different classes are associated with scales that have different properties. Furthermore, I show that the association of particular classes of adjectives with their particular kinds of scales is a direct consequence of the analysis developed for accounting for vagueness and context-sensitivity patterns in parts 1 and 2 of the dissertation. In other words, I show that, if we have an appropriate theory of both context-sensitivity and vagueness in the adjectival domain, we get a theory of scale structure patterns 'for free'. This is the main result of this work.

In chapter 6 (*Delineation TCS*), I lay out the proposed logical system (Delineation TCS) in more formal manner. As such, chapter 6 serves as a detailed summary of the proposals made in the dissertation.

Chapter 7 (*Comparison with Other Approaches*) presents a comparison between the account developed in this dissertation within the delineation approach and the currently dominant approach for analyzing the semantics and pragmatics of gradable expressions: degree semantics. I argue that certain empirical facts discussed in the text are unexpected within a degree approach, in which the gradability of a predicate is due its argument structure. I therefore propose that, in addition to it being more elegant, there are empirical reasons for preferring the theory presented here to current alternatives in the literature.

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.

Part 1

Context-Sensitivity

CHAPTER 2

Context-Sensitivity in the Adjectival Domain

2.1 Introduction

This chapter presents data and an analysis of two kinds of context-sensitivity patterns in the adjectival domain. Broadly speaking, we will call a predicate *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 *comparison classes*. Comparison classes are contextually given sets of individuals that influence (in a way to be discussed below) 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 comparison class-based context-sensitivity.

(1) **Relative Adjectives (RAs):**

tall, short, expensive, cheap, nice, friendly, intelligent, stupid, narrow, wide...

(2) **Total Absolute Adjectives (AA^Ts):**

bald, empty, full, clean, smooth, dry, straight, flat...

(3) **Partial Absolute Adjectives (AA^Ps):**

dirty, bent, wet, curved, crooked, dangerous, awake...

(4) **Non-Scalar Adjectives (NSs):**

atomic, geographical, polka-dotted, pregnant, illegal, dead, hexagonal...

Furthermore, following previous work, I argue that to properly understand this variation, it is useful to adopt two patterns of comparison class-based context-sensitivity: (what I will simply call) *type 1* context-sensitivity and *type 2* context-sensitivity. These patterns will be exemplified in great detail below; however, intuitively, predicates that are type 1 context-sensitive will show a greater range of meaning variation than predicates that are type 2 context-sensitive. In this chapter, I argue that the four scale-structure subclasses presented above display the following CS patterns (shown in table 2.1): RAs show both context-sensitivity patterns, both partial and total AAs are not type 1 context-sensitive, but are type 2 context-sensitive, and NSs are neither type 1 nor type 2 context-sensitive.

Pattern	Relative	Total	Partial	Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Table 2.1: Context-Sensitivity Patterns

The chapter is laid out as follows: in section 2.2, I present the CS patterns exhibited by the classes of adjectives in table 2.1. Then, in section 2.3, I present a discussion of the type 1/type 2 context-sensitivity distinction outside the adjectival domain. In particular, I suggest that this distinction broadly corresponds to the difference between the phenomena known as *indexicality/saturation* and *imprecision/loose talk/modulation* in the literature. Finally, in section 2.4, I give an analysis of the patterns in table 2.1 within *delineation* semantics (à la Klein (1980)), and I discuss the consequences of the observed variation in type 1 context-sensitivity for the scales associated with relative and absolute adjectives within this framework. The chapter ends by highlighting a longstanding puzzle in the semantics and pragmatics of scalar predicates: the paradox of the gradability of absolute adjectives. The solution to this puzzle and the analysis of variation in type 2 context-sensitivity will be given in part 2 of the dissertation.

2.2 Adjectival Context-Sensitivity Patterns

This section presents the data concerning the distribution of type 1 and type 2 context-sensitivity 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*, *bald*, and *empty* (5) on the one hand and non-scalar *atomic*, *pregnant* and *geographical* on the other (6)¹.

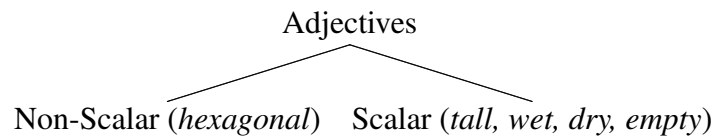
- (5) a. John is taller than Phil.
b. This watch is more expensive than that watch.
c. John is balder than Phil.
d. My cup is emptier than your cup.
- (6) a. ?This algebra is more atomic than that one.
b. ?Mary is more pregnant than Sue.
c. ?This map is more geographical than that one.
d. ?This shape 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*.

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

¹Note that the non-scalars can be very easily coerced into scalar adjectives (i.e. *Mary is more pregnant than Sue: she's farther along*). Coerced NSs will be discussed later in this chapter.

Therefore, we can partition the set of adjectives in the following way based on gradability:



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 like *tall* and *expensive* pattern differently from ones like *bald* and *empty* with respect to a variety of tests associated with how their denotations can vary across contexts². The tests that I present in this section are only a very small subset of the CS-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 Sassoon and Toledo (2011) for more information.

2.2.1 The Context-Independence Absolute Adjectives

The first way in which we can see the difference in context-sensitivity between relative adjectives and both total and partial AAs is through the *definite description* test. As observed by e.g. 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 comparison class when they appear in a definite description. For example, 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 (9), 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*, (10) is infelicitous.

(9) Pass me **the tall one**.

²Note that, by virtue of the fact that *bald* is generally considered to be a vague adjective, Kennedy (2007) classifies it as relative, not absolute. However, (as he notes and as we will see in this section), this adjective passes the context-sensitivity tests for being a total AA.

(10) Pass me **the empty one**.

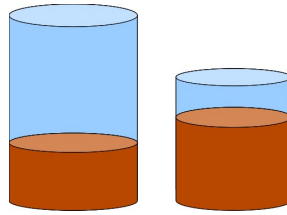


Figure 2.1: Pass me the **tall/# empty one**

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 *bald* are absolute, since (11a) is infelicitous if neither object is (close to) completely full/straight/bald. Likewise, we can make the argument that *dirty*, *wet*, and *bent* are also absolute, since (11b) 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).

(11) **Absolute Adjectives**

- a. Pass me the **full/straight/bald** one.
- b. Pass me the **dirty/wet/bent** one.

Furthermore, we can make the argument that *long*, *expensive*, and *nice* are relative, since the (12) is felicitous when comparing two objects when both or neither are particularly long/expensive/nice.

(12) Pass me the **long/expensive/nice** one.

A correlation of the observation that AAs are not context-sensitive 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 comparison class (Siegel 1979)³.

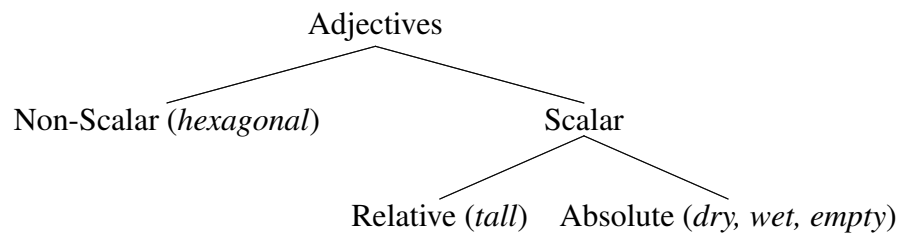
³The precise semantic contribution of *for*-phrases is still somewhat mysterious (see Bylinina (2011) for a recent proposal); however, I assume, as is standard, that they interact with comparison classes in some way and are restricting functions.

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

Absolute adjectives are more resistant to being modified by an expression that makes reference to a comparison class (cf. McNally (2011), Sassoon and Toledo (2011), and Bylinina (2011) for discussion.).

- (14) a. #This towel is **wet/dry** for a used towel.
 b. #This glass is **full/empty** for a plastic glass.
- (15) ?Compared to the glass on the table, this glass is full.
 McNally (2011) (p. 159)

In summary, based on the context-sensitivity tests above, we can make a second distinction:



2.2.2 The Context-Sensitivity of Absolute Adjectives

Of course, saying that absolute adjectives 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 theatre with two spectators in it, the same theatre might be considered *empty* in the context of evaluating attendance at a play (16a); however, it might not be considered so in the context of ensuring that no one is left inside during a fumigation or demolition process (16b).

- (16) a. Only two people came to opening night; the theatre was **empty**.
b. Two people didn't evacuate; the theatre 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. Finally, (to adapt an example from Fara (2000)) a particular man with a small amount of hair may be considered *bald* with respect to men that I would want to date; however, the same man may not be considered *bald* if he is auditioning to play Yul Brynner in a biographical film. 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 the previous section, the context-sensitivity of predicates like *empty* is more restricted than that of predicates like *tall*. Crucially, the examples above all 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, roads with few bends, men with little hair etc.). In other words, all these examples involve what are known in the literature as 'rough' (cf. Austin (1962)), 'loose' (Unger (1975), Sperber and Wilson (1985)), 'modulated' (Récanati (2004), Récanati (2010)), or 'imprecise' (Pinkal (1995), Kennedy and McNally (2005) a.o.) uses. Furthermore, as observed by McNally (2011), Sassoon and Toledo (2011), Bylina (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 (17) 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.

- (17) This restaurant is empty *for a Friday night*.

Data such as (17) and those discussed in the works of McNally and others suggest not only that comparison classes are relevant for some aspects of the meaning of absolute adjectives, 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 (18a) is awkward unless the context makes it very clear that *full* is being used imprecisely, (18b) with a scalar modifier that forces an imprecise use is fine.

- (18) 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 (18b) gives us a first connection between imprecision or ‘loose talk’ associated with AAs and the distribution of scalar modifiers. This connection will be further developed made more explicit throughout the dissertation, particularly in part 3 which presents a context-sensitivity and vagueness-based theory of scale structure.

2.2.2.1 Summary

Based on the empirical observations made in the previous two sections, I argue that we can identify two types of context-sensitivity in the adjectival domain: The first pattern (which I will descriptively call **type 1** context-sensitivity) corresponds to the ability to shift one’s threshold in any non-trivial comparison class. As we saw above, relative adjectives have this property⁴. Thus, *tall* can shift its criteria of application in the CC in figure 2.2, but *empty* cannot.

The second pattern (which I will call **type 2** context-sensitivity) corresponds to the ability to shift one’s threshold in some comparison class (but not necessarily the minimal CCs). Absolute adjectives have have this property, and I suggested that this kind of context-sensitivity is related

⁴Of 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 section 2.4

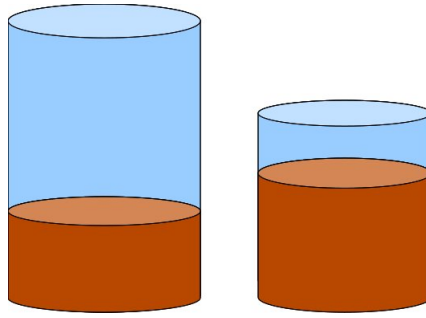


Figure 2.2: Type 1 context-sensitivity: “Give me the tall/# empty one”

to the pragmatic phenomenon of ‘loose talk’ or ‘imprecision’. If we consider again the predicate *empty*, now with reference to the two comparison classes in figure 2.3: while the leftmost container in this figure has a negligible amount of liquid in it and may not be considered empty in the upper comparison class, it may be considered empty in the lower comparison class.

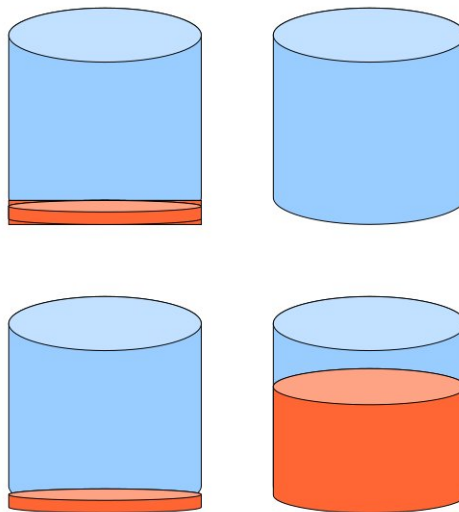


Figure 2.3: Type 2 context sensitivity: “Give me the empty one”

In summary, I have argued (following previous work) that relative and absolute scalar adjectives have different context-sensitivity patterns: RAs are both type 1 and type 2 context-sensitive, while AAs are only type 2 context-sensitive.

In the next section, I examine the context-sensitivity properties of non-scalar adjectives.

Pattern	Relative	Absolute
Type 1 CS	✓	×
Type 2 CS	(✓)	✓

Table 2.2: Context-Sensitivity Patterns of Scalar Adjectives

2.2.3 The Context-Independence of Non-Scalar Adjectives

It is easy to see that, like AAs, NS adjectives are not type 1 context sensitive. Firstly, as shown in (19), they uniformly fail the definite description test: they are only licit in contexts in which exactly one object is atomic/prime/hexagonal.

- (19) a. Pass me the atomic one.
 (But neither/both are atomic!)
 b. Pass me the prime one.
 (But neither/both are prime!)
 c. Pass me the hexagonal one.
 (But neither/both are hexagonal!)

Secondly, as shown in (20), they are much more awkward with *for* phrases than are relative adjectives.

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

What about type 2 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 (21).

- (21) France is hexagonal.

If we're comparing France to shapes in geometry textbooks, it will not be considered hexagonal (it's coastline has very many more 'sides' than six!); however, when we're 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 comparison class. Thus, we can conclude that *hexagonal*, in what Austin calls its 'rough' use (and what we have been calling its 'loose' use), is context-sensitive, and, at first glance, it may look as if we do find type 2 context-sensitive non-scalar adjectives.

But this conclusion would be premature. In fact, 'rough' *hexagonal* is perfectly natural in the comparative construction.

(22) France is more hexagonal than Canada.

So, as soon as we are licenced 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. (22) shows that, while the 'rough' use of *hexagonal* is context-sensitive, it is also scalar. In other words, it has been coerced into an absolute scalar adjective⁵. 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 (23).

(23) a. ?France is hexagonal, for a country.
b. France is **very** hexagonal, for a country.

I therefore propose that non-scalar adjectives are neither type 1 nor type 2 context-sensitive.

The observation that non-context-sensitive non-scalar predicates have properly type 2 CS scalar counterparts is a general one. For example, if we consider a context-sensitive use of *pregnant* as in (24a) (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 (24b).

⁵Note, of course, that even coerced 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.

- (24) 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*.

- (25) 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.
- (26) 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.
- (27) 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 deader than zombie A.

Data associated with the coercion 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:

(28) **Scalarity Generalization:**

An adjective is scalar iff it is type 2 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 dissertation.

2.2.4 Summary of Context-Sensitivity Data

In summary, in the first part of this chapter, I argued that we find the following context-sensitivity patterns in the adjectival domain:

Pattern	Relative	Total	Partial	(Non-Coerced) Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Table 2.3: Adjectival Context-Sensitivity Patterns

Furthermore, I proposed that there exists an important dependency between context-sensitivity and gradability in natural language (cf. (28)).

The final part of the chapter and some of part 2 (on vagueness) of the dissertation will be devoted to the analysis of the data set shown in table 2.3; however, in the next section, I look more closely at the type 1/type 2 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.

2.3 Type 1 vs Type 2 Context-Sensitivity

I have already mentioned, the patterns that distinguish type 1 and type 2 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); Lasnik (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.

(29) I am in Los Angeles.

Furthermore, we can observe that the referent of *I* can change depending on context: if I utter the sentence in (29), then *I* refers to Heather Burnett, and the sentence is true. However, if the speaker in the context of utterance of (29) 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 *8pm PST, April 25th, 2012*. Without using this expression in context, it is possible to assign some referent to it: the point of time that consists precisely of 8pm PST on April 25th. 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 (30a), *8pm PST on April 25th, 2012* might designate only exactly 8pm. However, if someone uses (30b) while recounting their weekend, *8pm PST on April 25th, 2012* will most likely designate a wider temporal interval consisting of multiple minutes around 8pm.

- (30) a. The deadline for the submission of abstracts is 8pm PST on April 25th, 2012.
b. The concert started at 8pm PST on April 25th, 2012.

There is a natural similarity between the extreme context-dependence of indexical pronouns and the extreme context-dependence of relative adjectives, on the one hand, and the more moderate context-dependence of temporal expressions and absolute scalar adjectives on the other. Without knowing what the appropriate contextually given comparison class 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, 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 (30), AAs have a context-independent prototypical core extension that may be broadened if the context requires it; whereas the extension of a relative adjective has no such context-independent meaning⁶.

In fact, (a subset of) the patterns discussed in the previous section 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,

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-H.B.], 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 judgement 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 having 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.

The account of the RA/AA distinction that I will develop in this dissertation 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:

⁶Indeed, a more cognitively oriented way of analyzing the difference between relative and absolute adjectives that reflects their context-sensitivity properties might be to say that AAs have certain distinguished members (i.e. prototypes, cf. Rosch (1973), Lakoff (1987), a.o.) that must always be included in its semantic or pragmatic extension; whereas, relative adjectives have no such members. I will briefly revisit this idea in chapter 4.

(31) **The RA/AA Distinction:**

1. Relative adjectives are indexical expressions; their semantic denotation is assigned relative to a contextually given comparison class (CC). As such, as the value of the CC indexical varies, so too will their semantic denotation.
2. Absolute scalar and non-scalar adjectives are not indexical expressions; their semantic denotation does not depend on a contextually given CC. As such, their semantic denotation does not vary across CCs.

2.3.1 An Alternative Style of Analysis

Although the proposal that RAs have an indexical semantics, while AAs have a context-independent semantics has been proposed before (see the works cited above), it is not the only style of analysis of the context-sensitivity patterns associated with the RA/AA distinction in the literature. In fact, it is not even the most dominant approach in linguistic semantics. Among other reasons, this is undoubtedly because, as will be outlined in section 2.5.1, proposals such as those in (31) give rise to certain puzzles associated with the gradability of AAs that alternative approaches avoid. I will argue in the rest of the dissertation that these puzzles can be given empirically and conceptually satisfying solutions within the framework that I develop in this work; however, for the moment, let us consider a recent influential proposal by Kennedy (2007) (and other accounts in the same vein such as Sassoon and Toledo (2011)) in which both RAs and AAs have a context-sensitive semantics, and the difference between the two classes comes from something else.

Kennedy (2007) proposes that the semantic denotations of AAs are context-sensitive. For Kennedy, who is working within the degree semantics framework (to be discussed in more detail in chapter 7), this boils down to proposing that all scalar predicates have a degree argument that can be bound by a context-sensitive covert operator called *POS*.

- (32) a. $\llbracket \text{tall} \rrbracket = \lambda d \lambda x. x \text{ is tall to degree } d.$
b. $\llbracket \text{POS} [\text{tall}] \rrbracket = \lambda x. x \text{ is tall to degree } d_s \text{ (contextual standard).}$

- (33) a. $\llbracket \text{empty} \rrbracket = \lambda d \lambda x. x \text{ is empty to degree } d.$

- b. $\llbracket \text{POS} [\text{empty}] \rrbracket = \lambda x. x \text{ is empty to degree } d_s \text{ (contextual standard)}$.

This analysis is appealing because it can immediately account for examples that show that AAs are context-sensitive (since POS is context-sensitive), as well as the similarities in the distribution and interpretation of scalar adjectives (i.e. unlike non-scalar adjectives, they are licensed in many of the same syntactic constructions etc.), since RAs and AAs have the same basic kinds of semantic denotations. Of course, now the puzzle for those who propose that AAs also have an indexical semantics is to explain the data that suggests that AAs are not CS in the same way as RAs. In order to solve this puzzle, Kennedy (2007) proposes that what enforces an absolute meaning with most uses of AAs is a meta-grammatical principle called *Interpretative Economy* (34). Furthermore, he proposes that AAs are lexically associated with scales with endpoints (i.e. the endpoint of the *empty* scale is zero objects), thus, (34) is designed to force the interpretation of an AA to be as close to the endpoint of the scale as possible.

(34) **Interpretative Economy:**

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

(Kennedy 2007) (p. 36)

We can therefore break down the analyses of the context-sensitivity patterns of AAs in the literature into two main groups: *semantics-heavy* analyses (i.e. analyses that put the CS of AAs into their semantics, cf. Kennedy and McNally (2005)⁷, Kennedy (2007) (and most work within

⁷I am classifying Kennedy and McNally (2005) as a semantics-heavy approach because they use scale structure and a ‘diachronic’ version of Interpretative Economy to explain the difference between AA and RA CS patterns. They say (p. 360),

The endpoints of the scale provide a fixed value as a potential standard, which in turn makes it possible to assign context-independent truth conditions to the predicate... The alternative-and the only option available to adjectives with open scales- is to compute the standard based on some context-dependent property of degrees... If we assume that interpretations that minimize context-dependence are in general preferred, then closed-scale adjectives should favor an absolute interpretation.

Thus, even though technically they propose that AAs have non-context-sensitive semantic denotations, their analysis is much more in the spirit of the semantics-heavy approach than the pragmatics-heavy approach that will be developed in this work.

degree semantics), Sassoon and Toledo (2011), a.o.) and *pragmatics-heavy* analyses (i.e. analyses that put the CS of AAs into their pragmatics, cf. Sapir (1944), Austin (1962), Unger (1975), Lewis (1979), Rotstein and Winter (2004)⁸, Syrett et al. (2010), Récanati (2010), **this work**, a.o.).

Approach	Semantics Heavy	Pragmatics Heavy
Are AAs indexical?	yes	no
Puzzle	Why are AAs less CS than RAs?	Why are AAs CS?
Solution	Scale structure + Interpretative Economy	Loose Talk

Table 2.4: Two Approaches to the CS of AAs

These two styles of analysis will be compared in great detail in chapter 7, and I will argue that the ‘pragmatics heavy’ approach has certain empirical and conceptual advantages over the ‘semantics heavy’ approach. For the moment, we can simply note that the two styles of analysis have been explored by many authors in the literature on adjectival meaning.

In summary, I proposed to adopt a style of analysis that goes back (at least) to Sapir (1944) in which what separates relative adjectives from absolute adjectives is that only members of the former class have context-sensitive semantic denotations. In the rest of the chapter, I will present a formal analysis of the type 1 context-sensitivity patterns in the adjectival domain (cf. second row of table 2.3) within a particular version of the *delineation* framework for the analysis of the semantics of gradable expressions (to be defined in the next section). It must be noted that there are a number of frameworks for analyzing scalarity currently available in the literature: in addition to delineation semantics, there is the very influential degree semantics approach and Moltmann (2009)’s *tropes* approach, among others. However, I suggest that we have already seen an empirical argument in favour of pursuing a delineation analysis of scale structure distinctions, namely, the descriptive generalization that I proposed in (28) (restated as (35)).

(35) **Scalarity Generalization:**

An adjective is scalar iff it is context-sensitive (either type 1 or type 2).

⁸Rotstein and Winter (2004) make a distinction between the total and partial AAs: they propose that total AAs (like *empty*) have context-independent semantic denotations, while partial AAs (like *wet*) have context-dependent semantic denotations. Therefore, they actually exemplify both styles of analysis.

As we will see, connections between context-sensitivity and scalarity/gradability are built into the architecture of delineation theory, and, thus, I believe that this framework gives us a very natural way to capture the dependency in (35). A more complete comparison between the framework developed in this dissertation (which builds on certain fundamental insights of delineation semantics) and other frameworks will be given in chapter 7.

2.4 Analysis of Type 1 Context-Sensitivity

This section gives a formal analysis of the variation in context-sensitivity (in particular, variation in type 1 context-sensitivity) between relative adjectives on the one hand and absolute and non-scalar adjectives on the other. Before I give the semantic/pragmatic analysis, however, some syntactic remarks are in order. The analyses that I will give in this dissertation will deal with a very small range of syntactic constructions: copular sentences with the positive form of relative adjectives (36a), the positive form of absolute adjectives (36b), comparatives with relative adjectives (36c), comparatives with absolute adjectives (36d), and, finally, positive non-scalar adjectives (36e).

- (36)
- a. John is tall.
 - b. John is bald.
 - c. John is taller than Peter.
 - d. John is balder than Peter.
 - e. John is dead.

Importantly, we will only deal with singular individual denoting subjects like *John* or *Mary*. Extensions of the approach developed in this thesis to plural and quantificational subjects are found in Burnett (2011a) and Burnett (2012a).

Since the data discussed so far deal with the context-sensitivity of the positive forms of the adjectives, in this chapter, I will focus primarily on the copular sentences with relative and absolute adjectives (sentences like (36a) and (36b)). Thus, syntactically, we are interested in three types of lexical items:

1. Singular individual denoting subjects (like *John*), which will generally be notated by lower-case letters of the alphabet ($a, b, c, j, m \dots$).
2. Predicates of the relative adjective class (RA), which will generally be notated by members of the P series: $P, P_1, P_2 \dots$
3. Predicates of the absolute adjectives class (AA), which will generally be notated by members of the Q series: $Q, Q_1, Q_2 \dots$
4. Predicates of the non-scalar class (NS), which will generally be notated by members of the S series: $S, S_1, S_2 \dots$

I will also refer to the entire class of scalar adjectives as SA ($RA \cup AA = SA$). Often, 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 do not believe this will cause confusion.

I first present a delineation analysis of relative adjectives as based on Klein (1980) and van Benthem (1982). I then present a new analysis of absolute and non-scalar predicates within this framework.

2.4.1 Relative Adjectives

In this section, I present the framework that I will adopt for analyzing the semantics of scalar adjectives (delineation semantics), and I apply a (simplified version of) Klein (1980)'s proposal to the analysis of the context-sensitivity and scalarity of relative adjectives. Delineation semantics is a framework for analyzing the semantics of gradable expressions that takes the observation that they are context sensitive to be their key feature. A delineation approach to the semantics of positive and comparative constructions was first proposed by Klein (1980), and has been further developed by van Benthem (1982), Keenan and Faltz (1985), Larson (1988), Klein (1991), van Rooij (2011a), Doetjes (2010), van Rooij (2011b), and Doetjes et al. (2011), 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 thesis. 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 language⁹.

In this framework, scalar adjectives denote sets of individuals and, furthermore, they are evaluated with respect to comparison classes¹⁰, 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 2.4.



Figure 2.4: Two-element (minimal) comparison class X

If we apply the predicate *tall* to the elements in the minimal two-element comparison class $X = \{u, v\}$, 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 2.5).

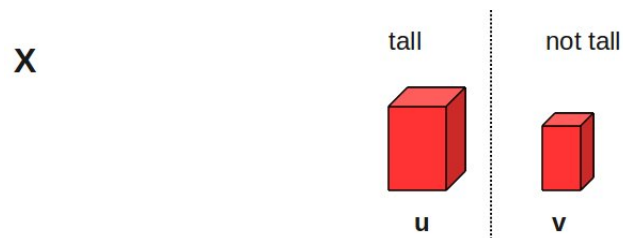


Figure 2.5: Application of *tall* in X

Now consider the larger comparison class X' (also based on Klein (1980), p.18; figure 2.6).

⁹See Kennedy (1997) for challenges raised by certain kinds of empirical phenomena (crosspolar anomaly, incommensurability etc.) for the delineation approach.

¹⁰Comparison classes are important in both delineation semantics and degree semantics. In fact, they are generally viewed unavoidable in any theory of scalar adjectives (Klein 1980). Note however that the comparison classes used in modern degree and delineation semantics are given by context, not entirely contributed by lexical material like the *for* phrase in (i). See Fara (2000) and Kennedy (2007) for discussion.

- (i) John is tall **for a basketball player**.

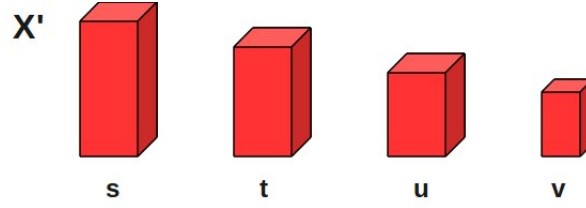


Figure 2.6: Four-element comparison class X'

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 2.7).

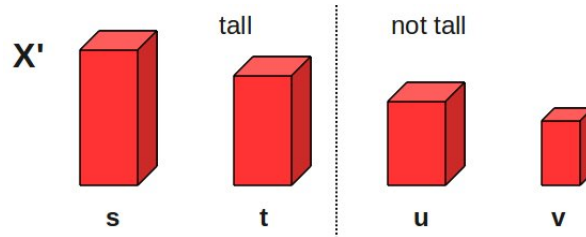


Figure 2.7: Application of *tall* in X'

These examples illustrate how the semantic denotation of a relative adjective can be relativized to and vary depending on comparison classes. More formally, we define our comparison class (CC) models and satisfaction in them as follows:

Definition 2.4.1 Comparison class model. A CC model is a tuple $\langle D, CC, \llbracket \cdot \rrbracket \rangle$ where D is a non-empty domain of individuals and CC is the set of comparison classes such that $CC = \mathcal{P}(D)$. Furthermore, $\llbracket \cdot \rrbracket$ is an interpretation function that assigns semantic values to syntactically atomic constituents.

1. For an individual denoting subject DP a , $\llbracket a \rrbracket \in D$.
2. For $P \in SA$, for every $X \in CC$, $\llbracket P \rrbracket_X \subseteq X$.

Observe that, unlike in first order logic where predicates are assigned any subset of the domain (cf. chapter 3), in the delineation analysis presented here, predicates are assigned different properties in different comparison classes (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 $\llbracket \cdot \rrbracket$ and the comparison classes on which $\llbracket \cdot \rrbracket$ is defined later in this section.

For a given utterance of a sentence containing the positive form of a scalar adjective, the relevant comparison class against which the statement is evaluated is given purely by context. Thus, truth in a CC model is always given with respect to a distinguished comparison class. Note that if the subject of the sentence is not included in the distinguished comparison class, the truth value of the sentence is undefined (as suggested by our judgements concerning sentences like (37)).

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

Definition 2.4.2 *Semantics of the positive form.* For a subject DP a , $P \in SA$, and some $X \in CC$,

$$(38) \quad \llbracket a \text{ is } P \rrbracket_X = \begin{cases} 1 & \text{if } \llbracket a \rrbracket \in \llbracket P \rrbracket_X \\ 0 & \text{if } \llbracket a \rrbracket \in X - \llbracket P \rrbracket_X \\ i & \text{otherwise} \end{cases}$$

This is a very simple analysis of the semantics of relative adjectives. In fact, it's 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 comparison classes, then we will allow some counter-intuitive results. For example, suppose that we apply *tall* in the comparison class X as in figure 2.5. So, in some CC, we have categorized u as *tall* in v as *not tall*. Suppose furthermore that, when we move to the larger comparison class X' , instead of applying *tall* as in figure 2.7, we apply it such that now v is tall and u is not tall (figure 2.8).

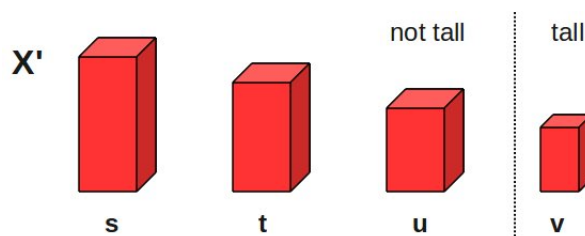


Figure 2.8: Weird application of *tall* in X'

As the moment, nothing in our definitions prohibits applying *tall* as in figure 2.8. 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) that put constraints on relations between degree individuals in the ontology (cf. chapter 7), in the Klein-ian framework, these constraints are put directly on how scalar predicates can be interpreted at different comparison classes. In other words, we can observe that the application of relative scalar predicates like *tall* in natural language is guided by certain coherence principles, and so we build these principles into the interpretation function.

Klein proposes two conditions to ensure that the kind of situation exemplified in figure 2.8 cannot occur. In this work, however, I will adopt another set of constraints on the application of relative adjectives: those presented in van Benthem (1982) and van Benthem (1990). Van Benthem proposes three axioms governing the categorization of individuals across comparison classes. They are the following (presented in my notation):

For $x, y \in D$ and $X \in CC$ such that $x \in \llbracket P \rrbracket_X$ and $y \notin \llbracket P \rrbracket_X$,

Axiom 2.4.1 No Reversal (NR:) *There is no $X' \in CC$ such that $y \in \llbracket P \rrbracket_{X'}$ and $x \notin \llbracket P \rrbracket_{X'}$.*

Axiom 2.4.2 Upward difference (UD): *For all X' , if $X \subseteq X'$, then there is some $z, z' : z \in \llbracket P \rrbracket_{X'}$ and $z' \notin \llbracket P \rrbracket_{X'}$.*

Axiom 2.4.3 Downward difference (DD): *For all X' , if $X' \subseteq X$ and $x, y \in X'$, then there is some $z, z' : z \in \llbracket P \rrbracket_{X'}$ and $z' \notin \llbracket P \rrbracket_{X'}$.*

No Reversal states that if there is some CC in which x is classified as P and y is classified as *not P*, there is no other CC in which they switch; in other words, *No Reversal* rules out the weird application of *tall* in figure 2.8 that we just discussed. **Upward Difference** states that if, in one comparison class, 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 some comparison class, we made a distinction

between some individuals w.r.t. to P , adding extra individuals to CCs cannot erase all distinctions (although they might shift). Finally, **Downward Difference** says that if in some comparison class, there is a $P/not P$ contrast involving x and y , then there remains a contrast in every smaller CC that contains both x and y .

In summary, in the very simple version of the delineation framework that I have presented, the semantic context-sensitivity of relative adjectives is modelled by having sentences involving them be evaluated with respect to a comparison class and allowing the extension of an RA to change depending the comparison classes, subject to only some basic ‘coherence’ constraints like van Benthem’s NR, UD and DD . Thus, the proposed analysis of relative adjectives is given in (39).

(39) **Semantic Analysis of Relative Adjectives:**

If $P \in RA$, then P satisfies NR, UD , and DD , **and nothing else**.

The proposal that the interpretation of all relative adjectives 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 (40) 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.

- (40) 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’ relative adjectives like *brilliant*: as shown in (41), a comparative with *tall* or *short* is not evaluative, but a comparative with *brilliant* is.

- (41) 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 scope of this work, the most promising way to analyze them in a delineation semantic framework would be to propose that *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* are 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 work, I assume that van Benthem's NR, UD, and DD give us a characterization of relative adjectives, at least where their context-sensitivity and scale structure properties are concerned.

The system presented so far allows us to be more explicit about the claim that RAs are type 1 context-sensitive. We can formalize what it means to have the type 1 context-sensitivity property: there are models in which a single individual can be classified as P in one comparison class and *not* P in another.

Definition 2.4.3 Type 1 Context-Sensitivity. *A predicate P is type 1 context-sensitive iff there is some CC model M in which, for some $x \in D$, there is some $X \in CC$ such that $x \in \llbracket P \rrbracket_X$ and there is some distinct $X' \in CC$ such that $x \notin \llbracket P \rrbracket_{X'}$.*

We can now prove that relative adjectives are type 1 context-sensitive.

Theorem 2.4.1 Type 1 Context-Sensitivity of RAs *If $P \in RA$, then P is type 1 context-sensitive.*

Proof Let P be a relative adjective (i.e. let every interpretation of P be constrained by NR, UD, and DD). Now consider the CC model $M = \langle \{a, b, c\}, \mathcal{P}(\{a, b, c\}), \llbracket \cdot \rrbracket \rangle$, where $\llbracket \cdot \rrbracket$ is defined (restricted to P) as follows:

1. $\llbracket P \rrbracket_{\{\}} = \{\}$.

2. $\llbracket P \rrbracket_{\{a\}} = \{a\}$; $\llbracket P \rrbracket_{\{b\}} = \{b\}$; $\llbracket P \rrbracket_{\{c\}} = \{c\}$.
3. $\llbracket P \rrbracket_{\{a,b\}} = \{a\}$.
4. $\llbracket P \rrbracket_{\{b,c\}} = \{b\}$.
5. $\llbracket P \rrbracket_{\{a,b,c\}} = \{a\}$.

Clearly, $\llbracket \cdot \rrbracket$ satisfies NR, UD, and DD, so it is a possible model, and there is some individual, namely b , such that $b \in \llbracket P \rrbracket_{\{b,c\}}$ and $b \notin \llbracket P \rrbracket_{\{a,b,c\}}$. So P is type 1 context-sensitive. \square

2.4.1.1 From Context-Sensitivity to Scalarity

Although the comparison-class-based variation restricted by van Benthem's axioms gives us a nice analysis of the extreme context-sensitivity of relative adjectives¹¹, 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. It is in this sense that, I argued, a simple delineation system already gives us a handle on dependencies like (35). The scales associated with particular adjectival predicates (as well as the comparative relation) are defined as follows:

Definition 2.4.4 *Semantics for the comparative.* For two DPs a, b and $P \in SA$, $\llbracket a \text{ is } P\text{-er than } b \rrbracket = 1$ iff $a >_P b$, where $>_P$ is defined as:

$$(42) \quad x >_P y \text{ iff there is some } X \in CC \text{ such that } x \in \llbracket P \rrbracket_X \text{ and } y \in X - \llbracket P \rrbracket_X.$$

Informally, in this framework, *John is taller than Mary* is true just in case there is some comparison class with respect to which John counts as tall and Mary counts as not tall. Thus, in the example in the figures above (figures 2.4 -2.7), we can establish the ordering $u >_{tall} v$ since u is tall in X and v is not tall in X (cf. figure 2.5). Furthermore, we can establish the orderings $t >_{tall} u$, $s >_{tall} u$, $t >_{tall} v$, and $s >_{tall} v$ from the CC X' (figure 2.7). More generally, van Benthem shows

¹¹Note that Klein and van Benthem do not make a distinction between relative and absolute adjectives, so presumably their analysis might be supposed to apply to all kinds of scalar adjectives. It has been observed, however, (cf. van Rooij (2011c), McNally (2011), Burnett (2012b), and section 2.5.1) that the delineation analysis in its form stated in Klein (1980) and here does not accurately model AAs. Therefore, I only assume that van Benthem's axiom set applies to relative adjectives.

that these axioms give rise to *strict weak orders*: irreflexive, transitive and almost connected relations.

Definition 2.4.5 *Strict weak order.* A relation $>$ is a strict weak order just in case $>$ is **irreflexive**, **transitive**, and **almost connected**.

The definitions of *irreflexivity*, *transitivity* and *almost connectedness* are given below.

Definition 2.4.6 *Irreflexivity.* A relation $>$ is irreflexive iff there is no $x \in D$ such that $x > x$.

Definition 2.4.7 *Transitivity.* A relation $>$ is transitive iff for all $x, y, z \in D$, if $x > y$ and $y > z$, then $x > z$.

Definition 2.4.8 *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 $z > y$.

As discussed in Klein (1980), van Benthem (1990) and van Rooij (2011a), strict weak orders (also known as *ordinal scales* in measurement theory) intuitively correspond to the types of relations expressed by many kinds of comparative constructions¹². For example, one cannot be taller than oneself; therefore $>_{tall}$ 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 $>_{tall}$ 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, $>_{tall}$ should be almost connected. Thus, the theorem in 2.4.2 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 2.4.2 *Strict Weak Order.* For all $P \in RA$, $>_P$ is a strict weak order.

¹²Note that we're talking only about explicit comparatives like *John is taller than Mary*. See van Rooij (2011b) for arguments that implicit comparatives, like *John is tall compared to Mary* lexicalize semi-orders. Furthermore, see van Rooij (2011a) and Sassoon (2010) for arguments that more restrictive orders are necessary to account for the interpretation of measure phrases in comparatives.

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 degree semantics, cf. chapter 7), but it is easy to construct a linear order from them in the following way¹³: first we define an equivalence relation \approx based on $>$.

Definition 2.4.9 Equivalent (\approx_P). For $x, y \in D$ and $P \in RA$, $x \approx_P y$ iff $x \not>_P y$ and $y \not>_P x$.

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 taller than John; that is, John and Mary ‘have the same height’¹⁴ just in case there is no comparison class in which the predicate *tall* distinguishes them. We first verify that \approx_P is an equivalence relation.

Theorem 2.4.3 For all predicates P , \approx_P is an equivalence relation.

Proof Reflexivity. Clearly, a single individual cannot be both in $\llbracket P \rrbracket_X$ and not in $\llbracket P \rrbracket_X$, for any $X \in CC$. Therefore, \approx_P is reflexive. ✓ **Transitivity.** Suppose $x \approx_P y$ and $y \approx_P z$. And suppose for a contradiction that there is some $X \in CC$ such that P distinguishes between x and z in X . **Case 1:** Suppose that $x \in \llbracket P \rrbracket_X$ and $z \notin \llbracket P \rrbracket_X$. Now consider $X \cup \{y\}$. If $y \in \llbracket P \rrbracket_X$, then $y \not\approx_P z$. ⊥ If $y \notin \llbracket P \rrbracket_X$, then $x \not\approx_P y$. ⊥ So $x \approx_P z$. ✓ **Case 2:** Suppose that $z \in \llbracket P \rrbracket_X$ and $x \notin \llbracket P \rrbracket_X$. Now consider $X \cup \{y\}$. If $y \in \llbracket P \rrbracket_X$, then $y \not\approx_P x$. ⊥ If $y \notin \llbracket P \rrbracket_X$, then $z \not\approx_P y$. ⊥ So $x \approx_P z$. ✓ **Symmetry.** Suppose $x \approx_P y$ to show $y \approx_P x$. Immediately from definition 2.4.9. ✓ □

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

Definition 2.4.10 Degree ordering (\gg). For all predicates P and individuals a, b :

$$(43) \quad [a]_{\approx_P} \gg_P [b]_{\approx_P} \text{ iff for all } x \in [a]_{\approx_P} \text{ and all } y \in [b]_{\approx_P}, x >_P y.$$

¹³This method of constructing linear orders from weaker relations (within the context of the analysis of scalar adjectives) is also used by Bale (2011)

¹⁴Note that this equivalence relation should not be taken as an analysis of the equative construction (ex. *John is as tall as Mary*), since this construction has its own particularities (cf. Rett (2008)) that go beyond the scope of this work.

Now we can show that this derived ordering is a strict linear order.

Theorem 2.4.4 *If $P \in RA$, then \gg_P is a linear order.*

Proof Irreflexivity. Immediately by the irreflexivity of $>_P$. **Transitivity.** Immediately by the transitivity of $>_P$. **Totality**¹⁵. Let $a, b \in D$ such that $[a]_{\approx_P} \neq [b]_{\approx_P}$. Suppose that $[b]_{\approx_P} \not\gg_P [a]_{\approx_P}$ to show that $[a]_{\approx_P} \gg_P [b]_{\approx_P}$. Suppose, for a contradiction, that there is some $x \in [a]_{\approx_P}$ such that $x \not>_P y$, for some $y \in [b]_{\approx_P}$. Since, by assumption, $[a]_{\approx_P}$ and $[b]_{\approx_P}$ are distinct, $x \not\approx_P y$. So $y >_P x$. Since $[x]_{\approx_P} = [a]_{\approx_P}$ and $[y]_{\approx_P} = [b]_{\approx_P}$, by the fact that \approx_P is an equivalence relation (cf. theorem 2.4.3), $[b]_{\approx_P} \gg_P [a]_{\approx_P}$. \perp So $[a]_{\approx_P} \gg_P [b]_{\approx_P}$. \square

Thus, if necessary, we can collapse the $>_P$ relations into linear orders to get a more ‘degree’-like structures (for example, one to which measure functions can apply etc.) in which each ‘degree’ is an equivalence class of individuals (see Bale (2011) for more discussion).

2.4.1.2 Positive vs Comparative Forms

With the axioms proposed above, we can observe that the comparison relation has the following property: if $x >_P y$, then, if we look at the minimal two-element comparison class (i.e. if we compare x directly with y), x will be P and y will be *not* P . I call this property *two element reducibility*.

Theorem 2.4.5 Two-element reducibility. $x >_P y$ iff $x \in \llbracket P \rrbracket_{\{x,y\}}$ and $y \notin \llbracket P \rrbracket_{\{x,y\}}$.

Proof \Rightarrow Suppose $x >_P y$ to show $x \in \llbracket P \rrbracket_{\{x,y\}}$ and $y \notin \llbracket P \rrbracket_{\{x,y\}}$. Since $x >_P y$, there is some $X \in CC$ such that $x \in \llbracket P \rrbracket_X$ and $y \notin \llbracket P \rrbracket_X$. Clearly $\{x,y\} \subseteq X$. So, by Downward Difference, there is some $z, z' \in \{x,y\}$ such that $z \in \llbracket P \rrbracket_{\{x,y\}}$ and $z' \notin \llbracket P \rrbracket_{\{x,y\}}$. By No Reversal, $x \in \llbracket P \rrbracket_{\{x,y\}}$ and $y \notin \llbracket P \rrbracket_{\{x,y\}}$. \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 (2011b), there are some cases in which we would like to apply the comparative form

15

Definition 2.4.11 Total. A relation $>$ is *total* iff for all $x, y \in D$, either $x > y$ or $y > x$.

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 2.9 (based on Kennedy's figure 1)), it seems appropriate to say both the sentences in (44).

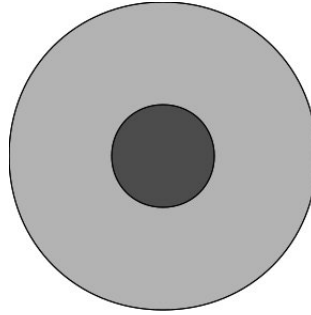


Figure 2.9: Uranus (51 118 km diameter) vs. Venus (12 100 km diameter)

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

This pattern is predicted by the analysis that I gave in the sections above: it predicts that Uranus is bigger than Venus iff, 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 2.10, based on Kennedy's figure 2).

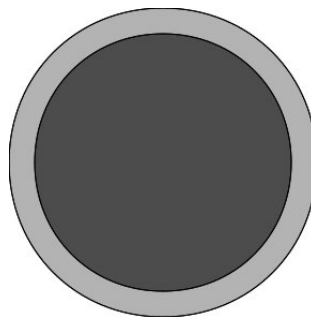


Figure 2.10: Uranus (51 118 km diameter) vs. Neptune (49 500 km diameter)

In this case, although we would assent to (45a), we would generally deny (45b).

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

Similarly, in the situation exemplified in figure 2.10, although (46a) is appropriate, (46b) is inappropriate.

- (46) 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 (2011b) gives an analysis of precisely this contrast within a delineation framework. I will not go through the details of van Rooij's analysis here, since it is somewhat complicated, but the reader is referred to the paper for the technical aspects of the theory. Briefly, he proposes to enrich van Benthem's analysis by adding a series of constraints on what can count as a 'pragmatically appropriate' comparison class. 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 the paper). Then he shows that the comparative relations, defined as in the previous section, gives 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 2.4.5 ceases to hold and we can account for contrasts like those in (46a) and (46b).

I will not incorporate van Rooij (2011b)'s analysis into the current system, particularly, because it implies a slightly different analysis of the properties of vague language than the one that will be adopted in the second part of the dissertation. 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 dissertation.

2.4.1.3 Dimensionality

The analysis of the semantics of relative adjectives and comparatives that I just presented is still very simple, and, indeed, it is still too simple. How the basic delineation system works was illustrated by using the predicate *tall*, and this choice was not arbitrary. The application of this predicate in English is very strongly influenced (if not completely determined) by the physical height of the

individuals that it applies to¹⁶; thus, it is relatively clear that the scale associated with *tall* should be (at least) a strict weak order (i.e. *taller* should be asymmetric etc.).

But what about other adjectives? As observed by Klein (among others), things are less clear with an adjective like *clever*. Suppose that we have two people: John and Mary. Suppose furthermore that Mary, unlike John, is very good with numbers and logical reasoning, while John is much better at dealing with people and making the most of social situations than Mary is. In this situation, it seems that both comparatives in (47) are true.

- (47) a. Mary is more clever than John.
b. John is more clever than Mary.

So we might think that the relations denoted by *more clever*, unlike those denoted by *taller*, are not asymmetric. However, I believe that this conclusion would be premature, particularly because taking a case like the one I just described to be a counterexample to the claim that comparatives denote orders misses an important empirical observation about the sentences in (47): in the situation just described, it also seems like they are both false.

- (48) a. Mary is more clever than John
(No! She's a social moron!)
b. John is more clever than Mary.
(No! He sucks at math!)

In other words, I claim that the sentences in (47) are ambiguous. Although ambiguity is a fundamental notion in semantics, surprisingly, there are not very many well-defined tests for this phenomenon in the literature. The diagnostic for ambiguity that I propose to adopt is Gillon's test (Gillon (1990); Gillon (2004)), stated as follows (2004: 166):

Alternate truth value judgement test. Let α be an expression. Let $\delta()$ be an expression frame such that $\delta()$ is a sentence liable to being judged with respect to a truth

¹⁶Note that this is not even the case in many other languages. For example, French *grand* can mean alternatively something like "tall", "large", or "great" depending on the linguistic and extra-linguistic context.

value. Let s be a state of affairs. If $\delta(\alpha)$ is alternately judged true and judged not true with respect to s , then α is *prima facie* ambiguous.

The pattern targeted by the test is exactly the pattern that we see in (47): we have a state of affairs (one in which Mary is better at math than John, and John is more personable than Mary) and a frame (*Mary is . . . than John*) into which we stick *more clever*. In this case, it is both true that Mary is more clever than John (with respect to book-smarts), and false that Mary is more clever than John (with respect to social smarts). Thus, we can conclude that, at least descriptively, the sentences in (47) are ambiguous. I therefore propose to analyze instances in which comparatives appear to fail to have the basic ordering properties as cases of some kind of lexical ambiguity.

However, this is clearly not the final word on multi-dimensional adjectives. For example, although *more clever* passes Gillon's test for being ambiguous, the relation between the *more clevers* in (47) seems qualitatively different from the classical cases of lexical ambiguity like *bank* (financial institution) and *bank* (riverbed). Furthermore, unlike in the *bank* case, certain linguistic expressions like *in many respects* can even make reference to these dimensions, as shown in (49).

(49) "I'm clever in many respects, but not academically. I wouldn't be doing this job if I could be a vet."

Male model David Gandy. (<http://jezebel.com/5906266/zoey-deschanel-is-the-new-scalp-of-pantene>)

Thus, it is reasonable to think that the relation between a relative adjective and a dimension is more similar to the relation between such an adjective and a contextually given comparison class: both dimensions and comparison classes influence the application of relative predicates. But, in the context of delineation semantics, what are 'dimensions' and how do they affect the interpretation of a relative predicate?

A full theory of multi-dimensionality and how it can be captured within the delineation framework is out of the scope of this work (see however van Rooij (2011a) for some ideas). Nevertheless, I suggest that one way of looking at dimensions are as context-dependent guidelines for the application of a relative predicate across comparison classes, i.e. instructions about which particular

aspects of individuals are relevant to the categorization process. Thus, if we are talking about social aptitude, then *clever* will apply to John in more CCs than it will apply to Mary; however, the opposite will be the case if we are talking about mathematical aptitude.

- (50) a. When book smarts are valued over social smarts,
 $\llbracket \text{clever} \rrbracket_{\{John, Mary\}} = \{Mary\}$; therefore $Mary >_{clever} John$.
- b. When social smarts are valued over book smarts,
 $\llbracket \text{clever} \rrbracket_{\{John, Mary\}} = \{John\}$; therefore $John >_{clever} Mary$.

Thus, the extensions of a relative predicate could be radically different depending on the dimension that is contextually selected; for instance, the extension of *clever (socially)* across CCs might look a lot more similar to the extension of the distinct adjective *personable*, than the extension of *clever (academically)*. Thus, we get the ambiguity effects discussed above. I hypothesize that what makes so-called ‘uni-dimensional’ adjectives like *tall* different from multi-dimensional adjectives like *clever* is that which features are relevant for the application of the former kind of predicate are more-or-less fixed across contexts. Note that even the denotation of an adjective like *tall* can vary depending on which features of individuals are considered more important for the application of the predicate, as shown in the following quotation (from <http://www.csmonitor.com/USA/Latest-News-Wires/2012/0430/World-Trade-Center-back-as-tallest-building-in-New-York-City-video>) describing the prospects of the new World Trade Center being the tallest building in the United States.

- (51) Crowning the world’s tallest buildings is a little like picking the heavyweight champion in boxing. There is often disagreement about who deserves the belt. In this case, the issue involves the 408-foot-tall needle that will sit on the tower’s roof. Count it, and the World Trade Center is back on top. Otherwise, it will have to settle for No. 2, after the Willis Tower in Chicago. “Height is complicated,” said Nathaniel Hollister, a spokesman for The Council on Tall Buildings and Urban Habitats, a Chicago-based organization considered an authority on such records.

In this situation, it seems that the sentences with *tall* in (52) are ambiguous in the way that the

sentences with *clever* in (47) are.

- (52) a. The World Trade Center is taller than the Willis Tower.
b. The Willis Tower is taller than the World Trade Center.

Of course, once it is made clear, for example, that roof needles are relevant for determining the height of a building (i.e we've 'picked a dimension'), (52b) is judged false, and the appearance of symmetry disappears.

In summary, I discussed the problem of multi-dimensional adjectives for the claim that relative comparatives denote (at least) orders that are irreflexive, transitive (so, asymmetric), and almost-connected. I suggested that adjectives like *clever* are not true counter-examples to this claim since the interpretation of relative adjectives is always relativized to a dimension, and, once we take into account the contribution of the dimension, the ordering properties are preserved. Furthermore, I suggested that a promising view to take about 'dimensions' within a delineation approach is that they are instructions to the interpretation function about which features of individuals are relevant for assigning the predicate within comparison classes. However a formalization of this analysis and a full exploration of its consequences is out of the scope of this work.

2.4.1.4 Summary

In summary, the delineation approach to the semantics of relative adjectives makes use of no mechanisms other than individuals and comparison classes (sets of individuals). The semantic denotation of the positive form of the adjective is context-sensitive and varies between CCs. The semantic denotation of a relative comparative is not context sensitive and is defined based on existential quantification over comparison classes. Finally, with van Benthem's axioms constraining the assignment of predicates across comparison classes, we can derive possibly non-trivial strict weak orders that are the scales lexicalized by the comparative.

I argued that Klein and van Benthem's analysis accurately captures the type 1 context-sensitivity of RAs; however, we saw in this chapter that neither AAs nor NSs behave in this way. Therefore, in the next section, I present an analysis of the non-context-sensitivity of the semantic

denotation of these predicates within the delineation system developed here.

2.4.2 Absolute/Non-Scalar Adjectives

I proposed in section 2.3 that both AAs and NSs have semantic denotations that are assigned independently of a contextually given comparison class. That is, in order to know who the bald people are or which rooms are empty, we don't need compare them to a certain group of other individuals; we just need to look at their properties. Similarly, for non-scalar adjectives: 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 comparison classes, what it means to be non-context-sensitive is to have your denotation be invariant across classes. Thus, for an absolute adjective Q and a comparison class 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 classical interpretation of the members of the absolute class that does not apply to the relative class: the *absolute adjective axiom (AAA)*.

Axiom 2.4.4 Absolute Adjective Axiom. *If $Q \in AA \cup NS$, then for all $X \in CC$ and $x \in X$, $x \in \llbracket Q \rrbracket_X$ iff $x \in \llbracket Q \rrbracket_D$.*

In other words, the semantic denotation of an absolute adjective is set with respect to the total domain, and then, by the AAA, the interpretation of Q in D is replicated in each smaller comparison class¹⁷. As an illustration, consider the absolute predicate *empty* and the comparison class $\{a, b\}$. In this example, only container a is truly empty, so when we apply the predicate (as in figure 2.11), only a is in its semantic denotation.

¹⁷At this point, one might wonder why I propose that AAs are evaluated with respect to CCs at all (since they do not affect their semantic denotation). However, recall that *for* phrases are possible with imprecise uses of AAs (i), so we the meaning of AAs must be relativized at some point to a CC. This will become important in the second and third parts of the dissertation.

- (i) This restaurant is very full for a Friday night.

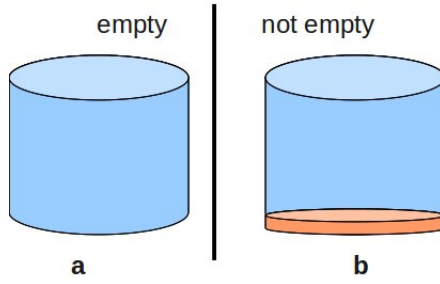


Figure 2.11: Application of *empty* in $\{a, b\}$

If we now move to a larger comparison class like $\{a, b, c\}$, when we reapply the predicate, the AAA tells us that, despite the fact that the CC has changed, a still has to be in the extension of *empty* and b still has to be in its anti-extension (cf. figure 2.12).

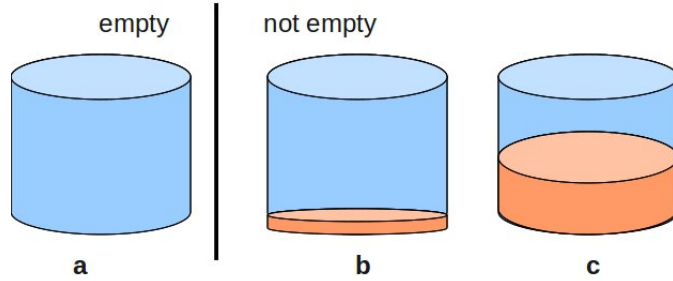


Figure 2.12: Application of *empty* in $\{a, b, c\}$

Note that if *empty* was a relative adjective, a possible interpretation of this predicate in $\{a, b, c\}$ would be $\{a, b\}$. But such an interpretation (in which $b \notin \llbracket \text{empty} \rrbracket_{\{a,b\}}$ and $b \in \llbracket \text{empty} \rrbracket_{\{a,b,c\}}$) is ruled out by the AAA. More generally, if we propose that both AAs and NSs are subject not to van Benthem's axioms, but to the AAA, we can show that the semantic denotation of these predicates is, in a sense, invariant across comparison classes, i.e. these adjectives are not type 1 context-sensitive.

Theorem 2.4.6 *If $Q \in AA \cup NS$, then Q is not type 1 context-sensitive.*

Proof Let $M = \langle D, CC, \llbracket \cdot \rrbracket \rangle$ be a CC model and let $a \in D$ and $Q \in AA \cup NS$. Suppose for a contradiction that there is some $X \subseteq D$ such that $a \in \llbracket Q \rrbracket_X$ and that there is some distinct $X' \subseteq D$ such that $a \notin \llbracket Q \rrbracket_{X'}$. Since $a \in \llbracket Q \rrbracket_X$, by the AAA, $a \in \llbracket Q \rrbracket_D$; however, since $a \notin \llbracket Q \rrbracket_{X'}$, $a \notin \llbracket Q \rrbracket_D$. \perp So Q is not type 1 context-sensitive. \square

Thus, the system that I have proposed derives the appropriate distribution of the type 1 context-sensitivity property that I presented in section 2.2.

Pattern	Relative	Total	Partial	Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Since, in the delineation framework, scalarity is derived from context-sensitivity, the proposal presented in this section already gives us some results 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 2.4.7 *For all $Q \in AA \cup NS$, $>_Q$ is a strict weak order.*

Proof *Irreflexivity.* An individual x cannot be both in $\llbracket Q \rrbracket_X$ and not in $\llbracket Q \rrbracket_X$. *Transitivity.* Trivially. *Almost Connected.* Let $x, y, z \in D$ and suppose $x >_Q y$. Since, by the AAA, all classical denotations are subsets of $\llbracket Q \rrbracket_D$, we have two cases: 1) if $z \in \llbracket Q \rrbracket_Q$, then $z >_Q y$, and 2) if $z \notin \llbracket Q \rrbracket_Q$, then $x >_Q z$.
□

Of course, as shown in the proof of theorem 2.4.7, although they are technically strict weak orders, 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 for the predicate to distinguish three distinct individuals.

Theorem 2.4.8 *If $Q \in AA \cup NS$, then there is no CC model M such that, for distinct $x, y, z \in D$, $x >_Q y >_Q z$.*

Proof Let $Q \in AA \cup NS$ (so it satisfies the AAA). Suppose for a contradiction that there is some CC model $M = \langle D, CC, \llbracket \cdot \rrbracket \rangle$ such that x, y, z are distinct members of D , and $x >_Q y >_Q z$. Then, by definition 2.4.4, there is some $X \in CC$ such that $x \in \llbracket Q \rrbracket_X$ and $y \notin \llbracket Q \rrbracket_X$. Therefore, by the AAA, $y \notin \llbracket Q \rrbracket_D$. Furthermore, since $y >_Q z$, there is some $X' \in CC$ such that $y \in \llbracket Q \rrbracket_{X'}$ and $z \notin \llbracket Q \rrbracket_{X'}$. Since $y \in \llbracket Q \rrbracket_{X'}$, by the AAA, $y \in \llbracket Q \rrbracket_D$. ⊥ So there is no CC model M such that, for distinct $x, y, z \in D$, $x >_Q y >_Q z$.
□

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 s) associated with AAs and NSs have at most two ‘degrees’.

Theorem 2.4.9 *If $Q \in AA \cup NS$, then there is no CC model M such that for $x, y, z \in D$, $[x]_{\approx_Q} \neq [y]_{\approx_Q} \neq [z]_{\approx_Q}$.*

Proof Let $Q \in AA \cup NS$ and suppose for a contradiction that there is some model in which, for $x, y, z \in D$, $[x]_{\approx_Q} \neq [y]_{\approx_Q} \neq [z]_{\approx_Q}$. Since $[x]_{\approx_Q} \neq [y]_{\approx_Q}$, $x \not\approx_Q y$. Without loss of generality, suppose $x >_P y$. So there is some $X \in CC$ such that $x \in \llbracket Q \rrbracket_X$ and $y \notin \llbracket Q \rrbracket_X$. Since $Q \in AA \cup NS$, Q satisfies the AAA. Therefore, $x \in \llbracket Q \rrbracket_D$ and $y \notin \llbracket Q \rrbracket_D$. Since $[y]_{\approx_Q} \neq [z]_{\approx_Q}$, $y \not\approx_Q z$. Suppose without loss of generality that $y >_Q z$. Then, by definition 2.4.4, there is some $X' \in CC$ such that $y \in \llbracket Q \rrbracket_{X'}$ and $z \notin \llbracket Q \rrbracket_{X'}$. So, by the AAA, $y \in \llbracket Q \rrbracket_D$. \perp So there is no CC model M such that for $x, y, z \in D$, $[x]_{\approx_Q} \neq [y]_{\approx_Q} \neq [z]_{\approx_Q}$. \square

In other words, all the elements that are completely *empty* or perfectly *hexagonal* are in one equivalence class (at the same ‘degree’), and all the elements that are not completely *empty* or perfectly *hexagonal* are all treated as equivalent. Note that this is very different from what we see with relative adjectives: these predicates can distinguish between more than two individuals (as an example, consider the model in the proof of theorem 2.4.1) and, correspondingly, the ‘degree-type’ scales that are associated with them (\gg_{PS}) can have more than two ‘degrees’. The difference between the degree scales derived from the context-sensitivity of the semantic denotation of both relative and absolute/non-scalar adjectives is illustrated in figure 2.13.

2.5 Conclusion

In this chapter, I gave a description of the data associated with context-sensitivity in the adjectival domain. I also gave an analysis of the variation in type 1 context-sensitivity between relative scalar adjectives and absolute scalar and non-scalar adjectives within a delineation approach to

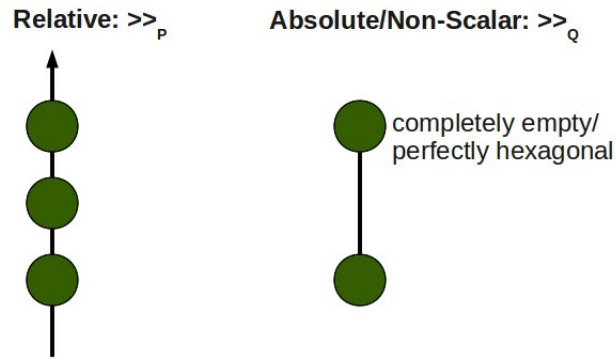


Figure 2.13: Degree scales associated with the semantic denotations of adjectives

the semantics of adjectival predicates. In particular, I proposed that the interpretation of relative adjectives across comparison classes is constrained only by van Benthem’s very weak ‘coherence’ axioms (*No Reversal (NR)*, *Upward Difference (UD)*, *Downward Difference (DD)*); however, the interpretations of absolute adjectives and non-scalar adjectives are constrained by the very strong *absolute adjective axiom (AAA)*. The analysis is schematized in table 2.5¹⁸.

Axiom	RA	AA	NS
No Reversal (NR)	✓	(✓)	(✓)
Upward Difference (UD)	✓	(✓)	(✓)
Downward Difference (DD)	✓	(✓)	(✓)
Absolute Adjective Axiom (AAA)	×	✓	✓

Table 2.5: Axioms governing the semantic denotation of adjectives

This analysis derives the variation in type 1 context-sensitivity between RAs and AAs/NSs; however, we also saw that it makes predictions about the structure of the scales that are associated with these predicates. Specifically, it predicts that although RAs may be associated with non-trivial strict weak orders, both AAs and NSs are only associated with trivial ones (cf. theorem 2.4.8). In the final section of this chapter, I examine the consequences of this prediction.

2.5.1 The Puzzle of Absolute 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 (i.e. non-coerced) non-scalar predicates. For example,

¹⁸Note that the van Benthem’s axioms with AAs/NSs can be proved as theorems, given theorem 2.4.7 and van Benthem (1990)’s theorem 1.5.4, (p.117).

if someone tells me (53), my reaction to this, after I have recovered from the strangeness of the statement, is to say, “Why yes; yes it certainly is.”

(53) 5 is **more prime** than 6.

But, on the other hand, it is clear that trivial scales are inappropriate for absolute adjectives: as discussed in section 2.2, it is perfectly acceptable and very natural to use absolute comparatives like those in (54).

- (54)
- a. Room A is emptier than room B, which is emptier than room C.
 - b. This road is flatter than that road.
 - c. This towel is wetter than that one.
 - d. Ottawa is cleaner than Montréal.

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 (54). To analyze certain differences between RAs and AAs, we proposed that AAs had a semantic denotation that was not context-sensitive. Within the delineation framework, scalarity is derived from context-sensitivity. Thus, as observed by van Rooij (2011c) and McNally (2011), AAs pose the following puzzle for the comparison class-based approach:

- (55) **The Puzzle of Absolute Adjectives:**
- a. If gradability is derived from context-sensitivity, **and**
 - b. AAs are not context-sensitive, **then**
 - c. **How can they be gradable?**

Although I have framed the puzzle of the gradability of absolute adjectives 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 constituents¹⁹. For example, the contradictory nature of AAs is summarized by Récanati (2010) (p. 117), who gives an analysis of the predicate *empty* within the degree semantics framework, in the following way:

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:

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*:

¹⁹Indeed, one way of viewing Kennedy's *Interpretative Economy* proposal is as a solution to the puzzle similar to that of the gradability of absolute adjectives. Although in degree semantics, there is no puzzle about how non-trivial scales get associated with AAs (scale structure is simply stipulated in the lexicon), every existing proposal (that I know of) claims that there is some relation between context-sensitivity and scale structure. Thus, every approach has to deal with questions similar to (55).

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 dissertation, I will give a new solution to the paradox of the gradability of absolute scalar adjectives within an extension of the delineation framework that I proposed in this chapter. Crucially, I will argue that the solution to this longstanding puzzle lies in the appropriate analysis of the type 2 context-sensitivity property exemplified in section 2.2 and its relation to the phenomena of imprecision and 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 the second part of the dissertation, I will make certain proposals about the nature of type 2 context-sensitivity, and then, in the third part of the dissertation, I will show that, with these proposals, 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.

Part 2

Vagueness

CHAPTER 3

Vagueness in Logic and Linguistics

3.1 Introduction

This chapter serves as an introduction to both the main empirical phenomenon to be analyzed in the second part of the dissertation and the formal tools that will be used in the analysis. As such, it has two main parts: in the first part, 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, I present the basic account of the puzzling properties of vague language that I will adopt in this thesis: Cobreros et al. (2011a)'s *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 exist¹. Secondly, and more importantly, many of the debates in the philosophical literature that have given rise to the wide range of theories of

¹See, 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. (2011a), and Cobreros et al. (2011b) for comparisons between TCS and other frameworks.

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)². From a linguist’s point of view, comparisons with FOL are pertinent only inasmuch as, as we will see in section 3.2, 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 the dissertation 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 3.3, 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 chapters 4-6. In section 3.4, I present the *Tolerant, Classical, Strict* account of vague language, and, finally, in section 3.5, I present a similar framework that has been very influential in linguistics: Lasnik (1999)’s *Pragmatic Halos* framework. I give a comparison between the two approaches and argue that, while Cobreros et al. (2011a)’s analysis (as applied to the interpretation of English) is empirically superior, they share many of the same driving intuitions. Thus, one way of looking at TCS is as a more nuanced version of the halos approach.

²**A concrete example:** As we will see in section 3.3, many speakers (the author included) judge sentences of the form *A is both P 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 $p \wedge \neg p$ both being true for many instances of p .

Thus, we can already see that theories of vagueness that, 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 language like English.

3.2 Our Classical Semantic Theory

Although the languages and the models that we will deal with in the rest of the dissertation 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.

3.2.1 Classical FOL

The language of FOL is defined as follows:

Definition 3.2.1 Vocabulary. *The vocabulary of FOL consists of a series of individual constants $a_1, a_2 \dots$, individual variables $x_1, x_2 \dots$, unary predicate symbols $P, Q, R \dots$,³ quantifiers \forall and \exists , and connectives \wedge, \vee, \neg and \rightarrow , plus parentheses.*

Definition 3.2.2 Syntax.

- *Variables and constants (and nothing else) are terms.*
- *If t is a term and P is a predicate symbol, then $P(t)$ is a well-formed formula (wff).*
- *If ϕ and ψ are wffs, then $\neg\phi$, $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \rightarrow \psi$, $\forall x\phi$, and $\exists x\phi$ are wffs.*
- *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 3.2.3 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:*

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

³In this chapter, for simplicity, I will limit the discussion to systems with unary predicates because the n -ary predication case is simply a straightforward generalization of the unary predicate case.

The interpretation of variables is given by *assignments*.

Definition 3.2.4 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 3.2.5 Interpretation. An interpretation \mathcal{I} is a pair $\langle M, g \rangle$, 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 3.2.6 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 3.2.7⁴. In what follows, for an interpretation $\mathcal{I} = \langle M, g \rangle$, a variable x_1 , and a_1 a constant, let $g[a_1/x_1]$ be the assignment in M which maps x_1 to a_1 and agrees with g on all variables that are distinct from x_1 . Also, let $\mathcal{I}[a_1/x_1] = \langle M, g[a_1/x_1] \rangle$.

Definition 3.2.7 Satisfaction (\models). For all interpretations $\mathcal{I} = \langle M, g \rangle$,

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. $\mathcal{I} \models \phi \wedge \psi$ iff $\mathcal{I} \models \phi$ and $\mathcal{I} \models \psi$
4. $\mathcal{I} \models \phi \vee \psi$ iff $\mathcal{I} \models \phi$ or $\mathcal{I} \models \psi$
5. $\mathcal{I} \models \phi \rightarrow \psi$ iff if $\mathcal{I} \models \phi$, then $\mathcal{I} \models \psi$

⁴As 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.

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 3.2.8 Consequence (\models). A set of formulas Ψ is a consequence of a set of formulas Φ (written $\Phi \models \Psi$) iff every interpretation which is a model of Φ is also a model of Ψ .

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

3.2.1.1 Aspects of FOL 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 3.2.7 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 3.2.7 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. In the case of unary predicates, this structure is 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 3.2.7, 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 3.2.7, 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 ,

$$\mathcal{I}(\exists x_1 P(x_1) \wedge \neg P(x_1)) = 0$$

In other words, there are no interpretations of FOL that can satisfy $\exists(x_1 P(x_1) \wedge \neg P(x_1))$. This fact has an important effect on the semantic consequences that we can draw from such sentences. In particular, since no interpretations satisfy $\exists(x_1 P(x_1) \wedge \neg P(x_1))$, by definition 3.2.8, any formula is a consequence of this sentence. In general, (2) follows immediately from the definitions given above.

(2) **Contradiction with Explosion:**

For all formulas ϕ, ψ ,

$$\{\phi, \neg\phi\} \models \psi$$

Secondly, 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 ,

$$\models \forall x_1 (P(x_1) \vee \neg P(x_1))$$

In other words, all interpretations satisfy $P(a_1) \vee \neg P(a_1)$, for all $a_1 \in D$.

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 ϕ, ψ ,

$$\{\phi \rightarrow \psi, \phi\} \models \psi$$

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

(5) **Deduction Theorem:**

For all sets of formulas Γ, Δ ,

$$\Gamma \models \Delta \text{ iff } \models \bigwedge \Gamma \rightarrow \bigvee \Delta$$

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

(6) **Transitivity: :**

For all formulas ϕ, ψ, χ ,

If $\phi \models \psi$ and $\psi \models \chi$, then $\phi \models \chi$

3.2.1.2 Summary

In summary, I have highlighted some basic features of classical first-order logic:

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 3.3, the semantic behaviour of vague predicates will appear to be in conflict with the picture described above.

3.2.2 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 NL 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 described above, 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 (1977), Bierwisch (1989) and very much subsequent work in the field).

(7) **Degree Analysis of *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). 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 (ex. as in Chierchia and McConnell-Ginet (2000), Heim and Kratzer (1998) etc.) or as a more general complement operator (ex. 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^{Eng} *John came to the party and, John came to the party* \models^{Eng} *John was at the party at some time*, then *John came to the party early* \models^{Eng} *John was at the party at some time*.

(9) *John came to the party early* \models^{Eng} *John came to the party*, iff \models^{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 first-order logical systems, but to the vast majority of semantic theories for natural language expressions that are proposed in philosophy and linguistics.

3.3 The Phenomenon of Vagueness

In every day language, the term *vague* has many uses. Not all of these uses refer to the particular linguistic phenomenon that will be studied in this dissertation. 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 dissertation 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.

3.3.1 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 of 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 feet. Similarly, it is clear that anyone under 5ft9” (the average) is not tall. But suppose that we look at John who is somewhere between 5ft9” and 6ft. Which one of the sentences in (12) is true?

(12) a. John is **tall**.

- b. John is **not tall**.

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 favoured by NL speakers. For example, Alxatib and Pelletier (2010) 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 judgements for the predicate *near*.

- (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 5ft11” 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 the previous section, 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 s’valuationist 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)⁵:

- (14) a. If x is less than four feet in height, then ‘ x is schort’ is true.
b. If x is more than six feet in height, then ‘ x is schort’ is false.
(The end)

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

⁵Similar predicates are discussed in Fine (1975), Soames (1999) and Tappenden (1993). Note, in fact, that Fine, contrary to Kamp and Smith a.o., judges underspecified predicates like *schort* to be paradigm cases of vagueness.

3.3.2 Fuzzy Boundaries

A second characterization of vague predicates going back to Frege (1904)'s *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 one cent 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 all the individuals in the domain up 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 (for example, 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 Θ if there is some degree of change in respect of Θ 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 (ex. Eklund (2005), Smith (2008), van Rooij (2010), Cobreros et al. (2011a)). 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.

3.3.3 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 (lit. the paradox of the ‘heaper’) is generally attributed to the Megarian philosopher Eubulides of Miletus⁶, and, informally, it can be laid out as below⁷:

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 be 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’ relation⁸.

⁶The Sorites is often discussed in parallel with another of Eubulides’ seven puzzles: the *Falakros* (‘the Bald Man’). We will discuss the Falakros in more detail in chapter 3.

⁷From the Stanford Encyclopedia of Philosophy.

⁸Note 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.

(17) **The Sorites Paradox**

- a. **Clear Case:** $P(a_1)$
- b. **Clear Non-Case:** $\neg P(a_k)$
- c. **Sorites Series:** $\forall i \in [1, n](a_i \sim_P a_{i+1})$
- d. **Tolerance:** $\forall x \forall y ((P(x) \wedge x \sim_P y) \rightarrow P(y))$
- e. **Conclusion:** $P(a_k) \wedge \neg P(a_k)$

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, though *universal instantiation* and repeated applications of *modus ponens*⁹ 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 6ft to satisfy (17a), and we can find someone who measures 5ft6” to satisfy (17b). In the previous subsection, we concluded that *tall* is tolerant, so it satisfies (17d), and, finally, we can easily construct a Sorites series based on height to fulfil (17c). Therefore, we would expect to be able to conclude that this 5ft6” 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 above, 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.

3.3.4 Summary

Throughout the rest of this dissertation, I will take empirical manifestations of vagueness in natural language to be the cluster of the three properties described in this section:

1. Borderline cases.
2. Fuzzy boundaries/tolerance.
3. Susceptibility to the Sorites paradox.

⁹Note that UI 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 MP that is important for the Sorites.

We can oppose predicates like *tall*, *expensive* and *heap* that display these three properties with “precise” or “non-vague” 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 example of precise predicates and constituents throughout the course of the dissertation, 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.

3.4 *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 dissertation: Cobreros et al. (2011a)’s *Tolerant, Classical, Strict* framework. In what follows, I provide a very succinct definition of the system for expository purposes; however, a more extended version will be given (in the notation adopted in the rest of the thesis) in chapter 4.

3.4.1 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. (2011a) 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 3.4.1 *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 3.4.2 *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:

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

Definition 3.4.3 *T(olerant) 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* (ex. 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 3.4.4 *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 3.4.5 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 3.4.6 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 $\mathcal{I}(a_1) = m(a_1)$.

The classical satisfaction relation in TCS is simply the satisfaction relation of FOL (extended to I_{ps} as shown below). In what follows, for an interpretation $\mathcal{I} = \langle M, g \rangle$, a variable x_1 , and a_1 a constant, let $g[a_1/x_1]$ be the assignment in M which maps x_1 to a_1 and agrees with g on all variables that are distinct from x_1 . Also, let $\mathcal{I}[a_1/x_1] = \langle M, g[a_1/x_1] \rangle$.

Definition 3.4.7 Classical Satisfaction (\models^c). Let M be a t -model such that $M = \langle D, m, \sim \rangle$, and let \mathcal{I} be an interpretation. For all predicates P and terms t_1, t_2 :

1. $\mathcal{I} \models^c P(t_1)$ iff $\mathcal{I}(t_1) \in m(P)$
2. $\mathcal{I} \models^c t_1 I_P t_2$ iff $\mathcal{I}(t_1) \sim_P \mathcal{I}(t_2)$
3. $\mathcal{I} \models^c \neg \phi$ iff $\mathcal{I} \not\models^c \phi$
4. $\mathcal{I} \models^c \phi \wedge \psi$ iff $\mathcal{I} \models^c \phi$ and $\mathcal{I} \models^c \psi$
5. $\mathcal{I} \models^c \phi \vee \psi$ iff $\mathcal{I} \models^c \phi$ or $\mathcal{I} \models^c \psi$
6. $\mathcal{I} \models^c \phi \rightarrow \psi$ iff if $\mathcal{I} \models^c \phi$, then $\mathcal{I} \models^c \psi$
7. $\mathcal{I} \models^c \forall x_1 \phi$ iff for every a_1 in D , $\mathcal{I}[a_1/x_1] \models^c \phi$
8. $\mathcal{I} \models^c \exists x_1 \phi$ iff there is some a_1 in D , $\mathcal{I}[a_1/x_1] \models^c \phi$

Definition 3.4.8 Tolerant/Strict satisfaction ($\models^{t/s}$). Let \mathcal{I} be an interpretation. For all predicates P and terms t_1, t_2 :

1. $\mathcal{I} \models^t P(a_1)$ iff $\exists a_2 \sim_P a_1 : \mathcal{I} \models^c P(a_2)$
2. $\mathcal{I} \models^t t_1 I_P t_2$ iff $\mathcal{I}(t_1) \sim_P \mathcal{I}(t_2)$
3. $\mathcal{I} \models^t \neg\phi$ iff $\mathcal{I} \not\models^s \phi$
4. $\mathcal{I} \models^t \phi \wedge \psi$ iff $\mathcal{I} \models^t \phi$ and $\mathcal{I} \models^t \psi$
5. $\mathcal{I} \models^t \phi \vee \psi$ iff $\mathcal{I} \models^t \phi$ or $\mathcal{I} \models^t \psi$
6. $\mathcal{I} \models^t \phi \rightarrow \psi$ iff if $\mathcal{I} \not\models^s \phi$ or $\mathcal{I} \models^t \psi$
7. $\mathcal{I} \models^t \forall x_1 \phi$ iff for every a_1 in D , $\mathcal{I}[a_1/x_1] \models^t \phi$
8. $\mathcal{I} \models^t \exists x_1 \phi$ iff there is some a_1 in D , $\mathcal{I}[a_1/x_1] \models^t \phi$
9. $\mathcal{I} \models^s P(a_1)$ iff $\forall a_2 \sim_P a_1 : \mathcal{I} \models^c P(a_2)$
10. $\mathcal{I} \models^s t_1 I_P t_2$ iff $\mathcal{I}(t_1) \sim_P \mathcal{I}(t_2)$
11. $\mathcal{I} \models^s \neg\phi$ iff $\mathcal{I} \not\models^t \phi$
12. $\mathcal{I} \models^s \phi \wedge \psi$ iff $\mathcal{I} \models^s \phi$ and $\mathcal{I} \models^s \psi$
13. $\mathcal{I} \models^s \phi \vee \psi$ iff $\mathcal{I} \models^s \phi$ or $\mathcal{I} \models^s \psi$
14. $\mathcal{I} \models^s \phi \rightarrow \psi$ iff if $\mathcal{I} \not\models^t \phi$ or $\mathcal{I} \models^s \psi$
15. $\mathcal{I} \models^s \forall x_1 \phi$ iff for every a_1 in D , $\mathcal{I}[a_1/x_1] \models^s \phi$
16. $\mathcal{I} \models^s \exists x_1 \phi$ iff there is some a_1 in D , $\mathcal{I}[a_1/x_1] \models^s \phi$

Note that the predicates that refer to indifference relations are interpreted ‘crisply’ (in the words of Cobreros et al. (2011a)): their interpretation is the same on all kinds of satisfaction.

3.4.1.1 Consequence Relations

The framework has three notions of satisfaction, and from these notions we can derive 9 consequence relations (defined in a similar manner to the consequence relation of FOL in definition

3.2.8). As discussed in Cobreros et al. (2011a), these relations are in the following lattice order (based on inclusion), where \models^m stands for reasoning from m interpreted premises to n interpreted conclusions. Note (as shown in Cobreros et al. (2011a)) 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 (1979)'s *Logic of Paradox* (LP), and \models^{ss} is equivalent to strong Kleene logic (K3).

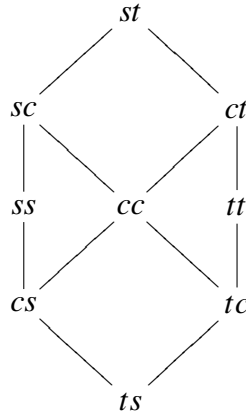


Figure 3.1: Consequence relations in TCS

How appropriate are these systems as basic semantic theories for natural languages? We saw in sections 3.2 and sections 3.3 that, to model a language like English, we want a system validates certain arguments and invalidated 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:**

$$\forall x \forall y (P(x) \wedge x \sim_P y \rightarrow P(y))$$

Secondly, we do not want the Sorites argument (17) to be valid because it is paradoxical. Thirdly, I suggested in section 3.2 that modus ponens (4) and the deduction theorem (5) ought to be valid for English. Furthermore, since, as discussed in section 3.3, 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 3.2, 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:

Argument	\models^{cc} (FOL)	\models^{tt} (LP)	\models^{ss} (K3)
Tolerance	×	✓	×
Sorites	✓	×	×
Modus Ponens	✓	×	×
Deduction Theorem	✓	×	×
Explosion	✓	×	✓
Transitivity	✓	×	×

Table 3.1: Non-Mixed Consequence Relations

From table 3.1, we can observe that the only non-mixed consequence relation that validates tolerance without also validating the Sorites is \models^{tt} (LP). However, systems with \models^{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 \models^{cc} was inadequate, and we can also note that \models^{ss} (K3) does not validate the tolerance principle and validates explosion¹⁰, which is undesirable.

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^{ct} and \models^{st} ¹¹.

¹⁰Note however that excluded middle is invalid in K3.

¹¹Because single premise validity for \models^{mm} is entirely dependent on n .

Argument	\models^{st}	\models^{ct}
Tolerance	✓	✓
Sorites	×	×
Modus Ponens	✓	✓
Deduction Theorem	✓	×
Explosion	×	×
Transitivity	×	×

Table 3.2: Mixed Consequence Relations Validating Tolerance

As shown in table 3.2, \models^{ct} and \models^{st} both validate tolerance but avoid the Sorites paradox (we will return to this point in the next subsection). The only difference between them is that only \models^{st} validates the deduction theorem. I therefore conclude (with Cobreros et al. (2011a)) that \models^{st} 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.

3.4.2 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 3.2). 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.

How this system avoids the Sorites paradox is a bit more complicated. Firstly, following Cobreros et al. (2011a) (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 3.2).

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

(22) **Sorites version 2:**

- a. $P(a_1)$
- b. $\forall i \in [1, n](a_i I P a_{i+1})$
- c. $\forall x \forall y ((P(x) \wedge x I P y) \rightarrow P(y))$

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, (22c) 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 dissertation. In the next section, I will briefly outline a very similar framework that has received a lot of attention in the field of linguistics: Lasersohn (1999)'s *Pragmatic Halos* framework.

3.5 Lasersohn (1999)'s *Pragmatic Halos*

Although it is sometimes presented as such (cf. Sauerland and Stateva (2007)), Lasersohn's *Pragmatic Halos* 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 *pragmatic slack*. Two examples of slack are the sentences in (23) and (24) used when, for example, the theatre has a couple of seats filled or a few irrelevant townspeople are awake.

(23) The theatre is empty.

(24) The townspeople are asleep.

In chapter 3, 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 *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 notation¹²) below.

3.5.1 Definition

In order to make the comparison with TCS optimally perspicuous, I will present 'halo' semantics for the language of first order logic (defined above). Of course, Lasersohn (1999) provides analyses

¹²Lasersohn does not give comprehensive definitions of his system; however, its architecture is easy to reconstruct given his remarks on pages 548-550.

of the semantics and pragmatics of many expressions of English that have no counterparts in FOL (definite plurals, slack regulators/hedges like *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

Definition 3.5.1 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 vocabulary¹³ satisfying:

- For a constant a_1 , $m(a_1) \in D$.
- 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 3.5.2 Halo Model. A halo model is a tuple $M = \langle D, m, h \rangle$, where $\langle D, m \rangle$ is a model (as defined above) 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 , $h(P_1) = \langle \mathcal{X}, \leq_{h(P_1)} \rangle$, where $\mathcal{X} \subseteq \mathcal{P}(D)$ and $\leq_{h(P_1)}$ is a reflexive, transitive and anti-symmetric relation.

¹³Technically 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.

Furthermore (cf. Lasersohn (1999) (p.548)),

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 3.5.3 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.

Definition 3.5.4 Variable Halos. For a variable x_1 , $h(x_1) = \langle X, \leq_{h(x_1)} \rangle$, 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 3.5.5 Interpretation. An interpretation \mathcal{I} is a pair $\langle M, g \rangle$, where M is a halo model and g is an assignment.

Definition 3.5.6 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 $\mathcal{I}(a_1) = m(a_1)$

‘Real’ truth/satisfaction is defined in the standard way¹⁴:

¹⁴Recall that, for an interpretation $\mathcal{I} = \langle M, g \rangle$, 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, $\mathcal{I}[a_1/x_1] = \langle M, g[a_1/x_1] \rangle$.

Definition 3.5.7 Satisfaction (\models^c). Let M be a halo model such that $M = \langle D, m, h \rangle$, and let \mathcal{I} be an interpretation. For all predicates P_1 and terms t_1, t_2 :

1. $\mathcal{I} \models^c P_1(t_1)$ iff $\mathcal{I}(t_1) \in m(P)$
2. $\mathcal{I} \models^c \neg\phi$ iff $\mathcal{I} \not\models^c \phi$
3. $\mathcal{I} \models^c \phi \wedge \psi$ iff $\mathcal{I} \models^c \phi$ and $\mathcal{I} \models^c \psi$
4. $\mathcal{I} \models^c \phi \vee \psi$ iff $\mathcal{I} \models^c \phi$ or $\mathcal{I} \models^c \psi$
5. $\mathcal{I} \models^c \phi \rightarrow \psi$ iff if $\mathcal{I} \models^c \phi$, then $\mathcal{I} \models^c \psi$
6. $\mathcal{I} \models^c \forall x_1 \phi$ iff for every a_1 in D , $\mathcal{I}[a_1/x_1] \models^c \phi$
7. $\mathcal{I} \models^c \exists x_1 \phi$ iff there is some a_1 in D , $\mathcal{I}[a_1/x_1] \models^c \phi$

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

1. $\mathcal{I} \models^e P_1(t_1)$ iff **there is some** $a_1 \in \text{dom}(h(t_1)) : a_1 \in P_2$, for some $P_2 \in \text{dom}(h(P_1))$
2. $\mathcal{I} \models^e \neg\phi$ iff $\mathcal{I} \not\models^e \phi$
3. $\mathcal{I} \models^e \phi \wedge \psi$ iff $\mathcal{I} \models^e \phi$ and $\mathcal{I} \models^e \psi$
4. $\mathcal{I} \models^e \phi \vee \psi$ iff $\mathcal{I} \models^e \phi$ or $\mathcal{I} \models^e \psi$
5. $\mathcal{I} \models^e \phi \rightarrow \psi$ iff if $\mathcal{I} \models^e \phi$, then $\mathcal{I} \models^e \psi$
6. $\mathcal{I} \models^e \forall x_1 \phi$ iff for every a_1 in D , $\mathcal{I}[a_1/x_1] \models^e \phi$
7. $\mathcal{I} \models^e \exists x_1 \phi$ iff there is some a_1 in D , $\mathcal{I}[a_1/x_1] \models^e \phi$

In the next section, I discuss the main ways in which PH and TCS converge and diverge.

3.5.2 Comparison with TCS

Two similarities between TCS (in the version presented in Cobreros et al. (2011a) and above) 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 true¹⁵. 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 denotation to include other objects that are considered to differ in only ‘pragmatically ignorable’ ways.

However, TCS and PH differ in, in my view, two principle ways. The first way involves the interpretation of negation. Lasersohn does not discuss how negation interacts with halos at all, and so we might assume that it is simply interpreted as in FOL (see section 3.2). TCS, on the other hand, interprets negation with reference to another kind of satisfaction: strict satisfaction. As we saw, the interaction between tolerant truth and its dual, strict truth, is what gives the system its account of the puzzling properties of vague language. Lasersohn’s system has no such account. Of

¹⁵This is Cobreros et al. (2011a)’s *Lemma 1*. For Lasersohn:

Theorem 3.5.1 $\mathcal{I} \models^c \phi \Rightarrow \mathcal{I} \models^e \phi$

Proof By induction on \models^c . Non-trivial case: Suppose $\mathcal{I} \models^c P_1(a_1)$. By definition 3.5.2, $m(a_1) \in \text{dom}(h(a_1))$ and $m(P_1) \in \text{dom}(h(P_1))$. So by definition 3.5.8, $\mathcal{I} \models^e P_1(a_1)$. \square

Theorem 3.5.2 *It is not the case that for all \mathcal{I}, ϕ , $\mathcal{I} \models^e \phi \Rightarrow \mathcal{I} \models^c \phi$.*

Proof Let $D = \{a_1, a_2\}$. Let $m(P_1) = \{a_1\}$ and let $P_2 = D$. Let $h(P_1) = \{\{P_1, P_2\}, \{P_1 \leq_{h(P_1)} P_2\}\}$ (+ reflexivity for $\leq_{h(P_1)}$). Suppose $h(a_2) = \{\{a_2\}, \{a_2 \leq_{h(a_2)} a_2\}\}$. Let g be an assignment. So, by the definition of \models^e , $\mathcal{I} \models^e P_1(a_2)$, but $\mathcal{I} \not\models^c P_1(a_2)$, because $a_2 \notin m(P_1)$. \square

course this is not surprising, given that he never addresses the question the proper analysis of vague predicates. However, in the chapters that follow, I will argue that the cases of ‘pragmatic slack’ that Lasersohn discusses are so similar to cases that more traditionally analyzed as vagueness that the phenomenon of “slack” should be subsumed under the phenomenon of vagueness. Thus, in the rest of the dissertation, I will extend the TCS approach rather than the PH approach to develop a unified theory of vague language in the adjectival domain.

The second way in which PH and TCS differ significantly is in the structure of the tolerant denotations/halos. In PH, halos come partially ordered. Both the halo and the ordering is given in the model in PH (so, by context in Lasersohn (1999)). In TCS, the tolerant denotations of constituents are unstructured; they have no ordering on them. While I will argue in chapters 3 and 4 that some kind of ordering of individuals with respect to a vague predicate is important, in Lasersohn’s analysis, it is not clear how the context provides these orders, nor why exactly they must be partial (as opposed to semi-orders, strict weak orders etc). I will return to this point in chapter 5, where we will see how we can *derive* the orders associated with tolerant and strict denotations of predicates within a comparison-class-based extension of TCS.

In summary, I argued that TCS (with its additional notion of *strict truth*) is better equipped to model both classic examples of vague predicates and what Lasersohn calls *pragmatic slack*, which I will argue later ought to also be analyzed as vagueness. However, I also argued that these systems share certain basic core features, and, thus, I view the analyses provided in this dissertation as being very much in line with the project described in Lasersohn (1999).

3.6 Conclusion

In conclusion, I have argued that the challenges that vague predicates raise for modelisation 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 dissertation (Cobreros et al. (2011a)’s *Tolerant, Classical, Strict*) and compared it favourably to another similar

influential framework: Lasersohn (1999)'s *Pragmatic Halos*.

CHAPTER 4

Potential Vagueness and Scalar Asymmetries

4.1 Introduction

This chapter motivates an important empirical connection between vagueness (i.e. the appearance of the properties described in chapter 3) and the ‘scale structure’ classes of adjectives whose context-sensitivity properties were studied in chapter 2. 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 favour of a closer relationship between the phenomena of vagueness and scale structure than is often assumed in the literature.

Concretely, in this chapter, I examine the distribution of the borderline cases, fuzzy boundaries/tolerance, and Sorites susceptibility properties with each of the four classes of adjectives (repeated below).

(1) **Relative Adjectives (RAs):**

tall, short, expensive, cheap, nice, friendly, intelligent, stupid, narrow, wide...

(2) **Total Absolute Adjectives (AA^T s):**

bald, empty, full, clean, smooth, dry, straight, flat

(3) **Partial Absolute Adjectives (AA^P s):**

dirty, bent, wet, curved, crooked, dangerous, awake...

(4) **Non-Scalar Adjectives (NSs):**

atomic, geographical, polka-dotted, pregnant, illegal, dead, hexagonal...

As mentioned in chapter 3, the relative adjectives are uncontroversial examples of vague constituents, and, indeed, that is why the discussion in the previous chapter was limited to them. However, as we will see later in this chapter, 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 absolute adjectives should be analyzed in a parallel manner to their appearance with relative adjectives. 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 *bald* 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 others) who observe that RAs and AAs display vagueness-based patterns.

A main proposal of this chapter is that it is possible to develop a unified analysis of the symptoms of vagueness with both relative and absolute adjectives, even taking into account the differences between the two classes (to be discussed below). Moreover, I argue that there are good empirical arguments to do so. Furthermore, I show that, if we adopt a particular analysis of vagueness (such as Cobreros et al. (2011a)'s *Tolerant, Classical, Strict* (outlined in chapter 3)), not only do we correctly derive the distribution of the *type 2 context-sensitivity* property that was described in chapter 2, but we also arrive at a new solution to the puzzle of the gradability of AAs.

The chapter is laid out as follows: in section 4.2, I present the proposal that the RA/AA distinction is relevant for the phenomenon of vagueness. In particular, I present the evidence discussed in Pinkal (1995) and Kennedy (2007) that absolute adjectives show a different pattern with respect to the context-based availability of the characteristic properties of vagueness than relative adjectives do. However, in section 4.3, I argue that this proposal is empirically incorrect. I provide new data showing that RAs also display contextual variability in the symptoms of vagueness, and, thus, I argue that this property should not be attributed solely to members of the AA class. Building on this empirical observation, I propose that we can arrive at a more accurate description of the phe-

nomenon 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* iff there is some context c such that P has borderline cases, fuzzy boundaries, and gives rise to the Sorites paradox in c .

In section 4.4, I show that relative and absolute adjectives do display variability in (potential) vagueness; however, it is *complement*-based, rather than context-based. In particular, I show that relative adjectives have both potentially vague positive forms (P) and negative forms ($not P$), while, for AAs, only one of the two are potentially vague. I show that whether an AA has a potentially 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 total AAs (ex. *empty*, *bald*, *straight*, *clean* etc.) have potentially vague positive forms and non-potentially vague negative forms; whereas, partial AAs (ex. *wet*, *dirty*, *bent* etc.) have potentially vague negative forms and non-potentially vague positive forms.

Class	P. Vague P	P. Vague $\neg P$
Relative	✓	✓
Total Absolute	✓	×
Partial Absolute	×	✓
Non-Scalar	×	×

Table 4.1: Potential Vagueness Typology of Adjectives

4.2 Contextual Variation in Vagueness

In this section, I present the proposal principally advocated by Pinkal (1995) and Kennedy (2007), and adopted by many authors following him, that the relative/absolute adjectival class distinction that was outlined in the previous section is relevant for the analysis of vague language. In particular, I illustrate the observation made in the literature that absolute adjectives display contextual

variation in the availability of the symptoms of vagueness and present the proposal that such variation is not found with relative adjectives. I then present the account of this proposed variation by Kennedy and others which involves analyzing absolute adjectives as exemplifying a phenomenon that is distinct from vagueness: “imprecision”. In section 4.3, I will go on to argue that relative adjectives also display contextual variation in the presence of the symptoms of vagueness and that this variation is governed by the same principles as with absolute adjectives. However, it is worthwhile first examining why we might think that there are two different phenomena at play in the adjectival domain.

4.2.1 Vagueness and Relative Adjectives

As discussed in chapter 3, the uncontroversial examples of vague predicates are relative adjectives 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. It is generally assumed in the literature that relative adjectives are not only vague in some contexts, but in all contexts. That is, it seems impossible to think of a context in which we can eliminate the borderline cases of *tall* or sharpen its boundaries in a non-artificial way¹. This apparent observation about relative adjectives as the generalization in (6).

(6) **Relative Adjective Generalization:**

If $P \in RA$, then for all contexts/situations c , P has borderline cases, fuzzy boundaries, and gives rise to the Sorites in c .

(6) proposes a link between the membership of an adjectival constituent in a particular lexical class, which we saw was established via vagueness independent syntactic and semantic tests, and the extent to which this constituent will display the symptoms of vagueness. This link is at the heart of the vagueness/imprecision debate. In section 4.3, I will argue that this generalization is false and that RAs display a different pattern, one that they share with AAs. However, we should

¹By ‘non-artificial’, I mean by means other than saying “By *tall*, I mean “has a height greater than exactly 6ft”, and thereby constructing a new precise predicate “taller than exactly 6ft”.

keep (6) in mind when comparing the examples of vagueness with relative predicates that we have discussed so far with similar examples with absolute predicates to be discussed below.

4.2.2 Vagueness and Absolute Adjectives

I now turn to absolute adjectives like *bald*, *empty* and *straight*. 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 seems like the answer to these questions is “yes.” It has been observed since Ancient Greece that adjectives like *bald* and *empty* display certain properties that are eerily similar to the properties displayed by *tall* and *expensive*. For example, if we take a normal case of the use of the word *bald*, talking about men on the street, we can easily identify clear cases of bald men (those with zero hairs on their head) and clear non-cases (those with a full head of hair). However, in this context, *bald* also seems to present the same properties as *tall*. For instance, what about people with a quarter head of hair? Are they bald? Not bald? Both or neither? Thus, in this context, *bald* appears to have borderline cases. Similarly, at what number of hairs does one go from being bald to not bald? The boundaries of *bald* appear fuzzy. Indeed, it seems bizarre to think that there is some point at which adding a single hair to a man’s head could take him from being bald to not bald; therefore, *bald* is tolerant in this context. Thus, we have the ingredients for a Sorites-type argument. In fact, *bald* is associated with one of Eubelides of Miletus’ seven puzzles known as the *Falakros* ‘the bald man’, which, in modern times, has been assimilated to the Sorites. The Falakros can be stated in informal terms as below (from Novak et al. (1999)).

A man having no hair or only one hair is bald. The same holds for a man with two hair etc. Hence, all men are bald.

We can see the same thing for *empty* and *straight*. Consider a context in which we are talking about theatres and whether or not a particular play was well-attended. In this kind of situation, we often apply the predicate *empty* to theatres that are not completely empty (i.e. those with a couple people in them), and, in this context, *empty* has borderline cases, has fuzzy boundaries, and is tolerant: If we are willing to call a theatre with a couple of people in it *empty*, then at what number of spectators does it become *not empty*? Likewise, in most situations, we can refer

to objects with slight bends as *straight*, provided the bends are not large enough to interfere with our purposes. And, in these contexts, *straight* is vague with respect to how big these bends are allowed to be before they make an object become not straight. In summary, we can conclude that, at least in some contexts, absolute adjectives also display the characteristic properties of vague language. It is this observation that has led to the unified treatment of RAs and AAs in many theories of vagueness throughout history. I assume, following the traditional view in the field, that the similarities between the fuzziness of *tall* and the fuzziness of *bald* strongly suggest that we are dealing with a single phenomenon at work in both cases.

On the other hand, it has been observed (by Pinkal (1995), Kennedy (2007) and others) that, in some other contexts, the symptoms of vagueness with AAs disappear. Thus, we might think that the answer to the question of the vagueness of absolute adjectives is “no.” As a first example, we might consider Kennedy (2007)’s 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 (7).

(7) The rod for the antenna needs to be **straight**, but this one has a 1mm 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. Thus, we have a context where *straight* stops being vague. 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 cutoff point between empty theatres and non-empty theatres would be sharply at ‘one or more spectators’. Finally, we can see that even *bald* can stop being vague in some contexts. To adapt an example from Fara (2000): suppose we are trying to cast a movie biography of the actor Yul Brynner. As shown in figure 1, Brynner is completely bald, and, indeed, his appearance is one of the things he is famous

for.



Figure 4.1: A picture of Yul Brynner's head

Thus, it is very important that the person that we pick to play him be completely bald (have zero hairs on their head). In this context, it would be appropriate to say something like (8).

- (8) The lead actor must be **bald**, but this guy has a hair on his head, so unfortunately, he won't work.

In this situation, *bald* has no borderline cases, and adding a single hair moves one sharply from *bald* to *not bald*². In summary, 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 absolute adjectives are vague, I conclude that the appropriate answer to this question is “sometimes (but not always).” This generalization can be stated as the generalization in (9).

(9) **Absolute Adjective Generalization:**

If $Q \in AA$, then there is some context c such that Q has no borderline cases, fuzzy boundaries nor gives rise to the Sorites in c .

As I mentioned above, at first glance, the behaviour of AAs appears to be different from that of

²Another example of a context forcing precise use of *bald* (suggested to me by Ed Stabler) might be one in which a doctor is evaluating a patient's suitability for a hair growth drug or a hair transplant. If we suppose that (for some reason) the procedure will only work if the patient has some of his original hair left, then the doctor might apply *bald* to just those patients that have no hair left.

- (i) This guy is bald, so he is unsuitable for the treatment.

RAs, because RAs seem to be vague in all contexts. We can state this proposed empirical observation that links lexical subclass membership to possible lack of vagueness, which I will henceforth call the *Pinkal/Kennedy Generalization*, as in (10).

(10) **The Pinkal/Kennedy Generalization**

Relative adjectives are vague in all contexts; whereas, there exist contexts in which absolute adjectives are not vague.

In other words, the claim is that absolute adjectives display contextual variation in the presence of vagueness; whereas, relative adjectives do not.

For some scholars, the P/K generalization has important consequences for the possibility of a unified analysis of the properties of vague language with both *tall* and *straight*. In particular, authors such as Kennedy and McNally (2005), Kennedy (2007), Moryzcki (2011), and Husband (2011) assume that the P/K generalization is enough to motivate a departure from the traditional approach to the unified analysis of vague language in the adjectival domain and to treat the appearance of borderline cases etc. with AAs as resulting from a different source. For example, when considering the examples like those discussed above or examples like those in (11), Kennedy (2007) (p. 24) says, “These examples illustrate a phenomenon that is distinct from vagueness, though typically exists alongside it: *imprecision*”.

- (11) a. I'm not awake yet.
b. The theater is empty tonight

Kennedy (2007) (p. 23)

In the next section, however, I will argue that (10) is false: relative adjectives also show contextual variation, and, therefore, I conclude that the contextual variation that we see with AAs does not eliminate the motivation for seeking a unified account of vague effects in the adjectival domain.

4.3 Contextual Variation with RAs

At first glance, the P/K generalization (10) appears correct: indeed, when we consider the classic example of a vague predicate, *tall*, it certainly seems difficult if not impossible to think of situations in which *tall* can be used precisely. However, if we consider a relative adjective like *expensive*, whose scale is built out of discrete units of value (i.e. cents, centimes etc.), we see a different pattern. In fact, it appears that relative adjectives with discrete scales can also be sharpened up. For example, in North America, there is a certain class of candies known as ‘penny’ candies because, historically, they were always sold for one cent in corner stores.



Figure 4.2: A heap of penny candy

When we're discussing the price of one of these candies, since the normal price is one cent, adding a single cent (the smallest amount of change that, for this scale, we could make) to the price of an object will cause it to move from *not expensive* to *expensive*, as shown in the felicitous dialogues in (12) and (13).

- (12) a. Speaker A: How much did you pay for your candy?
b. Speaker B: One cent.
c. Speaker A: Yeah. That's not expensive: that's what they usually cost.
- (13) a. Speaker A: How much did you pay for your candy?
b. Speaker B: 2 cents.
c. Speaker A: That's expensive!

When the appropriate comparison class for *expensive* is the set of penny candies, this predicate has no borderline cases, and its boundaries are sharp: one cent for a candy is not expensive, but

two cents for a candy is. Thus, we have our first counter-example to the generalization that relative adjectives always display the properties of vagueness. Note that it might be tempting to propose, in order to save (10), that *expensive* in (13) has been somehow transformed into a homophonous absolute adjective; however, this analysis would predict that we have two *expensives* in our lexicon: *expensive*₁, which is a gradable adjective, and *expensive*₂, which is a (total) absolute adjective³. However, in this case, we would expect *expensive*₂ to be licensed in the constructions that license AAs, and, as shown in (14), this prediction is not borne out.

(14) ?This watch is almost/completely expensive.

Instead, it seems that all we have done in (13) is provide the appropriate comparison class for the predicate, something that is necessary for all occurrences of relative adjectives.

However, ‘sharpening up’ is not only possible with adjectives that are commonly associated with discrete scales. A second counter-example to (10), which was suggested to me by an anonymous reviewer in another context, involves the relative adjective *long*. Suppose we are in a situation in which we are evaluating the length of trains that are composed of a number of cars, and the cars that are concatenated to form the train are sufficiently long themselves such that adding a single car can make a salient difference to the length of the train. In this situation, it might be appropriate to say something like (15).

(15) Train A, with 3 cars, is not long, but Train B, with 4 cars, is long.

Thus, in this context, *long* has no borderline cases: trains with less than four cars are not long, and

³Another option is that the absolute adjective *expensive*₂ is actually a member of the partial class. This analysis would predict that, like the other partial adjectives, *expensive*₂ would have an existential meaning. Presumably, in the same way that the partial adjectives *wet/dirty/sick* are generally proposed to hold of objects with non-zero degrees of wetness/dirt/sickness (cf. Yoon (1996); Rotstein and Winter (2004), among others), it would be true of an object just in case the object had a non-zero degree of value. However, then we would expect to be able to utter (i) using *expensive*₂ and then to conclude (ii). But this is simply not possible.

(i) This watch is more expensive than that watch.

(ii) This watch is expensive.

trains with more than 3 cars are long, and the boundaries of *long* (for a train) are sharp: between three and four cars. In other words, contra (10), *long* does not display the properties of vague language in this context.

A final counter-example (which was also suggested to me) involves the predicate *hot*. Suppose we want to bake a cake and, according to the recipe, we need to preheat the oven to 350 degrees Fahrenheit. In this situation, it is conceivable that we might consider an oven heated to any degree less than 350 as not hot; however, we move sharply to hot ovens as soon as the temperature hits 350 degrees.

I therefore conclude that (10) is incorrect, and that whether or not an adjective can be used with a precise meaning is not determined by its membership in the relative or absolute adjectival classes.

4.3.1 A New Generalization

What all the examples discussed above seem to have in common is that we have constructed a context in which making the minimal amount of change possible with respect to a dimension makes a big difference to our interests (whether we are being ripped off at the corner store; whether we can bake our cake etc.). In other words, in all these cases, the ‘indistinguishable difference’ relation (\sim_P) with which we construct the Sorites series relates no pairs that occupy different degrees on the scale associated with P ($>_P$). We can state this novel empirical generalization about the availability of the symptoms of vagueness with relative adjectives as in (16).

(16) Contextual Variation in Vagueness with RAs:

For a relative adjective P and its associated scale $>_P$,

- a. If, in some context, there is no $x \neq y$ such that $x \sim_P y$ and $x >_P y$, then P has no borderline cases, has sharp boundaries and does not give rise to the Sorites in that context.

In other words, when moving a single degree on a scale associated with P makes a noticeable difference to our interests, that's when we will not be able to construct a Soritical series along the scale associated with P , and this relative adjective will stop being vague.

Furthermore, we can observe that we have exactly the same pattern with absolute adjectives: in the Yul Brynner example, since adding one hair to someone's head (the smallest amount of change that we can make on the scale associated with *bald*) makes a big difference to who we're going to cast in the movie, \sim_{bald} will not relate people that have no hair with those that have even one hair in this context. Likewise, if there being a single person in a theatre results in the loss of a human life, then \sim_{empty} will not relate theatres with no people in them to theatres with people. I therefore conclude that the generalization in (16) holds not only of relative adjectives, but also of absolute adjectives. Thus, we should reformulate it as a generalization about the entire class of scalar adjectives: (17)⁴.

(17) **Contextual Variation in Vagueness with Scalar Adjectives:**

For a scalar adjective P of any lexical class and its associated scale $>_P$,

- a. If, in some context, there is no $x \neq y$ such that $x \sim_P y$ and $x >_P y$, then P has no borderline cases, has sharp boundaries and does not give rise to the Sorites in that context.

In other words, although they do differ at some level (as proposed in chapter 2), RAs and AAs obey the same basic principles when it comes to contextual variation in the presence of the characterizing properties of vagueness. Thus, I conclude that the Pinkal/Kennedy generalization (10) about the availability of borderline cases etc. across contexts should be discarded in favour of (17). I leave for future research the investigation of what kinds of (extra)linguistic factors create contexts in which objects at very close degrees on a scale are viewed as relevantly different, and whether or not there is something principled to be said about why it appears that examples of such contexts are easier to construct with absolute rather than relative adjectives.

⁴This generalization is, in some sense, the other side of the coin of contextualist proposals like Fara (2000)'s *Similarity Constraint* (see also proposals made by Kamp (1981) and Soames (1999)). These authors propose generalizations concerning when the boundaries of predicates can appear fuzzy. Here I give a generalization concerning when they can appear sharp.

I therefore conclude that we should look for a unified analysis of the puzzling properties of vague language with both *tall* and *bald*?

4.3.2 Potential Vagueness

In the previous sections, we observed that vagueness (even with (at least some) relative adjectives) is context-dependent. In other words, I argued that being *vague* (by which I mean “exhibiting the cluster of properties discussed in section 4.2”) is a stage-level property, 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, “imprecision”, and the absolute/relative distinction, we should employ a more nuanced notion, one that makes the contribution of the context fully explicit. I therefore introduce the term *potentially vague*, defined in (18)⁵.

(18) **Potential Vagueness:**

An adjective *P* is *potentially vague* iff there is some context *c* such that *P* has borderline cases, fuzzy boundaries, and gives rise to a Soritical argument in *c*.

In the next section, I argue that, while possible precision does not distinguish relative from absolute adjectives, something else does, namely whether both a predicate’s positive form and its negation are potentially vague.

4.4 (A)symmetric Vagueness

We saw in the previous sections that both relative and absolute predicates can be used in contexts in which they have a sharp meaning. It is because of this fact that, I argued, the notion of *potential vagueness* is more useful than the more commonly used context-independent term *vagueness*: for

⁵A more formal characterization of the *potentially vague* property is given in section 4.6; however, for the moment, (18) will suffice.

predicates that can display borderline cases, tolerance etc., it seems, in principle, to be possible to find contexts in which they can be used precisely. However, an important piece of the empirical puzzle that has not yet been discussed is that not all of these potentially vague predicates are potentially vague in the same way. In particular, it seems that the relative/absolute distinction **is**, in fact, relevant for vagueness, and we can see this by comparing positive potentially vague predicates with their negations.

Firstly, we can observe that, for relative adjectives, there is no difference in the potential vagueness of their positive form and their negation. We saw in section 4.2 that *tall* was potentially vague, and we can make the same observation about *not tall*: At what point does adding a millimetre to the height of a ‘not tall’ person change them into a tall person? In the contexts in which ‘± one millimetre’ counts as an irrelevant change, then *not tall* will also be tolerant; that is, we will generally assent to both the statements in (19).

(19) **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 a millimetre, then y is tall.
- b. **Not tall:** For all x, y , if x is not tall and x and y ’s heights differ by a millimetre, then y is not tall.

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

Definition 4.4.1 *Symmetric vagueness.* A predicate P is *symmetrically vague* iff P is potentially vague and ‘not P ’ is potentially vague.

However, absolute adjectives display a different pattern. Consider firstly total AAs like *bald* and *empty*. We saw in previous sections that these predicates were potentially vague, and we can think of contexts in which we would assent to the principle of tolerance for these predicates:

(20) **Tolerant *bald* and *empty*:**

- a. For all x, y , if x is bald and x and y ’s heads differ by a single hair, then y is bald.

- b. For all x, y , if x is empty and x and y 's contents differ by a single item, then y is empty.

If adding or subtracting one hair is viewed as an irrelevant change in the context, then whether y has one more or one fewer hair than x will not affect the application of *bald*. The same thing holds for *empty*: if adding or removing an object from a container is viewed as an irrelevant change, then we will always consider y empty if x is.

But we can observe that the negations of total AAs behave differently. In particular, even in the same contexts as described above, the principle of tolerance is not valid for *not bald* and *not empty* (21).

(21) **Intolerant *not bald* and *not empty*:**

- a. **False:** For all x, y , if x is not bald and x and y 's heads differ by a single hair, then y is not bald.
- 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.

The statements in (21) 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: if x has a single hair, it is conceivable that they would be considered *not bald*; however, if y has absolutely no hair, then they would never be considered *not bald*. Similarly with *empty*: (21b) is falsified by the case where x has one object and y has zero objects.

Thus, total AAs and their negations show a fundamental asymmetry with respect to potential vagueness: while it may be possible to find contexts in which an individual who is not completely bald/empty counts as *bald/empty*, someone (or something) who is completely *bald/empty* can never count as *not bald/not empty*. I will refer to the property of differing in vagueness with one's negation as being *asymmetrically vague*:

Definition 4.4.2 Asymmetric vagueness. A predicate P is asymmetrically vague iff one of $\{P, \text{not } P\}$ is not potentially vague.

We can now be more precise about how total predicates are potentially vague as in (22), and the reader is encouraged to verify that this correlation does hold of the entire list of total AAs in (2).

(22) **Total AA Generalization:**

Q is a total AA iff Q is potentially vague and *not* Q is not potentially vague.

Interestingly, although *not* Q does not satisfy tolerance and therefore does not give rise to the Sorites, it still allows for borderline cases. Thus, here we see another argument in favour of a Frege-ian/Wright-ian (i.e. fuzzy boundaries/tolerance) characterization of the essence of vagueness, rather than a Peirce-ian (i.e. borderline cases) characterization.

What about partial AAs? We can immediately see a difference between adjectives like *wet*, *dirty* etc. and *empty*, *bald* etc.: the negations of partial adjectives are potentially vague. For example, if we are in a situation where a single drop of water does not make a difference to our interests, then *not wet* will be tolerant (23a). Similarly with *not dirty*: this negated predicate will satisfy the principle of tolerance in cases where one speck of dirt is perceived as irrelevant (23b).

(23) **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 partial absolute adjectives, it is the positive form of the adjective that is not potentially 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.

(24) **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.

Thus, partial adjectives are also asymmetrically vague and conform to the generalization in (25).

(25) **Partial AA Generalization:**

Q is a partial absolute adjective iff Q is not potentially vague and *not* Q is.

In summary, with the new notion of potential vagueness, we can arrive at a series of new empirical generalizations concerning the distribution of the characterizing properties of vague language across the class of scalar adjectives.

4.5 A Vagueness-Based Characterization of Scale-Structure Distinction

We have now seen a new dimension upon which RAs can be differentiated from AAs: (a)symmetric vagueness. Furthermore, I propose that which scale structure class that an adjective belongs to can be determined through looking only at how the adjective behaves in a Soritical argument: the relative/absolute distinction corresponds to the symmetric/asymmetric vagueness distinction, and the total/partial subdistinction corresponds to the ‘vague positive’/‘vague negation’ distinction.

Furthermore, consider (non-coerced) non-scalar adjectives like *atomic* or *hexagonal*: we saw in chapter 3 that these constituents were typical examples of precise/non-vague predicates, and we can verify that both their positive forms and negative forms are intolerant in all contexts⁶.

(26) **Intolerance of (non-coerced) *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.

(27) **Intolerance of (non-coerced) *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

⁶Note that we are only talking about **non-coerced** non-scalar adjectives. We can observe that coerced non-scalars (i.e. “loose” hexagonal) display the characterizing properties of vague language: how many grooves does an object need to have before it cannot be considered loosely hexagonal?

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.

In summary, 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 4.2:

Pattern	RAs	(total) AAs	(partial) AAs	NSs
P. vague $\neg P$	✓	×	✓	×
P. vague P	✓	✓	×	×

Table 4.2: (Potential) Vagueness Patterns

4.5.1 Summary

In summary, I presented new data concerning the distribution of borderline cases, fuzzy boundaries and susceptibility to the Sorites paradox within the set of scalar and non-scalar adjectives. In the next section, I will give an analysis of the vagueness-based empirical patterns that we saw in this chapter (summarized in table 4.2 (repeated as table 4.3)). Furthermore, I will show that, given the proposed analysis of the data in table 4.3, we immediately have an analysis of the type 2 context-sensitivity data discussed in chapter 2; that is, we will derive the distribution of the *type 2* CS property shown in table 4.4. I therefore conclude that the system that will be proposed in the second part of this chapter can capture the strong link between vagueness and context-sensitivity that we have seen in the data so far.

Pattern	Relative	Total	Partial	Non-Scalar
P. vague $\neg P$	✓	×	✓	×
P. vague P	✓	✓	×	×

Table 4.3: Potential Vagueness Patterns

Pattern	Relative	Total	Partial	Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Table 4.4: Context-Sensitivity Patterns

4.6 Analysis of Vagueness Patterns

In this section, I present an analysis of the vagueness patterns presented in the first part of the chapter. In particular, I extend the delineation system proposed in chapter 2 to analyze the semantics of scalar and non-scalar adjectives with a version of Cobreros et al. (2011a)’s *Tolerant, Classical, Strict* non-classical logic for modelling the puzzling properties of vague language discussed in chapter 3 and above. In other words, in chapter 2, I developed a delineation account of the semantics of adjective phrases, and, in this chapter, I will develop a TCS account of their pragmatics. The section is organized as follows: in section 4.6.1, the semantic analysis proposed in chapter 2. Then, in the rest of the section, I integrate this analysis with a TCS-style analysis of the pragmatic denotations of adjectives. Firstly, in section 4.6.2, I give an analysis of the vagueness/imprecision of absolute scalar adjectives. Then (in sections 4.6.3 and 4.6.4 respectively), I give an analysis of the pragmatic denotations of relative and non-scalar adjectives.

4.6.1 Semantics of (Non)Scalar Adjectives

Recall that, in chapter 2, I proposed that the semantics of relative adjectives differs in an important way from the semantics of both absolute scalar adjectives and non-scalar adjectives: the semantic denotation of RAs can vary depending on which comparison class is chosen by the context, while the semantic denotations of both AAs and NSs are fixed across all comparison classes. This simple proposal was formalized in the following way: Adjectival predicates are interpreted in comparison class models (CC-models), and the interpretation of all adjectival predicates is done with respect to a comparison class (definitions are repeated in (28)).

- (28) a. **Comparison class model.** A CC model is a tuple $\langle D, CC, [\cdot] \rangle$ where D is a non-empty domain of individuals and CC is the set of comparison classes such that $CC =$

$\mathcal{P}(D)$. Furthermore, $\llbracket \cdot \rrbracket$ is an interpretation function that assigns semantic values to syntactically atomic constituents.

1. For an individual denoting subject DP a , $\llbracket a \rrbracket \in D$.

2. For $P \in SA \cup NS$, for every $X \in CC$, $\llbracket P \rrbracket_X \subseteq X$.

b. **Semantics of the positive form.** For a subject DP a , $P \in SA \cup NS$, and some $X \in CC$,

$$\llbracket a \text{ is } P \rrbracket_X = \begin{cases} 1 & \text{if } \llbracket a \rrbracket \in \llbracket P \rrbracket_X \\ 0 & \text{if } \llbracket a \rrbracket \in X - \llbracket P \rrbracket_X \\ i & \text{otherwise} \end{cases}$$

I proposed that the interpretation function is subject to different constraints when it applies to relative adjectives vs when it applies to absolute and non-scalar adjectives. In particular, the interpretation of RAs is constrained only by van Benthem’s ‘coherence’ axioms; whereas, the interpretation of AAs and NSs is constrained by the very powerful *Absolute Adjective Axiom* (AAA). The analysis is summarized in table 4.5.

Axiom	RA	AA	NS
No Reversal (NR)	✓	(✓)	(✓)
Upward Difference (UD)	✓	(✓)	(✓)
Downward Difference (DD)	✓	(✓)	(✓)
Absolute Adjective Axiom (AAA)	×	✓	✓

Table 4.5: Axioms governing the semantic denotation of adjectives (repeated)

Thus, we saw that, with this analysis, we could derive the variation in type 1 context-sensitivity between RAs and AAs/NSs.

4.6.2 Absolute Adjectives

In Klein (1980), vagueness is a semantic phenomenon; that is, Klein adopts a supervaluationist account (cf. Fine (1975), Kamp (1975), Keefe (2000), a.o.) and proposes that the semantic denotation of a scalar term can contain gaps within a comparison class. However, in the literature, there are many arguments in favour of being skeptical of s’valuationist approaches to vagueness in general⁷.

⁷See Hyde (2010), Smith (2008), Fara (2000), and Ripley (2011) and subsequent work for critiques of the s’valuationist approach to vagueness.

Additionally, we might observe that it is not obvious how to account for the phenomenon of asymmetric vagueness presented in this chapter within an s'valuationist (i.e. super or subvaluationist) approach because, if, as s'valuationists claim, vagueness is the result of deficiency of meaning, it is not clear why the meaning of a p. vague predicate *empty* would be deficient, but its precise negation *not empty* would not be deficient. Therefore, despite adopting a version of Klein's semantic analysis of relative adjectives, I adopt a different approach to adjectival vagueness.

Instead, the basic idea is to build pragmatic structures on top of semantic structures in the way done in Cobreros et al. (2011a), and use the properties of this logic to model the properties of vague language. We therefore extend our CC models to *tolerant* models by adding the function \sim in the way shown in definition 4.6.1.

Definition 4.6.1 Comparison Class t-model. $\langle D, CC, \llbracket \cdot \rrbracket, \sim \rangle$, where $\langle D, CC, \llbracket \cdot \rrbracket \rangle$ is a CC c-model and \sim is a two-place function that maps an ordered pair $\langle X, P \rangle$ to a binary relation on X : \sim_P^X .

There are two ways in which this definition differs from the original definitions in Cobreros et al. (2011a) (presented in chapter 3): Firstly, note that now, instead of mapping a predicate to an indifference relation as in classical TCS (see chapter 3), \sim maps a predicate and a comparison class to an indifference relation on the members of the class. Thus, indifference relations are also relativized to comparison classes. Secondly, in chapter 3, we immediately put constraints of reflexivity and symmetry on the \sim ps. In what follows, we will put similar constraints on the definition of \sim ; however, due to the fact that they now relate elements of comparison classes, these constraints will be more complicated. I will present each proposed constraint in detail below; however, we first define tolerant and strict denotations (relativized to comparison classes) as in definition 4.6.2.

Definition 4.6.2 Tolerant/Strict CC denotations. For $P \in SA \cup NS$ and $X \in CC$,

1. $\llbracket P \rrbracket_X^t = \{x : \exists d \sim_P^X x : d \in \llbracket P \rrbracket_X\}$.

2. $\llbracket P \rrbracket_X^s = \{x : \forall d \sim_P^X x, d \in \llbracket P \rrbracket_X\}$.

Finally, the tolerant/strict semantics for the positive form of an adjective with respect to a comparison class can be given as in 4.6.3.

Definition 4.6.3 Positive form. For a contextually given $X \in CC$,

$$1. \llbracket a \text{ is } P \rrbracket_X^t = \begin{cases} 1 & \text{if } \llbracket a \rrbracket \in \llbracket P \rrbracket_X^t \\ 0 & \text{if } \llbracket a \rrbracket \in X - \llbracket P \rrbracket_X^t \\ i & \text{otherwise} \end{cases}$$

$$2. \llbracket a \text{ is } P \rrbracket_X^s = \begin{cases} 1 & \text{if } \llbracket a \rrbracket \in \llbracket P \rrbracket_X^s \\ 0 & \text{if } \llbracket a \rrbracket \in X - \llbracket P \rrbracket_X^s \\ i & \text{otherwise} \end{cases}$$

Furthermore, we can define the tolerant comparative ($>_p^t$) and strict comparative ($>_p^s$) relations as follows:

Definition 4.6.4 Comparative form.

1. $\llbracket a \text{ is } P\text{-er than } b \rrbracket^t = 1$ iff $\llbracket a \rrbracket >_p^t \llbracket b \rrbracket$ iff there is some $X \in CC$ such that $a \in \llbracket P \rrbracket_X^t$ and $b \in X - \llbracket P \rrbracket_X^t$.
2. $\llbracket a \text{ is } P\text{-er than } b \rrbracket^s = 1$ iff $\llbracket a \rrbracket >_p^s \llbracket b \rrbracket$ iff there is some $X \in CC$ such that $a \in \llbracket P \rrbracket_X^s$ and $b \in X - \llbracket P \rrbracket_X^s$.

The analysis in definitions 4.6.3 and 4.6.4 is simply a straightforward implementation of the Klein-ian approach to the semantics of scalar adjectives within a TCS account of the pragmatics of vague predicates. Thus, in a particular situation with a particular comparison class X , a predicate P of either semantic class can have borderline cases (objects that in both $\llbracket P \rrbracket_X^t$ and $\llbracket \text{not } P \rrbracket_X^t$), and, provided that we put appropriate restrictions on the properties of \sim , members of X are can be related by \sim_P^X in a way that forms a Soritical series; therefore, we can construct a Sorites paradox. In the next section, I discuss what kinds of restrictions \sim should be subject to.

4.6.2.1 The Properties of Indifference

As it stands, we have not placed any constraints on the definition of \sim . However, like the case discussed in chapter 2 with the interpretation of relative predicates, if we do not say anything about

how indifference relations can be established across comparison classes, the \sim_{PS} 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 comparison classes. 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*⁸. This being said, it is important to note that the precise question of indifference or even similarity within and across adjectival comparison classes 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 indifference/close similarity into the logical system developed in this dissertation, one that will be appropriate for modelling the ‘loose’ uses of potentially vague predicates.

The first property that is generally proposed to characterize indifference/similarity relations is reflexivity (cf. Luce (1956), Pogonowski (1981), Cobreros et al. (2011a), among many others). Intuitively, every individual is indifferent from itself. Thus, we adopt the constraint in (29) that enforces reflexivity across CCs.

(29) **Reflexivity (R):** For all $P \in SA \cup NS$, all $X \in CC$ and all $x \in X$, $x \sim_P^X x$.

In addition to being reflexive, indifference and similarity relations are generally proposed to be symmetric (consider, for example, the original formulation of TCS presented in chapter 3). 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, judgements of similarity are directional (ex. 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 judgements of similarity and indifference particularly involve relations between individu-

⁸Presumably, indifference relations are just sub-types of similarity relations. That is, a and b are indifferent with respect to a property just in case they have an extremely high degree of similarity with respect to that property.

als that differ in terms of ‘prototypicality’ (cf. Tversky (1977), Rosch (1978), Ortony et al. (1985), Lakoff (1987)). The generalization concerning asymmetric judgements of similarity can be stated (in the words of Ortony et al. (1985) (p. 570)) as follows:

(30) **Prototypicality Generalization:**

Atypical members of categories tend to be judged as more similar to typical members than the other way around.

(30) is a robust generalization that has been observed in studies of judgements 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 (31a) over (31b), and similar results were obtained for pairs of sentences like (32) and (33).

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

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

- (33) a. Luxemburg is similar to Belgium.
b. Belgium is similar to Luxemburg.

This discussion of asymmetric similarity judgements is important for the present purposes because, in section 4.4, I argued that we saw a similar asymmetry in judgements of indifference with absolute adjectives. In particular, we saw that, with total AAs 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 (31a) and (31b) in (34).

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

We also saw that partial AAs display the opposite pattern:

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

In summary, I believe that it is a reasonable hypothesis that the patterns with AAs discussed in section 4.4 are instances of a more general phenomenon in which prototypical members of a predicate's denotation have a different status than less prototypical members⁹. Formally, I propose that these asymmetries are encoded into the indifference relations associated with total and partial AAs by means of the following two pragmatic axioms¹⁰. (Since the indifference relation is now not necessarily symmetric, $a \sim b$ can now be read as $\langle a, b \rangle \in \sim_Q^X$, 'b can count as a' or 'b approximates a').

(36) **Total Axiom:**

If Q is a total adjective ($Q \in AA^T$), then, for all $x, y \in D$, if $x \in \llbracket Q \rrbracket_D$ and $y \notin \llbracket Q \rrbracket_D$, then $y \not\sim_Q^X x$, for all $X \in CC$.

•Otherwise, if $y \sim_Q^X x$, then $x \sim_Q^X y$.

(37) **Partial Axiom:**

⁹Note that I am not claiming that every member of the denotation of an AA is necessarily a prototypical member, particularly in the case of partial AAs (this may be a possibility for total AAs). See Armstrong et al. (1983), Kamp and Partee (1995) and Osherson and Smith (1997) for a discussion about the relationship between a predicate's semantic denotation and the set of its typical members.

¹⁰A 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.

If Q is a partial adjective ($Q \in AA^P$), then, for all $x, y \in D$, if $x \in \llbracket Q \rrbracket_D$ and $y \notin \llbracket Q \rrbracket_D$, then $x \not\sim_Q^X y$, for all $X \in CC$.

•Otherwise, if $x \sim_Q^X y$, then $y \sim_Q^X x$.

The previous axioms made a distinction between the two subclasses of AAs that were identified; however, the rest of the pragmatic axioms that I will propose will apply to all AAs and, indeed, to all adjectival predicates in the same way. The first general axiom that I propose is called *tolerant no skipping*:

(38) **Tolerant No Skipping (T-NS):** For all $P \in SA \cup NS$, all $X \in CC$, and all $x, y \in X$, if $x \sim_P^X y$ and there is some $z \in X$ such that $x \geq_P^t z \geq_P^t y$, then $x \sim_P^X z$.

Where $x \geq_P^t y$ iff $x >_P^t y$ or $x \approx_P^t y$. Note that we can extend the definition of the equivalence relation \approx_P to the relations \approx_P^t and \approx_P^s as below:

Definition 4.6.5 Equivalent. (\approx) For $P \in SA \cup NS$, $a, b \in D$, and $x \in \{t, s\}$, $a \approx_P^x b$ iff $a \not\prec_P^x b$ and $b \not\prec_P^x a$.

Tolerant No Skipping 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. As we will see in the next section, T-NS performs a very similar function to van Benthem's *No Reversal*.

I propose a second axiom that is, in some sense (to be discussed below), the dual of T-NS: *Strict No Skipping*:

(39) **Strict No Skipping (S-NS):** For all $P \in SA \cup NS$, all $X \in CC$, and all $x, y \in X$, if $x \sim_P^X y$ and there is some $z \in X$ such that $x \geq_P^s z \geq_P^s y$, then $z \sim_P^X y$.

Strict No Skipping 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.

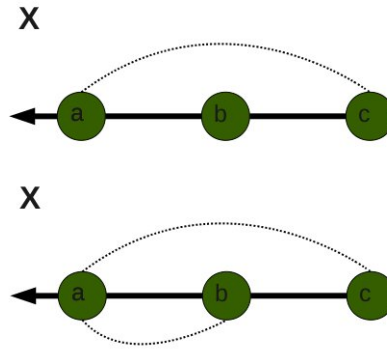


Figure 4.3: Tolerant No Skipping

The next axiom deals with how indifference relations can change across comparison classes. At the moment, \sim_{PS} can be established and destroyed in different comparison classes in a more or less arbitrary way, provided that R, TA, and PA are respected. But presumably we might want some more ‘coherence’ in the distribution of the \sim_{PS} . I therefore propose the following ‘granularity’ constraint:

- (40) **Granularity (G).** For all $P \in SA \cup NS$, all $X \in CC$ and all $x, y \in X$, if $x \sim_P^X y$, then for all $X' : X \subseteq X', x \sim_P^{X'} y$.

Granularity says that if person A and person B are indistinguishable in comparison class X , then they are indistinguishable in all supersets of X . This is meant to reflect the fact that the larger the domain is (i.e. the larger the comparison class is), the more things can cluster together. In other words, the larger the comparison class is, the more it is possible to collapse fine distinctions that were made in smaller comparison classes, and once you collapse such a ‘fine-grained’ distinction, you cannot make it again at a more ‘coarse-grained’ level¹¹. It is loosely inspired by theories of granularity such as Hobbes (1985) in which distinctions made in a more complex theory are collapsed in more simple ‘coarse-grained’ theories.

While *granularity* talks about how indifference is preserved, the final two axioms deal with the

¹¹One might wonder whether *Granularity* is perhaps a bit strong. Maybe there are some (albeit) restricted situations in which the addition of new individuals into a comparison class would cause two individuals that were indifferent in a smaller comparison class to be distinguished in a larger one. For a proposal that contains a weaker version of *granularity* that would allow such cases and a discussion of its effect on the relations denoted by absolute comparatives, see Burnett (2012c).

preservation of differences across comparison classes.

- (41) **Minimal Difference (MD).** For all $P \in SA \cup NS$, and all $x, y \in D$, if $x >_P y$, then $x \not\sim_P^{\{x,y\}} y$.
- (42) **Contrast Preservation (CP).** For all $P \in SA \cup NS$ all $X, X' \in CC$, if $X \subset X'$ and, for $x, y \in X$, $x \not\sim_P^X y$ and $x \sim_{P'}^{X'} y$, then $\exists z \in X' - X : x \not\sim_{P'}^{X'} z$.

Minimal Difference 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 *Downward Difference* because it allows us to preserve contrasts down to the smallest comparison classes.

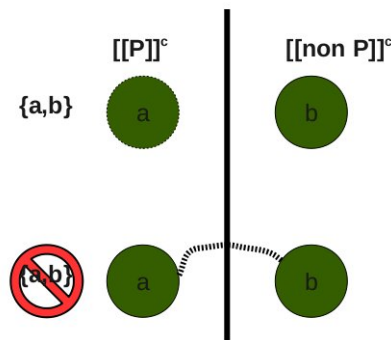


Figure 4.4: Minimal difference

Contrast Preservation 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. This axiom is similar in spirit to van Benthem's *Upward Difference* in that it ensures that, if there is a contrast/distinction in one comparison class, the existence of a contrast is maintained in all the larger CCs.

In summary, I proposed that the definition of \sim is constrained by the axioms in table 4.6.

4.6.2.2 Context-Sensitivity Results

Although I have put many constraints on the \sim function, all of them combined have a much weaker effect on the pragmatic denotations of AAs than the AAA had on their semantics denotations. In

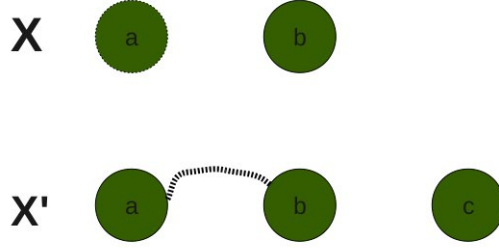


Figure 4.5: Contrast Preservation (CP)

Axiom	Total AA	Partial AA
Reflexivity (R)	✓	✓
Total Axiom (TA)	✓	×
Partial Axiom (PA)	×	✓
Tolerant No Skipping (T-NS)	✓	✓
Strict No Skipping (S-NS)	✓	✓
Granularity (G)	✓	✓
Minimal Difference (MD)	✓	✓
Contrast Preservation (CP)	✓	✓

Table 4.6: Pragmatic Axioms for Absolute Adjectives

particular, the analysis given in 4.6 still allows for the pragmatic denotations of AAs to vary across comparison classes. Thus we are now in a position to have an analysis of the distribution of the type 2 context-sensitivity property described in chapter 2. We first adopt the formal definition of type 2 CS in definition 4.6.6: informally, a predicate is type 2 context-sensitive if either its tolerant or its strict denotation can vary across comparison classes.

Definition 4.6.6 Type 2 Context-Sensitivity *An adjective P is type 2 context-sensitive iff there is some CC t-model M in which, for some $x \in D$, there is some $X \in CC$ such that, for $n \in \{t, s\}$, $x \in \llbracket P \rrbracket_X^n$ and there is some distinct $X' \in CC$ such that $x \notin \llbracket P \rrbracket_{X'}^n$.*

With this definition, we can now show that both total AAs and partial AAs are type 2 context-sensitive. (Note that, as shown in chapter 2, they are not type 1 context-sensitive).

Theorem 4.6.1 *If $Q \in AA^T$, then Q is type 2 context-sensitive.*

Proof Let M be a CC t-model: $\langle \{a, b, c\}, CC, \sim, \llbracket \cdot \rrbracket \rangle$. Let $Q \in AA^T$ such that:

1. $\llbracket Q \rrbracket_{\{\}} = \{\}$.
2. $\llbracket Q \rrbracket_{\{a\}} = \{a\}$.
3. $\llbracket Q \rrbracket_{\{b\}} = \{\}$.
4. $\llbracket Q \rrbracket_{\{c\}} = \{\}$.
5. $\llbracket Q \rrbracket_{\{a,b\}} = \{a\}$.
6. $\llbracket Q \rrbracket_{\{b,c\}} = \{\}$.
7. $\llbracket Q \rrbracket_{\{a,c\}} = \{a\}$.
8. $\llbracket Q \rrbracket_{\{a,b,c\}} = \{a\}$.

So $\llbracket \cdot \rrbracket$ satisfies the AAA. Then, let \sim be defined as follows:

1. $\sim(Q, \{\}) = \{\}$.
2. $\sim(Q, \{a\}) = \{\langle a, a \rangle\}$.
3. $\sim(Q, \{b\}) = \{\langle b, b \rangle\}$.
4. $\sim(Q, \{c\}) = \{\langle c, c \rangle\}$.
5. $\sim(Q, (\{a, b\})) = \{\langle a, a \rangle, \langle b, b \rangle\}$.
6. $\sim(Q, (\{b, c\})) = \{\langle c, c \rangle, \langle b, b \rangle\}$.
7. $\sim(Q, (\{a, c\})) = \{\langle a, a \rangle, \langle c, c \rangle\}$.
8. $\sim(Q, (\{a, b, c\})) = \{\langle a, a \rangle, \langle b, b \rangle, \langle c, c \rangle, \langle \mathbf{a}, \mathbf{b} \rangle\}$.

In this model, there is some CC $(\{a, b, c\})$ such that $b \in \llbracket Q \rrbracket_{\{a,b,c\}}^t$ and there is some CC $(\{a, b\})$ such that $b \notin \llbracket Q \rrbracket_{\{a,b\}}^t$. Therefore, by definition 4.6.6, Q is type 2 context-sensitive. \square

Theorem 4.6.2 *If $Q \in AA^P$, then Q is type 2 context-sensitive.*

Proof Let M be a CC t-model: $\llbracket \{a, b, c\}, CC, \sim, \llbracket \cdot \rrbracket \rrbracket$. Let $Q \in AA^P$ such that:

1. $\llbracket Q \rrbracket_{\{\}} = \{\}$.
2. $\llbracket Q \rrbracket_{\{a\}} = \{a\}$.
3. $\llbracket Q \rrbracket_{\{b\}} = \{b\}$.
4. $\llbracket Q \rrbracket_{\{b\}} = \{\}$.
5. $\llbracket Q \rrbracket_{\{a,b\}} = \{a, b\}$.
6. $\llbracket Q \rrbracket_{\{b,c\}} = \{b\}$.
7. $\llbracket Q \rrbracket_{\{a,c\}} = \{a\}$.
8. $\llbracket Q \rrbracket_{\{a,b,c\}} = \{a, b\}$.

So $\llbracket \cdot \rrbracket$ satisfies the AAA. Then, let \sim be defined as follows:

1. $\sim (Q, \{\}) = \{\}$.
2. $\sim (Q, \{a\}) = \{\langle a, a \rangle\}$.
3. $\sim (Q, \{b\}) = \{\langle b, b \rangle\}$.
4. $\sim (Q, \{c\}) = \{\langle c, c \rangle\}$.
5. $\sim (Q, (\{a, b\})) = \{\langle a, a \rangle, \langle b, b \rangle\}$.
6. $\sim (Q, (\{b, c\})) = \{\langle c, c \rangle, \langle b, b \rangle\}$.
7. $\sim (Q, (\{a, c\})) = \{\langle a, a \rangle, \langle c, c \rangle\}$.
8. $\sim (Q, (\{a, b, c\})) = \{\langle a, a \rangle, \langle b, b \rangle, \langle c, c \rangle, \langle \mathbf{c}, \mathbf{b} \rangle\}$.

In this model, there is some CC $(\{b, c\})$ such that $b \in \llbracket Q \rrbracket_{\{b,c\}}^s$ and there is some CC $(\{a, b, c\})$ such that $b \notin \llbracket Q \rrbracket_{\{a,b,c\}}^s$. Therefore, by definition 4.6.6, Q is type 2 context-sensitive. \square

In summary, we now have an analysis for the CS patterns associated with both partial and total AAs (repeated in table 4.7).

Pattern	Relative	Total	Partial	Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Table 4.7: Context-Sensitivity Patterns

4.6.2.3 From Context-Sensitivity to Scalarity

In chapter 2, it was shown that the constraints placed on the application of relative and absolute predicates across comparison classes had important effects on the scales that were associated with them via the semantics of the comparative relation. In the TCS extension of the Klein-ian system that I proposed in this chapter, in addition to a semantic denotation, all adjectival predicates are assigned two pragmatic denotations: a tolerant denotation ($\llbracket \cdot \rrbracket^t$) and a strict denotation ($\llbracket \cdot \rrbracket^s$). Furthermore, we defined the tolerant and strict comparative relations ($>_p^t$) and ($>_p^s$) in a parallel manner to the regular semantic denotation of the comparative ($>_p$) (cf. definition 4.6.4). 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 ‘coherence’ 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. In other words, the proposals that I made about how indifference relations can be established and change across comparison classes give us a pragmatic solution to the paradox of absolute adjectives.

As a short illustration of how this works, consider the following example: if I compare Homer Simpson, who has exactly two hairs, with Yul Brynner (figure 4.6), the two would not be considered indifferent with respect to baldness (Homer has hair!).

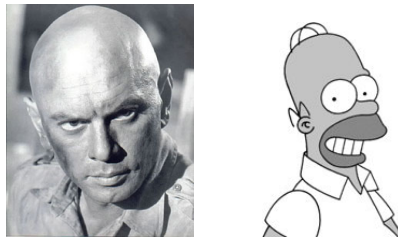


Figure 4.6: Yul Brynner vs Homer Simpson

However, if I add Marge Simpson into the comparison class (figure 4.7), then Yul and Homer

start looking much more similar, when it comes to baldness.



Figure 4.7: Yul Brynner vs Homer Simpson vs Marge Simpson

However, given our axioms, it would never be the case that, in some CC, Yul and Marge could be considered to be indifferent with respect to baldness, but not Homer. Thus, I propose that it should be possible to order individuals with respect to how close to being completely bald they are by looking at in which comparison classes they are considered indifferent to completely bald people.

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.118-119),

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 (1999)’s pragmatic halo framework

that was outlined in chapter 3. 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 orders¹². More importantly, Récanati proposes that the relation between *empty*₁ and *empty*₂ is that of homophony, meaning that, in the lexicon, for every absolute adjective, there are two versions, a non-scalar version (the absolute version) and a scalar version (presumably a relative adjective¹³). In addition to being rather inelegant, this analysis would seem to make wrong predictions with respect to the distribution of the relative adjective *empty*₁ 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 (43) could have the reading in (43a).

- (43) Pass me the empty one.
- a. Pass me the *empty*₁ one. (less full)
 - b. Pass me the *empty*₂ one. (completely empty)

But clearly, it is not possible to use (43) in this way¹⁴. I therefore conclude that a homophony analysis of the gradability of AAs is empirically insufficient.

But 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 *empty*₁ (gradable *empty*) and *empty*₂ (non-gradable *empty*) is simply the difference between the pragmatic and semantic denotations of a single lexical item, *empty*. In the rest of this subsection, I will show that, by adopting the axioms that characterize AAs presented in table 4.6, we can associate non-trivial scales with absolute predicates.

¹²Although see Bale (2011) for such a proposal.

¹³Of 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 *empty*₁ is an absolute adjective; however, if he did so, it would be unclear why we would need two *emptys* in the first place.

¹⁴Note that Récanati (2010)'s main proposal, which is that the gradability of absolute adjectives 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.

We can first note a series of facts about the $>_Q^t$ and $>_Q^s$ relations and their relationship to $>_Q$. For example, Minimal Difference (MD) ensures that classical absolute denotations are subsets of tolerant denotations:

Theorem 4.6.3 Tolerant Subset. *If $Q \in AA$, then $>_Q \subseteq >_Q^t$.*

Proof Let $x, y \in D$ such that $x >_Q y$ to show that $x >_Q^t y$. Since $x >_Q y$, there is some $X \in CC$ such that $x \in \llbracket Q \rrbracket_X$ and $y \notin \llbracket Q \rrbracket_X$. Now consider $\{x, y\} \in CC$. By downward difference, $x \in \llbracket Q \rrbracket_{\{x, y\}}$ and $y \notin \llbracket Q \rrbracket_{\{x, y\}}$. By the definition of $\llbracket \cdot \rrbracket^t$, $x \in \llbracket Q \rrbracket_{\{x, y\}}^t$. Furthermore, by Minimal Difference, $x \not\sim_Q^{\{x, y\}} y$. So $y \notin \llbracket Q \rrbracket_{\{x, y\}}^t$. By the definition of $>_Q^t$ (definition 4.6.4), $x >_Q^t y$. \square

This is a welcome result, since it shows that the classical and tolerant denotations of absolute comparatives are in the same relationship as the classical and tolerant denotations of the positive forms of the adjectives, namely inclusion (cf. Cobreros et al. (2011a)'s Corollary 1). However, a difference between the system described here and the original TCS is that classical denotations of absolute comparatives are also included in the strict denotations of these constituents:

Theorem 4.6.4 Strict Subset. $>_Q \subseteq >_Q^s$.

Proof Let $x >_Q y$ to show $x >_Q^s y$. Since $x >_Q y$, by two-element reducibility, $x \in \llbracket Q \rrbracket_{\{x, y\}}$ and $y \notin \llbracket Q \rrbracket_{\{x, y\}}$. Therefore, by MD, $x \not\sim_Q^{\{x, y\}} y$. So, by the definition of $\llbracket \cdot \rrbracket^s$, $x \in \llbracket Q \rrbracket_{\{x, y\}}^s$ and $y \notin \llbracket Q \rrbracket_{\{x, y\}}^s$. So $x >_Q^s y$. \square

Secondly, with only T/S-No Skipping, we can prove that van Bentham's No Reversal holds at the tolerant and strict levels. It is in this sense that, as I mentioned, T/S-NS can be viewed as the tolerant/strict correspondent of *No Reversal*.

Theorem 4.6.5 No Tolerant Reversal (T-NR): *For $X \in CC$, if $x \in \llbracket Q \rrbracket_X^t$ and $y \notin \llbracket Q \rrbracket_X^t$, then there is no $X' \in CC$ such that $y \in \llbracket Q \rrbracket_{X'}^t$ and $x \notin \llbracket Q \rrbracket_{X'}^t$.*

Proof Suppose $x \in \llbracket Q \rrbracket_X^t$ and $y \notin \llbracket Q \rrbracket_X^t$. Suppose, for a contradiction that there is an $X' \in CC$ such that $y \in \llbracket Q \rrbracket_{X'}^t$ and $x \notin \llbracket Q \rrbracket_{X'}^t$. Therefore, $x >_Q^t y$ and $y >_Q^t x$. Furthermore, by assumption and the

definition of $\llbracket Q \rrbracket_X^t$, there is some $d \sim_Q^X x$ such that $d \in \llbracket Q \rrbracket_X$, and $d \not\sim_Q^X y$. Thus $d >_Q^t y$ and so $d >_Q^t y >_Q^t x$. Since $d \sim_Q^X x$, by No Skipping, $d \sim_Q^X y$. \perp \square

Theorem 4.6.6 No Strict Reversal. *For some $X \in CC$, let $x \in \llbracket Q \rrbracket_X^s$ and $y \notin \llbracket Q \rrbracket_X^s$. Then, there is no distinct $X' \in \llbracket Q \rrbracket_{X'}^s$, such that $y \in \llbracket Q \rrbracket_{X'}^s$, and $x \notin \llbracket Q \rrbracket_{X'}^s$.*

Proof Suppose, for a contradiction that there is an $X' \in CC$ such that $y \in \llbracket P \rrbracket_{X'}^s$ and $x \notin \llbracket P \rrbracket_{X'}^s$. Since $x \notin \llbracket Q \rrbracket_{X'}^s$, there is some $d \sim_Q^{X'} x$ such that $d \notin \llbracket Q \rrbracket_{X'}$. Therefore, since $y \in \llbracket Q \rrbracket_{X'}^s$, $y >_Q^s d$. By assumption $x >_Q^s y >_Q^s d$, and $x \sim_Q^{X'} d$. Therefore, by S-NS, $y \sim_Q^{X'} d$. But $y \in \llbracket Q \rrbracket_{X'}^s$. \perp \square

Using the complete axiom set, we can show that, for all $Q \in AA$, $>_Q^t$ is a strict weak order (irreflexive, transitive and almost-connected).

Lemma 4.6.7 Irreflexivity. *For all $x \in D$, $x \not>_Q^t x$.*

Proof Since it is impossible, for any $X \in CC$, for an element to be both in $\llbracket Q \rrbracket_X^t$ and not in $\llbracket Q \rrbracket_X^t$, by the definition of $\llbracket \rrbracket^t$, $>_Q^t$ is irreflexive.

We now prove transitivity for $>_Q^t$.

Lemma 4.6.8 Transitivity. *For all $x, y, z \in D$, if $x >_Q^t y$ and $y >_Q^t z$, then $x >_Q^t z$.*

Proof Suppose $x >_Q^t y$ and $y >_Q^t z$ to show that $x >_Q^t z$. Then there is some $X \in CC$ such that $x \in \llbracket Q \rrbracket_X^t$ and $y \notin \llbracket Q \rrbracket_X^t$. Thus, there is some $d \in \llbracket Q \rrbracket_X$ such that $d \sim_Q^X x$. Now consider $X \cup \{z\}$. By the AAA and the assumption that $x >_Q^t y$ and $y >_Q^t z$, $y, z \notin \llbracket Q \rrbracket_{X \cup \{z\}}$.

Case 1: $X \cup \{z\} = X$. Since $x \in \llbracket Q \rrbracket_X^t$ and $z \notin \llbracket Q \rrbracket_X^t$, $x >_Q^t z$. \checkmark

Case 2: $X \subset X \cup \{z\}$. Since $X \subset X \cup \{z\}$ and $d \sim_Q^X x$, by G1', $d \sim_Q^{X \cup \{z\}} x$. By the AAA, $d \in \llbracket Q \rrbracket_{X \cup \{z\}}$. So $x \in \llbracket Q \rrbracket_{X \cup \{z\}}^t$. Suppose, for a contradiction that $z \in \llbracket Q \rrbracket_{X \cup \{z\}}^t$. Then there is some $d' \in \llbracket Q \rrbracket_{X \cup \{z\}}$ such that $d' \sim_Q^{X \cup \{z\}} z$. By assumption and since $y \notin \llbracket Q \rrbracket_X$, by MD, $d' >_Q^t y >_Q^t z$. So by Tolerant No Skipping, $d' \sim_Q^{X \cup \{z\}} y$. Since $y \notin \llbracket Q \rrbracket_X^t$, $d' \not\sim_Q^X y$. So by G2, since $X \cup \{z\} - X = \{z\}$, $d' \not\sim_Q^{X \cup \{z\}} z$. \perp . So $z \notin \llbracket Q \rrbracket_{X \cup \{z\}}^t$, and $x >_Q^t z$. \checkmark \square

Finally, we can prove almost connectedness.

Lemma 4.6.9 Almost Connected. *For all $x, y \in D$, if $x >^t_Q y$ then for all $z \in D$, either $x >^t_Q z$ or $z >^t_Q y$.*

Proof Let $x >^t_Q y$ and $z \not>^t_Q y$ to show $x >^t_Q z$.

Case 1: $x \in \llbracket Q \rrbracket_D$. Since $x >^t_Q y$ and $z \not>^t_Q y$, $z \notin \llbracket Q \rrbracket_D$. So $x >_Q z$, and, by theorem 4.6.3, $x >^t_Q z$. ✓

Case 2: $x \notin \llbracket Q \rrbracket_D$. Since $x >^t_Q y$, there is some $X \in CC$ such that $x \in \llbracket Q \rrbracket'_X$ and $y \notin \llbracket Q \rrbracket'_X$. So there is some $d \in \llbracket Q \rrbracket_X$ such that $d \sim^X_Q x$. Consider $X \cup \{z\}$. Since $z \not>^t_Q y$, $x, y, z \notin \llbracket Q \rrbracket_{X \cup \{z\}}$. Since $d \sim^X_Q x$, by Granularity, $d \sim^{X \cup \{z\}}_Q x$ and by the AAA, $d \in \llbracket Q \rrbracket'_{X \cup \{z\}}$. So $x \in \llbracket Q \rrbracket'_{X \cup \{z\}}$. Now suppose for a contradiction that $z \in \llbracket Q \rrbracket'_{X \cup \{z\}}$. Then there is some $d' \in \llbracket Q \rrbracket_{X \cup \{z\}}$ such that $d' \sim^{X \cup \{z\}}_Q z$. Since $d' \in \llbracket Q \rrbracket_{X \cup \{z\}}$ and $y \notin \llbracket Q \rrbracket_{X \cup \{z\}}$, $d' >_Q y$; so by theorem 4.6.3, $d' >^t_Q y$. Furthermore, since, by assumption, $z \not>^t_Q y$, $y \geq^t_Q z$. Since $d' \geq^t_Q y \geq^t_Q z$ and $d' \sim^{X \cup \{z\}}_Q z$, by Tolerant No Skipping, $d' \sim^{X \cup \{z\}}_Q y$. However, since $y \notin \llbracket Q \rrbracket'_X$, and by the AAA, $d' \in \llbracket Q \rrbracket_X$, $d' \not\sim^X_Q y$. Since $X \subset X \cup \{z\}$ and $d' \sim^{X \cup \{z\}}_Q y$, by Contrast Preservation, there is some $a \in X \cup \{z\} - X$ such that $d' \not\sim^{X \cup \{z\}}_Q a$. Since $X \cup \{z\} - X = \{z\}$, $d' \not\sim^{X \cup \{z\}}_Q z$. ⊥ So $z \notin \llbracket Q \rrbracket'_{X \cup \{z\}}$ and $x >^t_Q z$. ✓ □

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

Theorem 4.6.10 *If $Q \in AA$, $<^t_Q$ is a strict weak order.*

Proof Immediate from lemmas 4.6.7, 4.6.8 and 4.6.9. □

Using the same axiom set, we can also show that the strict scale ($>^s_Q$) is a strict weak order.

Lemma 4.6.11 Irreflexivity. *If $Q \in AA$, $>^s_Q$ is irreflexive.*

Proof Immediately from the definition of $>^s_Q$. □

Lemma 4.6.12 Transitivity. *If $Q \in AA$, $>^s_Q$ is transitive.*

Proof Suppose $x >_Q^s y$ and $y >_Q^s z$ to show $x >_Q^s z$. Since $y >_Q^s z$, there is some $X \in CC$ such that $y \in \llbracket Q \rrbracket_X^s$ and $z \notin \llbracket Q \rrbracket_X^s$. So there is some $d \sim_Q^X z$ such that $d \notin \llbracket Q \rrbracket_X$. Clearly, $y >_Q^s d$. Now consider $X \cup \{x\}$. By the AAA, $d \notin \llbracket Q \rrbracket_{X \cup \{x\}}$ and, by Granularity, $z \sim_Q^{X \cup \{x\}} d$. So $z \notin \llbracket Q \rrbracket_{X \cup \{x\}}^s$. Suppose for a contradiction that $x \notin \llbracket Q \rrbracket_{X \cup \{x\}}^s$. So there is some $d \notin \llbracket Q \rrbracket_{X \cup \{x\}}$ such that $x \sim_Q^{X \cup \{x\}} d$. Since $x >_Q^s y$, $x \in \llbracket Q \rrbracket_{X \cup \{x\}}$. So $x >_Q d$ and by theorem 4.6.4, $x >_Q^s d$. So $x >_Q^s y >_Q^s d$. Since $x \sim_Q^{X \cup \{x\}} d$, by S-NS, $y \sim_Q^{X \cup \{x\}} d$. Since $y \in \llbracket Q \rrbracket_X^s$, $y \not\sim_Q^X d$. So, by G2, $x \not\sim_Q^{X \cup \{x\}} d$. \perp So $x \in \llbracket Q \rrbracket_{X \cup \{x\}}^s$. \square

Lemma 4.6.13 Almost-Connectedness. *If $Q \in AA$, $>_Q^s$ is almost-connected.*

Proof Suppose $x >_Q^s y$ and $z \not>_Q^s y$ to show $x >_Q^s z$. Then there is some $X \in CC$ such that $x \in \llbracket Q \rrbracket_X^s$ and $y \notin \llbracket Q \rrbracket_X^s$. Now consider $X \cup \{z\}$. By the AAA and G1', $y \notin \llbracket Q \rrbracket_{X \cup \{z\}}^s$. Since $z \not>_Q^s y$, $z \in \llbracket Q \rrbracket_{X \cup \{z\}}^s$.

Case 1: $x \in \llbracket Q \rrbracket_{X \cup \{z\}}^s$. Then $x >_Q^s z$. \checkmark

Case 2: $x \notin \llbracket Q \rrbracket_{X \cup \{z\}}^s$. Then there is some $d \sim_Q^{X \cup \{z\}} x$ such that $d \notin \llbracket Q \rrbracket_{X \cup \{z\}}$.

Suppose for a contradiction that $d \neq z$. So $x, d \in X$. Since $x \in \llbracket Q \rrbracket_X^s$, $x \not\sim_Q^X d$. So, by CP, $d \not\sim_Q^{X \cup \{z\}} z$. Since $d \notin \llbracket Q \rrbracket_{X \cup \{z\}}$, for all $a \in X \cup \{z\}$, $a \geq_Q^s d$. So $y \geq_Q^s d$ and $z \geq_Q^s d$. Since $x >_Q^s y \geq_Q^s d$ and $x \sim_Q^{X \cup \{z\}} d$, by S-NS, $y \sim_Q^{X \cup \{z\}} d$. Since $z \not>_Q^s y$, $y \geq_Q^s z$, so $y \geq_Q^s z \geq_Q^s d$. Since $y \sim_Q^{X \cup \{z\}} d$, by S-NS, $d \sim_Q^{X \cup \{z\}} z$. \perp So $d = z$. Since $x >_Q^s y$, $x \in \llbracket Q \rrbracket_{X \cup \{z\}}$, so $x >_Q z$ and, by theorem 4.6.4, $x >_Q^s z$. \checkmark \square

The second main theorem of this section is the following:

Theorem 4.6.14 *If $Q \in AA$, $>_Q^s$ is a strict weak order.*

Proof Immediately from lemmas 4.6.11, 4.6.12, and 4.6.13. \square

In summary, the axioms presented in this section allow us to extract strict weak orders from the behaviour of tolerant and strict denotations across comparison classes. Additionally, we can see that, unlike the orders extracted from the semantic denotations of AAs, the tolerant orders associated with total AAs and the strict orders associated with partial AAs are not necessarily trivial. Consider, for instance, the model in the proof of theorem 4.6.1. In this model, $a >_Q^t b >_Q^t c$, despite the fact that $b \not>_Q c$. That is, if we define the pragmatic ‘degree’ scale associated with Q as

in (44a), the tolerant scale associated with this predicate will be $\{a\} \gg_Q^t \{b\} \gg_Q^t \{c\}$, while the degree scale associated with its semantic denotation will be $\{a\} \gg_Q \{b, c\}$.

Definition 4.6.7 *Pragmatic degree ordering* ($\gg^{t/s}$). For all predicates P and individuals a, b :

- (44) a. $[a]_{\approx_P^t} \gg_P^t [b]_{\approx_P^t}$ iff for all $x \in [a]_{\approx_P^t}$ and all $y \in [b]_{\approx_P^t}$, $x >_P^t y$.
b. $[a]_{\approx_P^s} \gg_P^s [b]_{\approx_P^s}$ iff for all $x \in [a]_{\approx_P^s}$ and all $y \in [b]_{\approx_P^s}$, $x >_P^s y$.

Similarly, consider the model in the proof of theorem 4.6.2: in this model, $a >_Q^s b >_Q^s c$, despite the fact that $b \not>_Q c$, and $\{a\} \gg_Q^s \{b\} \gg_Q^s \{c\}$ while $\{a, b\} \gg_Q \{c\}$.

In the proposal presented here, in addition to a semantic denotation, 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 two non-trivial pragmatic scales. However, the possibility that a particular absolute predicate may be associated with both a tolerant and a strict scale is actually ruled out by the total and partial axioms proposed above. In particular, we can prove that if Q is a total AA, then it is necessarily associated with a trivial strict scale, and if Q is a partial AA, then it is necessarily associated with a trivial tolerant scale.

Theorem 4.6.15 *If $Q \in AA^t$, then there is no CC t-model M such that, for distinct $x, y, z \in D$, $x >_Q^s y >_Q^s z$.*

Proof Let $Q \in AA^t$ and suppose for a contradiction that there is some model M such that there are $x, y, z \in D$ such that $x >_Q^s y >_Q^s z$. Since $y >_Q^s z$, there is some $X \in CC$ such that $y \in \llbracket Q \rrbracket_X^s$ and $z \notin \llbracket Q \rrbracket_X^s$. Since $y \in \llbracket Q \rrbracket_X^s$, $y \in \llbracket Q \rrbracket_X$. Since $x >_Q^s y$, there is some $X' \in CC$ such that $y \notin \llbracket Q \rrbracket_{X'}^s$. Since $y \in \llbracket Q \rrbracket_X$, by the AAA, $y \in \llbracket Q \rrbracket_{X'}$. Therefore there is some $d \in X'$ such that $d' \notin \llbracket Q \rrbracket_{X'}$ and $d' \sim_Q^{X'} y$. However, by the total axiom, $d' \not>_Q^{X'} y$. \perp So there is no CC t-model M such that, for distinct $x, y, z \in D$, $x >_Q^s y >_Q^s z$. \square

We can also prove that, if Q is a partial AA, then its tolerant scale is trivial.

Theorem 4.6.16 *If $Q \in AA^P$, then there is no CC t-model M such that, for distinct $x, y, z \in D$, $x >_Q^t y >_Q^t z$.*

Proof Let $Q \in AA^P$ and suppose for a contradiction that there is some model M such that there are $x, y, z \in D$ such that $x >_Q^t y >_Q^t z$. Since $x >_Q^t y$, there is some $X \in CC$ such that $x \in \llbracket Q \rrbracket_X^t$ and $y \notin \llbracket Q \rrbracket_X^t$. Since $y \notin \llbracket Q \rrbracket_X^t$, $y \notin \llbracket Q \rrbracket_X$. Since $y >_Q^s z$, there is some $X' \in CC$ such that $y \in \llbracket Q \rrbracket_{X'}^t$. Since $y \notin \llbracket Q \rrbracket_X$, by the AAA, $y \notin \llbracket Q \rrbracket_{X'}$. Therefore there is some $d \in X'$ such that $d' \in \llbracket Q \rrbracket_{X'}$ and $d' \sim_Q^{X'} y$. However, by the partial axiom, $d' \not\sim_Q^{X'} y$. \perp So there is no CC t-model M such that, for distinct $x, y, z \in D$, $x >_Q^t y >_Q^t z$. \square

In summary, the analysis presented in this chapter additionally provides us with some results concerning the scalarity of absolute predicates. We can illustrate the predictions about the association of different types of predicates with non-trivial strict weak orders (SWOs) as in table 4.8. Note that the blank cells will be filled in by the end of the chapter.

Adjective	$>_p$: non-trivial SWO?	$>_p^t$: non-trivial SWO?	$>_p^s$: non-trivial SWO?
Relative	✓		
Total Absolute	×	✓	×
Partial Absolute	×	×	✓
Non-Scalar	×		

Table 4.8: Scalarity Patterns with Absolute Adjectives

Observe furthermore that we now have a solution to the paradox of absolute adjectives; that is, we now have an explanation for how an adjective like *empty* can be, at the same time, gradable and not gradable and context-sensitive and not context-sensitive. Thus, my new solution (which is in the spirit of proposals made by Sapir and Récanati) is stated as in (45).

(45) **Solution to the Paradox of Absolute Adjectives:**

Although absolute adjectives have neither context-sensitive nor gradable semantic denotations, they have both context-sensitive and gradable pragmatic denotations.

In the next section, I show that the analysis of the pragmatics of AAs that I've given so far derives the potential vagueness patterns that we saw in section 4.4.

4.6.2.4 Potential Vagueness Results

Within the framework developed in this chapter, we can now have a formal definition of the *potentially vague* property that was introduced in this chapter.

Definition 4.6.8 Potentially vague adjective (formal) An adjective P is potentially vague just in case there is some CC t-model M such that there is some $X \in CC$ such that,

1. **Clear Case:** There is some $a_1 \in X$ such that $a_1 \in \llbracket P \rrbracket_X^s$.
2. **Clear Non-Case:** There is some $a_n \in \llbracket \text{not } P \rrbracket_X^s$.
3. **Sorites Series:** There are $a_1 \dots a_n \in X$ such that $a_1 \sim_P^X a_2$, and $a_2 \sim_P^X a_3 \dots a_{n-2} \sim_P^X a_{n-1}$, and $a_{n-1} \sim_P^X a_n$.

Recall that AAs are asymmetrically potentially vague: Total adjectives have potentially vague positive forms and not potentially vague negative forms; while partial adjectives display the reverse pattern. These patterns are direct consequences of the analysis presented so far.

Theorem 4.6.17 Vagueness of total absolute adjectives. If $Q \in AA^T$, then Q is potentially vague.

Proof Consider the CC t-model M such that $D = \{a, b, c, d, e\}$. Consider $X \in CC$ such that $X = \{a, b, c, d, e\}$. Suppose $\llbracket Q \rrbracket_X = \{a\}$. Suppose $\sim_Q^X = \{\langle a, b \rangle, \langle b, c \rangle, \langle c, d \rangle\}$ + reflexivity and symmetry for every pair except $\langle a, b \rangle$. Suppose furthermore that $a >_Q^t b >_Q^t c >_Q^t d >_Q^t e$ in M . Therefore,

1. **Clear Case:** $a \in \llbracket Q \rrbracket_X^s$.
2. **Clear Non-Case:** $d \in \llbracket \text{not } Q \rrbracket_X^s$.
3. **Sorites Series:** The sequence $\langle a, b, c, d, e \rangle$.

Therefore, Q is potentially vague. □

Note that the element e , that is both at the bottom of the tolerant scale associated with Q and indifferent from both c and d is required to satisfy the axiom *contrast preservation* (CP).

We can also show that partial negated adjectives are potentially vague.

Theorem 4.6.18 *Vagueness of the negation of partial absolute adjectives.* If $Q \in AA^P$, then ‘not Q ’ is potentially vague.

Proof Consider the CC t-model M such that $D = \{a, b, c, d, e\}$. Consider $X \in CC$ such that $X = \{a, b, c, d, e\}$. Suppose $\llbracket Q \rrbracket_X = \{a, b, c, d\}$. Suppose $\sim_Q^X = \{\langle b, c \rangle, \langle c, d \rangle, \langle d, e \rangle\}$ + reflexivity and symmetry for every pair except $\langle d, e \rangle$. Suppose furthermore that $a >_Q^s b >_Q^s c >_Q^s d >_Q^s e$ in M . Therefore,

1. **Clear Case:** $e \in \llbracket \text{not } Q \rrbracket_X^s$.
2. **Clear Non-Case:** $b \in \llbracket Q \rrbracket_X^s$.
3. **Sorites Series:** The sequence $\langle e, d, c, b, a \rangle$.

Therefore, not Q is potentially vague. □

Finally, we can prove that both the negations of total AAs and the positive forms of partial AAs are not potentially vague.

Theorem 4.6.19 *Precision of total ‘not Q ’.* If $Q \in AA^T$, then ‘not Q ’ is not potentially vague.

Proof Suppose ‘not Q ’ is potentially vague. Then there is some model and some comparison class X and some sequence $a_1 >_Q^t a_2 >_Q^t \dots a_n$ such that $a_1 \in \llbracket Q \rrbracket_X^s$, $a_n \in \llbracket \text{not } Q \rrbracket_X^s$, and $a_n \sim_Q^X a_{n-1} \dots a_2 \sim_Q^X a_1$. Since $a_1 \in \llbracket Q \rrbracket_X^s$, $a_1 \in \llbracket Q \rrbracket_X$. Since $a_n \in \llbracket \text{not } Q \rrbracket_X^s$, $a_n \notin \llbracket Q \rrbracket_X$. Finally, since $a_n \dots a_1$ form a Soritical series, there are some $a_i, a_{i+1} : a_1 \geq_Q^t a_i >_Q^t a_{i+1} \geq_Q^t a_n$, and $a_i \in \llbracket Q \rrbracket_X$ and $a_{i+1} \notin \llbracket Q \rrbracket_X$. Furthermore, $a_{i+1} \sim_Q^X a_i$. But, by the total axiom, $a_{i+1} \not\sim_Q^X a_i$. \perp So not Q is not potentially vague. □

Theorem 4.6.20 *Precision of Q .* If $Q \in AA^P$, then Q is not potentially vague.

Proof Suppose for a contradiction that Q is potentially vague. Then there is some model M and some $X \in CC_M$ such that there are some $a_1, a_2 \dots a_n : a_1 >_Q^s a_2 \dots a_{n-1} >_Q^s a_n$. Furthermore, $a_1 \in \llbracket Q \rrbracket_X^s$ and $a_n \in \llbracket \text{not } Q \rrbracket_X^s$, and $a_1 \sim_Q^X a_2 \dots a_{n-1} \sim_Q^X a_n$. Finally, since $a_1 \dots a_n$ form a Soritical series, there are some $a_i, a_{i+1} : a_1 \geq_Q^s a_i >_Q^s a_{i+1} \geq_Q^s a_n$, and $a_i \in \llbracket Q \rrbracket_X$ and $a_{i+1} \notin \llbracket Q \rrbracket_X$. Furthermore, $a_i \sim_Q^X a_{i+1}$. But, by the partial adjective axiom, $a_i \not\sim_Q^X a_{i+1}$. \perp So Q is not potentially vague. □

I therefore conclude that we can give an appropriate analysis of the asymmetric potential vagueness pattern within the framework developed in this dissertation.

Pattern	Relative	Total	Partial	Non-Scalar
P. vague \negP	✓	×	✓	×
P. vague P	✓	✓	×	×

Table 4.9: Potential Vagueness Patterns (AAs)

4.6.2.5 Other Consequences of the Proposal

The proposals made in this chapter make certain empirical predictions that, I argue, are borne out in the data associated with absolute comparatives. Comparatives and the properties of the scales associated with AAs will be the major focus of part 3 of the dissertation; however, we can observe that both total and partial comparatives display certain patterns associated with the semantic property of *evaluativity*¹⁵, which, I argue, follow naturally from the analysis given in this chapter.

Firstly, it is often observed that comparatives formed from most relative adjectives make no claims about whether the greater of the two individuals satisfy the positive form of the predicate to a high degree¹⁶.

- (46)
- 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.

¹⁵Rett (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 involve 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)

¹⁶Exceptions 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 degree semantics, see Rett (2008).

- (i)
 - a. Mary is more beautiful than Sue. (\Rightarrow Mary is beautiful.)
 - b. Mary is more brilliant than Sue. (\Rightarrow Mary is brilliant.)

However, comparatives formed from total absolute adjectives 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 (47).

- (47)
- a. John is **balder** than Peter.
If both people have full heads of hair, even if John actually has a smaller number of hairs than Peter.
 - b. 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.
 - c. This twisty staircase is **straighter** than this other twisty staircase.
If they are both really twisty, even if one has one fewer twist than the other.

The 'evaluativity' of total comparatives is straightforwardly predicted by the theory. By the definition of the tolerant comparative (def. 4.6.4) a will only be tolerantly greater than b if there is some context and some comparison class 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 comparison class in which a is indifferent from some individual at the top endpoint of the scale, then for all $b \in D$, $\langle a, b \rangle \notin >^t_Q$.

A second set of data involving evaluativity concerns similar effects with partial adjectives. 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 partial adjective. That is, the inferences in (48) hold:

- (48)
- a. This shirt is dirtier than that shirt \Rightarrow This shirt is dirty.
 - b. This towel is wetter than that towel \Rightarrow This towel is wet.
 - c. John is sicker than Mary \Rightarrow John is sick.
 - d. This stick is more bent than that stick \Rightarrow This stick is bent.

Again, these data are predicted by the theory. As shown by theorem 4.6.16, the only non-trivial

scale that can be associated with a partial adjective is the strict one. By the definition of the tolerant/strict comparative (def. 4.6.4), subjects of strict comparatives must be in the strict denotation of the predicate in some comparison class. Since strict denotations are subsets of semantic denotations, subjects of strict comparatives must be in the semantic denotation of the predicate in that comparison class. Since the semantic denotations of AAs are invariant across comparison classes, if a comparative with a partial AA Q is true under its strict interpretation, then its subject must always be in Q 's semantic denotation. Thus, the inferences in (48) 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 Sassoon and Toledo (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 total AA. For example, Kennedy and McNally (2005) (following Unger (1975)) claim that (49a) is necessarily false; whereas, Rotstein and Winter (2004) say that the almost identical example in (49b) can be true.

- (49) a. # The red towel is cleaner than the blue one, but both are clean.
 (Kennedy and McNally (2005))
- b. Both towels are clean, but the red one is cleaner than the blue one.
 (Rotstein and Winter (2004))

The approach developed so far has a straightforward account of why the examples in (49a) and (49b) could alternatively be viewed as both contradictions and non-contradictions: The only non-trivial comparative relation for a total AA is the tolerant comparative ($>_{clean}^t$). 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 comparison class is (i.e. $blue \notin \llbracket clean \rrbracket_X$, for all $X \in CC$). 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. $blue \in \llbracket clean \rrbracket_X^t$,

for some contextually given X)¹⁷. 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)), (49a)/(49b) 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 (49a) and (49b) 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 partial and total adjectives in comparatives.

4.6.2.6 Summary

In this section, I have presented an analysis of the asymmetric potential vagueness of absolute adjectives. I proposed a simple set of constraints that characterize the distribution of indifference relations across comparison classes with these predicates, and I showed that from the analysis of the vagueness patterns seen in this chapter, we arrive at both an analysis of the type 2 context-sensitivity patterns discussed in chapter 2 and a new solution to the puzzle of the gradability of absolute adjectives.

In the next section, I extend my analysis of the pragmatic denotations of AAs to the pragmatic denotations of relative adjectives.

4.6.3 Relative Adjectives

Since the axioms that I proposed in the previous section are meant to characterize the behaviour of indifference relations across comparison classes, and, in most cases, I believe that it is reasonable to think that relative and absolute predicates do not differ in these respects, I propose that the axioms *Tolerant/Strict No Skipping*, *Granularity*, *Minimal Difference*, and *Contrast Preservation* also apply to indifference relations associated with relative predicates. For example, there is nothing about a constraint like *Contrast Preservation* (CP) (repeated as (50) below) that would make it specific to absolute adjectives.

¹⁷See also Sassoon and Toledo (2011) for an account of the contrast in (49a) and (49b) in similar terms.

(50) **Contrast Preservation (CP).** For all $P \in SA \cup NS$ all $X, X' \in CC$, if $X \subset X'$ and, for $x, y \in X$, $x \not\sim_P^X y$ and $x \sim_P^{X'} y$, then $\exists z \in X' - X : x \not\sim_P^{X'} z$.

On the other hand, the total and partial axioms (repeated as in (51) and (52)) are different from the more general ‘coherence’ constraints on \sim_{PS} : in particular, I proposed that they make a distinction between total and partial AAs. Furthermore, I proposed that they were motivated by the special ‘prototypical’ status of the members of the context-independent denotations of AAs.

(51) **Total Axiom:**

If Q is a total adjective ($Q \in AA^T$), then, for all $x, y \in D$, if $x \in \llbracket Q \rrbracket_D$ and $y \notin \llbracket Q \rrbracket_D$, then $y \not\sim_Q^X x$, for all $X \in CC$.

•Otherwise, if $y \sim_Q^X x$, then $x \sim_Q^X y$.

(52) **Partial Axiom:**

If Q is a partial adjective ($Q \in AA^P$), then, for all $x, y \in D$, if $x \in \llbracket Q \rrbracket_D$ and $y \notin \llbracket Q \rrbracket_D$, then $x \not\sim_Q^X y$, for all $X \in CC$.

•Otherwise, if $x \sim_Q^X y$, then $y \sim_Q^X x$.

Relative adjectives, however, do not have such prototypical members, presumably because their semantic denotations are so highly context-sensitive¹⁸. Therefore, I propose that indifference relations associated with relative predicates are not subject to the limited asymmetry induced by the total and partial axioms. Instead, in line with traditional views of tolerance relations, I propose that \sim_{PS} are always symmetric.

(53) **Symmetry (S):** For all $P \in RA$, all $X \in CC$, and $x, y \in X$, if $x \sim_P^X y$, then $y \sim_P^X x$.

Therefore, the analysis of the pragmatic denotations of RAs (compared with AAs) is given as in table 4.10.

¹⁸The 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.

Axiom	Relative	Total AA	Partial AA
Reflexivity (R)	✓	✓	✓
Tolerant No Skipping (T-NS)	✓	✓	✓
Strict No Skipping (S-NS)	✓	✓	✓
Granularity (G)	✓	✓	✓
Minimal Difference (MD)	✓	✓	✓
Contrast Preservation (CP)	✓	✓	✓
Symmetry (S)	✓	×	×
Total Axiom (TA)	×	✓	×
Partial Axiom (PA)	×	×	✓

Table 4.10: Pragmatic Axioms for Scalar Adjectives

4.6.3.1 Context-Sensitivity and Vagueness Results

This analysis makes predictions about the context-sensitivity and potential vagueness properties of relative adjectives. Firstly, we can show that RAs are correctly predicated to be both type 1 and type 2 context-sensitive (theorem 4.6.22). In particular, this is because, based on the definitions of tolerant and strict denotations, being type 1 context-sensitive implies being type 2 context-sensitive (lemma 4.6.21).

Lemma 4.6.21 *For all $P \in SA \cup NS$, if P is type 1 context-sensitive, then P is type 2 context-sensitive.*

Proof Suppose P is type 1 context-sensitive. Then there is some CC t-model $M = \langle D, CC, \sim, \llbracket \cdot \rrbracket \rangle$ such that, for some $x \in D$, there is some $X \in CC$ such that $x \in \llbracket P \rrbracket_X$ and there is some distinct $X' \in CC$ such that $x \notin \llbracket P \rrbracket_{X'}$. Suppose that, for all P and all $X'' \in CC$, $\sim (\langle P, X'' \rangle)$ is the identity relation on X'' . So, by the definition of $\llbracket \cdot \rrbracket^t$, for all $X'' \in CC$, $\llbracket P \rrbracket_{X''}^t = \llbracket P \rrbracket_{X''}$. So $x \in \llbracket P \rrbracket_X^t$ and $x \notin \llbracket P \rrbracket_{X'}^t$. So P is type 2 context-sensitive, since its tolerant denotation is context sensitive. (Furthermore, by the definition of $\llbracket \cdot \rrbracket^s$, $x \in \llbracket P \rrbracket_X^s$ and $x \notin \llbracket P \rrbracket_{X'}^s$, so P 's strict denotation is also context-sensitive. \square)

Theorem 4.6.22 *If $P \in RA$, P is type 2 context-sensitive.*

Proof Immediately from the type 1 context-sensitivity of P (theorem 2.4.1) and lemma 4.6.21. \square

Furthermore, with this definition and our theory of the semantics and pragmatics of relative adjectives, we can show that both positive and negative forms of RAs are correctly predicted to have this property. In other words, we predict the symmetric potential vagueness of relative adjectives.

Theorem 4.6.23 *Potential vagueness of relative adjectives.* If $P \in RA$, P is potentially vague.

Proof Consider the CC t-model M such that $D = \{a, b, c, d\}$. Consider $X \in CC$ such that $X = \{a, b, c, d\}$. Suppose $\llbracket P \rrbracket_X = \{a, b\}$. Suppose $\sim_P^X = \{\langle a, b \rangle, \langle b, c \rangle, \langle c, d \rangle\}$ + reflexivity and symmetry. Suppose furthermore that $a >_P b >_P c >_P d$ in M . Therefore,

1. **Clear Case:** $a \in \llbracket P \rrbracket_X^s$.
2. **Clear Non-Case:** $d \in \llbracket \text{not } P \rrbracket_X^s$.
3. **Sorites Series:** The sequence $\langle a, b, c, d \rangle$.

Therefore, P is potentially vague. □

Theorem 4.6.24 *If $P \in RA$, then ‘not P ’ is potentially vague.*

Proof Consider the model in the proof of theorem 4.6.23. □

Thus, we predict both the CS and p.vagueness patterns discussed in the text so far.

Pattern	Relative	Total	Partial	Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Pattern	Relative	Total	Partial	Non-Scalar
P. vague ¬P	✓	×	✓	×
P. vague P	✓	✓	×	×

Finally, I end the examination of p.vagueness and context-sensitivity with RAs with a short discussion of the predictions that my analysis makes with respect to the possibility of having precise interpretations of both RAs and AAs, as discussed in sections 4.2.2 and 4.3. Clearly

the simple fact that tolerant and strict denotations are calculated on the basis of semantic denotations and contextually given indifference relations does not automatically mean that for every predicate P in every situation, P will be vague. By assumption, indifference relations are reflexive (you're always indifferent from yourself), so all the members of a predicate's classical extension will always be in its tolerant extension. Although, in a given context, it is possible that $\llbracket P \rrbracket \subset \llbracket P \rrbracket^t$, it is not necessary. In contexts in which it is important to be very precise, that is, in contexts in which differences between individuals that are close to each other with respect to $>_P$ have an import for our purposes, \sim_P^X might not relate individuals that are on either side of the borderline of $\llbracket P/\text{not } P \rrbracket$. In the models and comparison classes in which \sim_P^X respects the border of $\llbracket P \rrbracket$, the predicate 'sharpens up'. Thus, in a situation in which people with a single hair are perceived as relevantly different from those with no hair (like in the Yul Brynner example), \sim_{bald}^X will not relate completely bald individuals with individuals that have any hair, and $\llbracket bald \rrbracket_X^s = \llbracket bald \rrbracket_X = \llbracket bald \rrbracket_X^t$. Likewise, if a single cent makes a significant difference to how we view price of a penny candy, then candies costing one cent will not be indifferent from those costing two cents, and $\llbracket expensive \rrbracket_{penny\ candies} = \llbracket expensive \rrbracket_{penny\ candies}^t = \llbracket expensive \rrbracket_{penny\ candies}^s$.

4.6.3.2 Relative Pragmatic Scales

Finally, in this section, I examine what implications the analysis given above has for the pragmatic scales associated with relative adjectives. However, so that we are sure that such an analysis is needed, let us first consider what would happen if we did not put any constraints on the \sim_{PS} associated with relative adjectives. The scales constructed out of the semantic denotation of RAs form possibly non-trivial strict weak orders; therefore, we might wonder what consequences having a more complex structure at the semantic level might have for the scales constructed out of tolerant and strict denotations. One might think that we would be able to see some reflexes of the structure created by van Benthem's axioms for $>_P$ in the $>_P^t$ or $>_P^s$ relations. However, surprisingly¹⁹, despite the fact that they hold at the semantic level, none of van Benthem's axioms are theorems at the tolerant level. Consider, for example, van Benthem's simplest axiom: *No Reversal*, stated in its

¹⁹In fact, this is a difference between Klein-ian TCS and simple TCS, since classical validities imply tolerant validities ($\models^c \Rightarrow \models^t$) in the latter system (cf. Cobreros et al. (2011a)'s Corollary 1).

classical/semantic and tolerant forms in (54).

(54) **No Reversal:**

a. **Semantic (NR):**

For $x \in \llbracket P \rrbracket_X$ and $y \in \llbracket P \rrbracket_X$,

There is no $X' \in CC$ such that $y \in \llbracket P \rrbracket_{X'}$ and $x \notin \llbracket P \rrbracket_{X'}$.

b. **Tolerant (T-NR):**

For $x \in \llbracket P \rrbracket_X^t$ and $y \in \llbracket P \rrbracket_X^t$,

There is no $X' \in CC$ such that $y \in \llbracket P \rrbracket_{X'}^t$ and $x \notin \llbracket P \rrbracket_{X'}^t$.

Theorem 4.6.25 $NR \not\Rightarrow T-NR$. *It is not the case that, if P satisfies NR, then P satisfies T-NR.*

Proof Suppose $P \in RA$ (so P satisfies NR). Let $X \in CC$ such that $X = \{a, b, c\}$ and $\llbracket P \rrbracket_X = \{a\}$ and $\sim_P^X = \{\langle a, b \rangle, \langle b, a \rangle\} + \text{refl.}$ So $b \in \llbracket P \rrbracket_X^t$ and $c \notin \llbracket P \rrbracket_X^t$. Let $X' \in CC$ such that $X' = \{a, b, c, d\}$ and $\llbracket P \rrbracket_{X'} = \{a\}$ and $\sim_{P'}^{X'} = \{\langle a, c \rangle, \langle c, a \rangle\} + \text{refl.}$ So $b \notin \llbracket P \rrbracket_{X'}^t$ and $c \in \llbracket P \rrbracket_{X'}^t$. \square

So, without any axioms constraining what kind of indifference relations can be established in various CCs, even though NR holds at the semantic level, it does not tolerantly hold. So the basic ‘coherence’ constraints proposed above are necessary.

Now consider what happens in the analysis proposed in table 4.10. Firstly, we can observe that, if we adopt Tolerant No Skipping, it allows us to prove the tolerant version of van Benthem’s No Reversal, just like it did with absolute adjectives.

Theorem 4.6.26 *No Tolerant Reversal (T-NR):* *For $P \in RA$ and $X \in CC$, if $x \in \llbracket P \rrbracket_X^t$ and $y \notin \llbracket P \rrbracket_X^t$, then there is no $X' \in CC$ such that $y \in \llbracket P \rrbracket_{X'}^t$ and $x \notin \llbracket P \rrbracket_{X'}^t$.*

Proof Suppose $x \in \llbracket P \rrbracket_X^t$ and $y \notin \llbracket P \rrbracket_X^t$. Suppose, for a contradiction that there is an $X' \in CC$ such that $y \in \llbracket P \rrbracket_{X'}^t$ and $x \notin \llbracket P \rrbracket_{X'}^t$. Therefore, $x >_P^t y$ and $y >_{P'}^t x$. Furthermore, by assumption and the definition of $\llbracket P \rrbracket_X^t$, there is some $d \sim_P^X x$ such that $d \in \llbracket P \rrbracket_X$, and $d \not\sim_{P'}^{X'} y$. Thus $d >_P^t y$ and so $d >_{P'}^t y >_P^t x$. Since $d \sim_P^X x$, by Tolerant No Skipping, $d \sim_{P'}^{X'} y$. \perp \square

We saw in the section about AAs that the axioms in table 4.10 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 $>_P^t$ relations that are not even transitive. This fact is stated below and the counter-model is presented in the appendix.

(55) **Theorem 4.8.1:**

For $P \in RA$, $>_P^t$ is not necessarily transitive.

Crucially, however, the intransitivity of the $>_P^t$ s is extremely limited. In particular, transitivity does not go through only for cases where three individuals are all equivalent with respect to $>_P^{20}$, but there are some very restricted situations in which they all look to have different heights. We can state this observation as theorem 4.6.27.

Theorem 4.6.27 *Let $P \in RA$ and let M be a CC t-model such that $x, y, z \in D$. Suppose $x >_P^t y$ and $y >_P^t z$, but $x \not>_P^t z$. Then $x \approx_P y \approx_P z$.*

Proof Suppose $P \in RA$ and let M be a CC t-model such that $x, y, z \in D$. Suppose $x >_P^t y$ and $y >_P^t z$, but $x \not>_P^t z$.

1. Show $x \approx_P z$. Suppose for a contradiction that $x \not\approx_P z$. By theorem 4.6.3, if $x >_P z$, then $x >_P^t z$, but by assumption $x \not>_P^t z$. So $z >_P x$. By theorem 2.4.9, $z \in \llbracket P \rrbracket_{\{x,z\}}$ and $x \notin \llbracket P \rrbracket_{\{x,z\}}$. Now consider $\{x, y, z\} \in CC$. **Case 1:** $y \in \llbracket P \rrbracket_{\{x,y,z\}}$. Since, by assumption, $x >_P^t y$, by theorems 4.6.3 and 4.6.5, $x \in \llbracket P \rrbracket_{\{x,y,z\}}$. By Upward Difference, $z \notin \llbracket P \rrbracket_{\{x,y,z\}}$. So $x >_P z$ and by thm. 4.6.3, $x >_P^t z$. \perp So $x \approx_P z$. \checkmark **Case 2:** $y \notin \llbracket P \rrbracket_{\{x,y,z\}}$. Since $y >_P^t z$, by thm. 4.6.3 and thm. 4.6.5, $z \notin \llbracket P \rrbracket_{\{x,y,z\}}$. By Upward Difference, $x \in \llbracket P \rrbracket_{\{x,y,z\}}$. So $x >_P z$ and, by theorem 4.6.3, $x >_P^t z$. \perp So $x \approx_P z$. \checkmark
2. Show $x \approx_P y$. Suppose for a contradiction that $x \not\approx_P y$. Since $x >_P^t y$, $x >_P y$. So, by thm. 2.4.9, $x \in \llbracket P \rrbracket_{\{x,y\}}$ and $y \notin \llbracket P \rrbracket_{\{x,y\}}$. Now consider $\{x, y, z\} \in CC$. **Case 1:** $y \in \llbracket P \rrbracket_{\{x,y,z\}}$. Since $x >_P y$,

²⁰That is to say, when we are speaking precisely, we would consider them all as having the same height.

by No Reversal, $x \in \llbracket P \rrbracket_{\{x,y,z\}}$. Then, by Upward Difference, $z \notin \llbracket P \rrbracket_{\{x,y,z\}}$. So $x >_P z$ and, by theorem 4.6.3, $x >_P^t z$. \perp So $x \approx_P y$. \checkmark **Case 2:** $y \notin \llbracket P \rrbracket_{\{x,y,z\}}$. Since $y >_P^t z$, by theorems 4.6.3 and 4.6.5, $z \notin \llbracket P \rrbracket_{\{x,y,z\}}$. So by Upward Difference, $x \in \llbracket P \rrbracket_{\{x,y,z\}}$. So $x >_P z$ and, by theorem 4.6.3, $x >_P^t z$. \perp So $x \approx_P y$. \checkmark

□

As I see it, given this result, there are two possibilities. One is to introduce a new axiom that only relative \sim_P s are subject to: something that rules out cases where we would make more distinctions ‘speaking loosely’ than when we are ‘speaking precisely’. In a sense, such an axiom would be similar to Kennedy (2007)’s *Interpretative Economy* in that we would force the tolerant denotation of a relative predicate to respect the distinctions made by its semantic denotation²¹. However, another possibility is to simply allow the restricted intransitivity that the analysis predicts. Although, at the moment, I cannot think of concrete examples in which we would want to view individuals that we consider to have the same height to be distinct with respect to the property of being *tall*, we might later find arguments that such situations could arise. I therefore suggest that we allow the limited intransitivity with $>_P$ s in the analysis.

This proposal also 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. Thus, we can fill in more of table 4.8 as shown in table 4.11. In other words, my analysis predicts that RAs are uniquely associated with scales derived from their semantic denotations, total AAs are uniquely associated with scales derived from their tolerant denotations, and partial AAs are uniquely associated with scales derived from their strict denotations.

Adjective	$>_P$: non-trivial SWO?	$>_P^t$: non-trivial SWO?	$>_P^s$: non-trivial SWO?
Relative	✓	×	×
Total Absolute	×	✓	×
Partial Absolute	×	×	✓
Non-Scalar	×		

Table 4.11: Scalarity Patterns with Scalar Adjectives

²¹This approach was adopted by Burnett (2012c).

4.6.3.3 Are comparatives vague?

One of the main focuses of this work is the (potential) vagueness and precision of the positive form of adjectives; however, now we might wonder about the p.vagueness properties of comparatives (i.e. *John is taller than Mary*). Is the binary predicate *taller than* potentially vague?

When it comes to the vagueness of comparatives, the literature is divided. By far, the dominant view in linguistics is that comparatives are precise (see the discussion in Kennedy (2007) and van Rooij (2011a)). In part, this is because, when we compare sentences like (56a) with sentences like (56b), we have the intuition that, unlike for (56a), (56b) has sharp truth conditions: to know whether this sentence is true, all we have to do is compare Mary and John's heights and see which one is greater.

- (56) a. Mary is tall.
b. Mary is taller than John.

However, it has been argued (ex. by Kamp (1975), Williamson (1994), and Keefe (2000)) that comparatives can also be vague. From what I can see, the cases that are proposed to attest to the potential vagueness of comparatives have already been discussed in the section on multi-dimensionality in the first part of the dissertation. For example, given that, as I discussed in chapter 2, if Mary is clever at math and John is clever with people, we might want to say both sentences in (57) are true. Thus, in this situation, the pair ⟨John, Mary⟩ would be a borderline case of *more clever than*.

- (57) a. Mary is more clever than John.
b. Mary is not more clever than John.

Similarly, if the new World Trade Center has a greater height than the Willis Tower only if we measure its roof needle, then we might want to assent to both sentences in (58). So ⟨World Trade Center, Willis Tower⟩ would be a borderline case of *taller than*.

- (58) a. The World Trade Center is taller than the Willis Tower.

- b. The World Trade Center is not taller than the Willis Tower.

This being said, I believe that there are arguments against the claim that comparatives are potentially vague. These arguments rejoin the discussion of multi-dimensionality and the asymmetry of the comparative relation in chapter 2. In that chapter, I argued that comparatives like (57) and (58) are ambiguous in that, at the same time, they can be both true and false depending on the dimension that we pick. So, in this way of thinking, comparatives would not be potentially vague and $\langle \text{WTC, Willis Tower} \rangle$ would not be a true borderline case of the *taller than* relation.

We can distinguish the tolerance of contradictions with comparatives from contradictions with borderline cases of (potentially) vague predicates by comparing sentences like (59) and (60).

- (59) The Willis Tower is both taller than the WTC and not taller than the WTC.
 - a. The Willis Tower is taller than the WTC (because the main building is taller) and not taller than the WTC (because the WTC's needle makes it taller).
 - b. When we take needles into account:
 - # The Willis Tower is both taller than the WTC and not taller than the WTC.
- (60) Mary is both tall and not tall.
 - a. Mary is both tall (because she has huge hair) and not tall (because her body is short).
 - b. Mary is both tall and not tall (when we take into account hair length).

Acceptable contradictions with comparatives can be disambiguated into clearly true and clearly false conjuncts, and once the dimensions that are relevant to the application of the predicate are made explicit, the contradictions become unacceptable. Observe that this is a different pattern from the 'tolerant' contradictions with true borderline cases the positive form of the adjective. We can be very clear about how we are measuring height (from the top of the hair; in centimetres; with this ruler. . .), and, provided that Mary is still in the borderline region for a vague predicate, (60b) is still acceptable.

But even if I am wrong, and no principled difference exists between the contradiction in (59)

and the one in (60), it is undoubtedly possible to extend the type of account developed in this thesis for unary predicates like *tall* to binary predicates like *taller than*. However, such an extension is out of the scope of this work.

4.6.3.4 Summary

In summary, in this section, I extended my analysis of vagueness and imprecision with absolute adjectives to relative adjectives. I showed that this straightforward extension correctly predicts the potential vagueness patterns and the type 2 context-sensitivity patterns associated with these predicates, as well as certain observations about their scalarity. In the next section, I develop an analysis of the vagueness and CS patterns associated with non-scalar predicates within the current framework.

4.6.4 Non-Scalar Adjectives

In this section, I complete the analysis of vagueness and context-sensitivity patterns presented in parts one and two of the dissertation. First, I give an analysis of both the non-context-sensitivity and the precision of non-coerced non-scalar adjectives, and then I give an analysis of the properties of coerced NS.

As discussed in section 4.4, 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 force their pragmatic denotations to replicate exactly the distinctions made by their semantic denotations. This constraint is stated *Be precise* as in (61).

(61) **Be Precise (BP):**

For all $S \in NS$, all $X \in CC$, and all $x, y \in X$, if $x \sim_S^X y$ and there is some $z \in X$ such that $x \geq_S z \geq_S y$, then $x \sim_S^X z$ and $z \sim_S^X y$.

Thus, my analysis of the pragmatic constraints associated with non-coerced non-scalar adjectives is given (alongside the constraints associated with other kinds of adjectival predicates) in

table 4.12.

Axiom	Relative	Total AA	Partial AA	Non-Scalar
Reflexivity (R)	✓	✓	✓	✓
Tolerant No Skipping (T-NS)	✓	✓	✓	✓
Strict No Skipping (S-NS)	✓	✓	✓	✓
Granularity (G)	✓	✓	✓	✓
Minimal Difference (MD)	✓	✓	✓	✓
Contrast Preservation (CP)	✓	✓	✓	✓
Symmetry (S)	✓	×	×	✓
Total Axiom (TA)	×	✓	×	×
Partial Axiom (PA)	×	×	✓	×
Be Precise (BP)	×	×	×	✓

Table 4.12: Pragmatic Axioms for (Non)Scalar Adjectives

4.6.4.1 Vagueness/Context-Sensitivity/Scalarity Results

Firstly, we can show that this analysis correctly predicts that true NSs are not potentially vague.

Theorem 4.6.28 *If $S \in NS$, then S is not potentially vague.*

Proof Suppose $S \in NS$ and suppose for a contradiction that S is potentially vague. So there is some CC t-model M such that there is some $X \in CC$ and some some series $a_1 \dots a_n \in X$ such that $a_1 \in \llbracket S \rrbracket_X^s$ and $a_n \in \llbracket \text{not } S \rrbracket_X^s$. Furthermore, there are a_i and a_{i+1} such that $a_i \in \llbracket S \rrbracket_X$ and $a_{i+1} \notin \llbracket S \rrbracket_X$, and $a_i \sim_S^X a_{i+1}$. Since $a_i \in \llbracket S \rrbracket_X$ and $a_{i+1} \notin \llbracket S \rrbracket_X$, $a_i >_S a_{i+1}$. Furthermore, since $a_n \in \llbracket \text{not } S \rrbracket_X^s$, $a_n \notin \llbracket S \rrbracket_X$, so $a_i >_S a_n$. Finally, since $a_{i+1}, a_n \notin \llbracket S \rrbracket_X$, since S satisfies the AAA, $a_n \approx_S a_{i+1}$ (cf. theorem 2.4.8). So $a_i \geq_S a_n \geq_S a_{i+1}$. Since $a_i \sim_S^X a_{i+1}$, by Be Precise, $a_1 \sim_S^X a_n$. So, by the definition of $\llbracket \cdot \rrbracket^s$, $a_n \notin \llbracket S \rrbracket_X^s$. \perp So S is not potentially vague. \square

The analysis also gives us some results about the scalarity of non-scalar adjectives: in particular, we predict that (non-coerced) NSs also have necessarily trivial scales. First, we can show that, thanks to the *Be Precise* axiom, NSs' pragmatic scales are the same as their semantic scales.

Lemma 4.6.29 *If $S \in NS$, then $>_S^t = >_S = >_S^s$.*

Proof Let $S \in NS$,

1. Show $>^t_S = >_S$. \subseteq Let $x >^t_S y$ to show $x >_S y$. Suppose $x \not>_S y$. By theorems 4.6.5 and 4.6.3, $x \approx_S y$. Since $x >^t_S y$, there is some $X \in CC$ such that $x \in \llbracket S \rrbracket'_X$ and $y \notin \llbracket S \rrbracket'_X$. So there is some $d \in \llbracket S \rrbracket_X$ such that $d \sim^X_S x$ but $d \not\sim^X_S y$. Furthermore, since $x \approx_S y$, $x \notin \llbracket S \rrbracket_X$. So $d >_S y \geq_S x$. Since $d \sim^X_S x$, by Be Precise, $d \sim^X_S y$. \perp So $x >_S y$. $\checkmark \supseteq$ Immediately from theorem 4.6.3. \checkmark
2. Show $>_S = >^s_S$. \subseteq Immediately from theorem 4.6.4. $\checkmark \supseteq$ Let $x >^s_S y$ to show $x >_S y$. Suppose for a contradiction that $x \not>_S y$. By theorems 4.6.6 and 4.6.4, $x \approx_S y$. Since $x >^s_S y$, there is some $X \in CC$ such that $x \in \llbracket S \rrbracket^s_X$ and $y \notin \llbracket S \rrbracket^s_X$. By the definition of $\llbracket \cdot \rrbracket^s$ and since $x \approx_S y$, $x, y \in \llbracket S \rrbracket_X$. Since $y \notin \llbracket S \rrbracket^s_X$, there is some $d \sim^X_S y$ such that $d \notin \llbracket S \rrbracket_X$. Furthermore, by the definition of $\llbracket \cdot \rrbracket^s$, $x \not\sim^X_S d$. So $y \geq_S x >_S d$. Since $d \sim^X_S y$, by BP, $x \sim^X_S d$. \perp So $x >_S y$. \checkmark

□

Now, since non-scalar adjectives have trivial semantic scales, they also have trivial pragmatic scales.

Theorem 4.6.30 *If $S \in NS$, then,*

1. *There are no CC t-models M such that, for distinct $x, y, z \in D$, $x >^t_S y >^t_S z$.*
2. *There are no CC t-models M such that, for distinct $x, y, z \in D$, $x >^s_S y >^s_S z$.*

Proof Immediately from theorem 2.4.8 (triviality of $>_S$) and lemma 4.6.29. □

Finally, we can show that non-scalar adjectives are not type 2 context-sensitive.

Theorem 4.6.31 *If $S \in NS$, then S is not type 2 context-sensitive.*

Proof T-context-sensitive. Suppose $S \in NS$. Then S satisfies the AAA. Suppose for a contradiction that S is t-context sensitive. Then there is some model M such that there is some $x \in D$ and there are some $X, X' \in CC$ such that $x \in \llbracket S \rrbracket'_X$ and $x \notin \llbracket S \rrbracket'_{X'}$. Since $x \notin \llbracket S \rrbracket'_{X'}$, by the AAA, $x \notin \llbracket S \rrbracket_D$. Since $x \in \llbracket S \rrbracket'_X$, there is some $d \sim^X_S x$ such that $d \in \llbracket S \rrbracket_X$. By the AAA, $d \in \llbracket S \rrbracket_D$. Therefore $d >_S x$, so $d \in \llbracket S \rrbracket_{d,x}$ and $x \notin \llbracket S \rrbracket_{d,x}$. Furthermore, by MD, $d \not\sim^{d,x}_S x$. Additionally, since $\{x, d\} \subset X$ and

$d \sim_S^X x$, by Granularity, there is some $z \in X - \{x, d\}$ such that $d \not\sim_S^X z$ and $x \not\sim_S^X z$. **Case 1:** $z \in \llbracket S \rrbracket_X$. Then $d \geq_S z >_S x$. Since $d \sim_S^X x$, by Be Precise, $d \sim_S^X z$. \perp **Case 2:** $z \notin \llbracket S \rrbracket_X$. Then $d >_S z \geq_S x$. Since $d \sim_S^X x$, by Be Precise, $d \sim_S^X z$. \perp So S is not t-context-sensitive. \checkmark .

S-context-sensitive. By parallel reasoning and the duality of $\llbracket \cdot \rrbracket^t$ and $\llbracket \cdot \rrbracket^s$. \checkmark □

In summary, the analysis presented so far in this dissertation derives the context-sensitivity, potential vagueness, and scalarity patterns that (I argued) we observe in the adjectival domain. A summary of the empirical predictions of the model is given in the following tables:

Predicted Context-Sensitivity, Vagueness, and Scalarity Patterns

Pattern	Relative	Total	Partial	Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Table 4.13: Context-Sensitivity Patterns

Pattern	Relative	Total	Partial	Non-Scalar
P. vague ¬P	✓	×	✓	×
P. vague P	✓	✓	×	×

Table 4.14: Potential Vagueness Patterns

Adjective	$>_p$: non-trivial SWO?	$>'_p$: non-trivial SWO?	$>^s_p$: non-trivial SWO?
Relative	✓	×	×
Total Absolute	×	✓	×
Partial Absolute	×	×	✓
Non-Scalar	×	×	×

Table 4.15: Scalarity Patterns

4.6.5 Coerced Non-Scalar Adjectives

I proposed that pragmatic denotations of non-scalar adjectives 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 semantic analysis in chapter 2, I propose that what differentiates non-scalar adjectives from absolute scalar adjectives is in their pragmatics, not their semantics.

(62) **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 ‘coerced’ non-scalar adjectives:

(63) **Scalar Coercion:**

Coerced non-scalar adjectives are subject to all the same constraints as regular NSs, except *Be Precise*.

This approach is very different from certain other current views (to be discussed in chapter 7) that propose a semantic and even a syntactic difference between these predicates. In this section, I give three arguments in favour 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 →AA dependency*.

4.6.5.1 The Technical Nature of NSs

The first argument that the AA/NS distinction has to do with precision concerns the nature of the inventory of non-scalar adjectives. In particular, if we look at the (extended) inventory of NSs discussed throughout the dissertation (64), we can notice that the vast majority (if not all of them)

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*).

(64) **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 than that one* vs *✓This room is emptier than that one*), some readers may have different judgements about the non-scalar status of some of the words in (64). So they might not find the generalization concerning scalar/non-scalar lexicalization patterns so convincing. But this observation about variation in judgements of non-scalarity brings me to the second argument: the ease of coercion.

4.6.5.2 The Ease of Coercion

As I mentioned, a characteristic property of non-scalar adjectives 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 (64) sounds strange in the comparative out of context, it is perfectly natural to use many of them as follows:

(65) 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 *deader* than zombie B.
 - f. France is *more hexagonal* than Canada.
- and so on. . .

Even some of the more ‘mathematical’ terms in (64) can acquire a gradable 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 (66a) and (66b), as well as in the example (66c) from Rett (2012) (p.9), which is also inspired by the results of Armstrong et al. (1983).

- (66)
- 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 non-scalar adjectives that, with very little effort, they can appear in degree constructions and (as discussed in chapter 2) when they do so, they become context-sensitive and vague²². Of course, an analysis in which there was a scalar coercion process in the grammar (perhaps some sort of degree argument adding operation) could account for the ease of coercion (the coercion process could be highly productive). However, the pragmatic

²²In 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.

analysis that I have given needs no such morpho-semantic operation: non-scalar adjectives are simply absolute scalar adjectives that tend to be used with a higher level of precision. Alternatively, in my analysis, we could describe AAs as simply non-scalar adjectives that tend to be used loosely.

4.6.5.3 The $NS^c \rightarrow AA$ Dependency

My final argument in favour of an analysis in which non-scalar adjectives are semantically identical to absolute scalar adjectives comes from an empirical observation (already discussed in chapter 2) about the properties of coerced NSs. The generalization is the following:

(67) **The $NS^c \rightarrow AA$ Dependency:**

Non-scalar adjectives are coerced into absolute scalar adjectives.

For example, we saw in chapter 2 that, although they are type 2 context-sensitive, NS^c appear to uniformly fail the definite description test:

- (68)
- a. Pass me the hexagonal one.
(But both/neither are hexagonal!)
 - b. Show me the illegal one.
(But both/neither are illegal!)
 - c. Show me the dead one.
(But both/neither are dead!)

An analysis in which coercion is a morpho-semantic operation would have to build (67) into the operation, and this would raise the question of why we cannot coerce a NS into a relative scalar adjective. 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 ‘coerced’ NSs.

Finally, we can observe that the pragmatic loosening operation that NS 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 total AAs and partial AAs are models for coerced NSs). Thus, the results concerning the (non)scalarity of adjectival predicates are shown in table 4.16.

Adjective	$>_p$: non-trivial SWO?	$>_p^t$: non-trivial SWO?	$>_p^s$: non-trivial SWO?
Relative	✓	×	×
Total Absolute	×	✓	×
Partial Absolute	×	×	✓
Non-Scalar	×	×	×
Coerced Non-Scalar	×	✓	✓

Table 4.16: Scalarity Patterns

4.6.5.4 Summary

In summary, in this section, I presented a new analysis of the pragmatics of non-scalar adjectives like *hexagonal* and *prime*. I proposed that NSs differ from AAs in that they are subject to a pragmatic principle that ensures that they are applied precisely in all contexts. I showed that this analysis makes the correct predictions with respect to the vagueness, context-sensitivity, and scalarity patterns exhibited by this class of adjectives. Furthermore, I proposed a new analysis of the operation of ‘scalar coercion’ as a pragmatic loosening process through which the constraint *Be Precise* is lifted, and I argued that such an analysis can account for a series of empirical observations about the close relationship between ‘coerced’ non-scalars and absolute scalars.

4.7 Conclusion

In this chapter, I presented new data concerning the behaviour of adjectives of various scale structure classes in Soritical arguments. I argued that the scale structure classes that are the topic of this work can be entirely distinguished based on their potential vagueness patterns, and I gave an analysis of these patterns within a TCS extension of the delineation framework that I proposed in chapter 2. The bulk of the analysis consisted in proposing a series of constraints on the definition of

the \sim as it applies to adjectives of different classes. The constraints are summarized in table 4.12 (repeated as table 4.17 below).

Axiom	Relative	Total AA	Partial AA	Non-Scalar
Reflexivity (R)	✓	✓	✓	✓
Tolerant No Skipping (T-NS)	✓	✓	✓	✓
Strict No Skipping (S-NS)	✓	✓	✓	✓
Granularity (G)	✓	✓	✓	✓
Minimal Difference (MD)	✓	✓	✓	✓
Contrast Preservation (CP)	✓	✓	✓	✓
Symmetry (S)	✓	×	×	✓
Total Axiom (TA)	×	✓	×	×
Partial Axiom (PA)	×	×	✓	×
Be Precise (BP)	×	×	×	✓

Table 4.17: Pragmatic Axioms for (Non)Scalar Adjectives

As the table suggests, the constraints can be grouped into three distinct classes. The first class (*R*, *T-NS*, *S-NS*, *G*, *MD*, and *CP*) is meant to characterize cognitive indifference relations and their basic distribution in different comparison classes. 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_{PS} 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 this chapter and the first part of the dissertation. Furthermore, I showed that my analysis provides a solution to the puzzle of the gradability of absolute adjectives that was presented in chapter 2: the non-trivial scales that are associated with AAs are constructed from their pragmatic meanings not their 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 the next and final part

of the dissertation, I will show that the analysis developed in parts 1 and 2 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 the proposals made so far to account for 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.

4.8 Appendix: Transitivity Countermodel

Theorem 4.8.1 *For $P \in RA$, $>_P^t$ is not necessarily transitive.*

Show there is some model M such that $x >_P^t y$ and $y >_P^t z$, but there is no $X \in CC$ such that $x \in \llbracket P \rrbracket_X^t$ and $z \notin \llbracket P \rrbracket_X^t$.

Proof Let $D = \{d, d', x, y, z, a\}$. Let $x \approx_P y \approx_P z$. So x, y, z have to behave the same way w.r.t. classical denotations.

1. $\{d, x, y\} : \llbracket P \rrbracket_{\{d, x, y\}} = \{d\}$. And $\sim_P^{\{d, x, y\}}$ is empty (except for refl.). So $\llbracket P \rrbracket_{\{d, x, y\}}^t = \{d\}$.
2. $\{d, x, y, a\} : \llbracket P \rrbracket_{\{d, x, y, a\}} = \{d\}$. Now $d \sim_P^{\{d, x, y, a\}} x$. So $\llbracket P \rrbracket_{\{d, x, y, a\}}^t = \{d, x\}$.

$$(69) \quad \text{So } x >_P^t y.$$

3. $\{d', y, z\} : \llbracket P \rrbracket_{\{d', y, z\}} = \{d'\}$. And $\sim_P^{\{d', y, z\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d', y, z\}}^t = \{d'\}$.
4. $\{d', y, z, a\} : \llbracket P \rrbracket_{\{d', y, z, a\}} = \{d'\}$. And now $y \sim_P^{\{d', y, z, a\}} z$. So $\llbracket P \rrbracket_{\{d', y, z, a\}}^t = \{d', y\}$.

$$(70) \quad \text{So } y >_P^t z.$$

We now have to consider all the comparison classes (subsets of D) that contain both x and z to show that it is possible for x, z to not be a tolerant difference pair.

1. $\{x, z\} : \text{Since } x \approx_P z, \llbracket P \rrbracket_{\{x, z\}} = \{x, z\} \text{ and } \llbracket P \rrbracket_{\{x, z\}}^t = \{x, z\}$.

2. $\{x, y, z\} : \llbracket P \rrbracket_{\{x, y, z\}} = \{x, y, z\}$, so $\llbracket P \rrbracket_{\{x, y, z\}}^t$.
3. $\{x, z, a\} : \llbracket P \rrbracket_{\{x, z, a\}} = \{x, z\}$ and $\sim_P^{\{x, z, a\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{x, z, a\}}^t = \{x, z\}$.

$$(71) \quad x, z >_P^{c/t} a.$$

4. $\{d, x, z\} : \llbracket P \rrbracket_{\{d, x, z\}} = \{d\}$. And $\sim_P^{\{d, x, z\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d, x, z\}}^t = \{d\}$.
5. $\{d', x, z\} : \llbracket P \rrbracket_{\{d', x, z\}} = \{d'\}$. And $\sim_P^{\{d', x, z\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d', x, z\}}^t = \{d'\}$.
6. $\{d, d', x, z\} : \llbracket P \rrbracket_{\{d, d', x, z\}} = \{d, d'\}$. And $\sim_P^{\{d, d', x, z\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d, d', x, z\}}^t = \{d, d'\}$.
7. $\Rightarrow \{d, x, y, z\} : \llbracket P \rrbracket_{\{d, x, y, z\}} = \{d\}$ and $\sim_P^{\{d, x, y, z\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d, x, y, z\}}^t = \{d\}$.
 - $d \sim_P^X x$ only when you have both $y, a \in X$.
8. $\{d', x, y, z\} : \llbracket P \rrbracket_{\{d', x, y, z\}} = \{d'\}$ and $\sim_P^{\{d', x, y, z\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d', x, y, z\}}^t = \{d'\}$.
9. $\Rightarrow \{d, x, z, a\} : \llbracket P \rrbracket_{\{d, x, z, a\}} = \{d\}$. And $\sim_P^{\{d, x, z, a\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d, x, z, a\}}^t = \{d\}$.
 - $d \sim_P^X x$ only when you have both $y, a \in X$.
10. $\{d', x, z, a\} : \llbracket P \rrbracket_{\{d', x, z, a\}} = \{d'\}$. And $\sim_P^{\{d', x, z, a\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d', x, z, a\}}^t = \{d'\}$.
11. $\{x, y, z, a\} : \llbracket P \rrbracket_{\{x, y, z, a\}} = \{x, y, z\}$. And $\sim_P^{\{x, y, z, a\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{x, y, z, a\}}^t = \{x, y, z\}$.
12. $\Rightarrow \{d, d', x, y, z\} : \llbracket P \rrbracket_{\{d, d', x, y, z\}} = \{d, d'\}$. And $\sim_P^{\{d, d', x, y, z\}}$ is non-trivially empty. So $\llbracket P \rrbracket_{\{d, d', x, y, z\}}^t = \{d, d'\}$.
13. $\Rightarrow \{\mathbf{d}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{a}\} : \llbracket P \rrbracket_{\{d, x, y, z, a\}} = \{d, x, y, z\}$. Because $d \sim_P^{\{d, x, y, a\}} x$, by Granularity, $d \sim_P^{\{d, x, y, z, a\}} x$. However, this doesn't matter because $x, z \in \llbracket P \rrbracket_{\{d, x, y, z, a\}}$. So $\llbracket P \rrbracket_{\{\mathbf{d}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{a}\}}^t = \{\mathbf{d}, \mathbf{x}, \mathbf{y}, \mathbf{z}\}$.
 - Because $a \in \{d, x, y, z, a\}$, Upward Difference is satisfied.
 - Because $d \sim_P^{\{d, x, y, z, a\}} x$, Granularity is satisfied.

- But x, z are not a tolerant difference pair.

14. $\Rightarrow \{d', x, y, z, a\} : \llbracket P \rrbracket_{\{d', x, y, z, a\}} = \{d', x, y, z\}$. Since $d' \sim_P^{\{d', x, y, a\}} y$, by Granularity, $d' \sim_P^{\{d', x, y, z, a\}} y$. And $\llbracket P \rrbracket_{\{d', x, y, z, a\}}^t = \{d', x, y, z\}$.

15. $\{d, d', x, z, a\} : \llbracket P \rrbracket_{\{d, d', x, z, a\}} = \{d, d', x, z\}$. $\sim_P^{\{d, d', x, z, a\}}$ is non-trivially empty. So $\{d, d', x, z, a\} : \llbracket P \rrbracket_{\{d, d', x, z, a\}}^t = \{d, d', x, z\}$.

16. $\{d, d', x, y, z, a\} : \llbracket P \rrbracket_{\{d, d', x, y, z, a\}} = \{d, d', x, y, z\}$. By Granularity, $d \sim_P^{\{d, d', x, y, z, a\}} x$ and $d' \sim_P^{\{d, d', x, y, z, a\}} y$. And $\llbracket P \rrbracket_{\{d, d', x, y, z, a\}}^t = \{d, d', x, y, z\}$.

Therefore, $x >_P^t y$ and $y >_P^t z$, but $x \not\approx_P^t z$. So $>_P^t$ is not transitive. \square

Part 3

Scale Structure

CHAPTER 5

Adjectival Scale Structure

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. These classes are repeated (with examples) below:

(1) **Relative Adjectives (RAs):**

tall, short, expensive, cheap, nice, friendly, intelligent, stupid, narrow, wide...

(2) **Total Absolute Adjectives (AA^Ts):**

bald, empty, full, clean, smooth, dry, straight, flat

(3) **Partial Absolute Adjectives (AA^Ps):**

dirty, bent, wet, curved, crooked, dangerous, awake...

(4) **Non-Scalar Adjectives (NSs):**

atomic, geographical, polka-dotted, pregnant, illegal, dead, hexagonal, opaque, transparent...

Following much previous work, I argue that the adjectives in each of the classes shown above are associated with scales that have different properties. In particular, as we will see, there are empirical arguments for proposing that *absolute total* adjectives, like those in (2), are associated with scales that have maximal elements, *absolute partial* adjectives, like those in (3), are associated with scales that have minimal elements, and *relative* adjectives, like those in (1), are associated with

scales that have neither minimal nor maximal elements. Additionally, we will see that, when non-scalar adjectives (4) are coerced into scalar adjectives, they can be associated with scales that have both minimal and maximal elements. Thus, I will argue that we see the following scale structure patterns in languages like English:

Pattern	Relative	Total	Partial	Coerced Non-Scalar
Maximal Element?	×	✓	×	✓
Minimal Element?	×	×	✓	✓

Table 5.1: Scale Structure Patterns

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 chapters 2 and 4. In other words, once we have an (independently necessary) analysis of context-sensitivity and (potential) vagueness in the adjectival domain, we get an analysis of adjectival scale structure ‘for free’. As will be discussed in chapter 7, the link between (lack of) context-sensitivity and scalar endpoints has been proposed before as an empirical generalization (for example, by Kennedy and McNally (2005) and Kennedy (2007)); however, (to my knowledge) the theory developed in the first parts of this dissertation is the only one that can account for this observation without invoking additional semantic or pragmatic principles, like *Interpretative Economy* (discussed in chapter 2).

The chapter is laid out as follows: in section 5.2, I present the data associated with the scale structure patterns shown in table 5.1. I first introduce tests for the presence of scalar endpoints and I apply these tests to RAs, total AAs and partial AAs. Then, I apply them to coerced 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 the first two parts of the dissertation. Finally, in section 5.4, I explore some further empirical consequences of the proposed framework, particularly those dealing with the phenomenon of antonymy. I therefore conclude that the present theory can account for a wide range of empirical phenomena associated with adjectival predicates in a simple and elegant way.

5.2 Scale Structure Patterns

This section presents the data (or at least an illustrative subset of it) concerning the boundedness or the unboundedness of the orders associated with adjectival predicates. I will first consider only scalar adjectives, leaving a discussion of coerced and non-coerced NSs to section 5.2.4. In the literature, we can identify two classes of scalar structure diagnostics: tests that (I argue) show that scales are associated with maximal elements and tests that show that scales are associated with minimal elements.

5.2.1 Tests for a Maximal Element

In this section, I present some empirical arguments that total AAs are associated with scales that have a maximal element, while both RAs and partial AAs are associated with scales with no maximal element.

5.2.1.1 The Accentuation Test

The first test for a scale with a maximal element is the *accentuation* test. As discussed in Kennedy (2007) (p. 46) (based on Unger (1975)), accenting an adjective (in a non-contrastive focus situation) forces the adjective to pick out those individuals that lie at a very high degree on the adjective's scale¹.

- (5) a. Sarah is TALL!
b. This towel is WET!
c. This line is STRAIGHT!

Once the adjective is stressed, with both RAs (6) and partial AAs (7), it is still possible to hypothetically increase the degree to which the predicate holds of the subject.

¹Unger suggests that accenting an adjective actually forces a more precise interpretation. I believe that there is something to this idea, but it would need to be developed better to apply to relative adjectives.

- (6) a. This film is INTERESTING, but it could be more interesting.
 b. Sarah is TALL, but she could be taller if she wore heels.
- (7) a. This towel is WET, but it could be wetter.
 b. This neighbourhood is DANGEROUS, but it could be more dangerous.

However, Unger and Kennedy observe that if the adjective that is stressed is a total AA, increasing the degree to which the subject satisfies the predicate results in a contradiction (8).

- (8) a. My glass is FULL # but it could be fuller.
 b. This line is STRAIGHT # but you could make it straighter.

We can therefore conclude that the scales associated with total AAs are bounded in a way that the scales associated with relative and partial absolute adjectives are not.

5.2.1.2 Strong Resultatives

A second argument that the scales associated with total AAs have an upper bound (unlike the scales associated with RAs and partial AAs) comes from the aspectual properties of *strong adjectival resultative secondary predication constructions*. Resultative adjectival secondary predication constructions are complex verbal predicates composed of an atelic activity verb like *hammer* or *wipe* (9), 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 (10).

- (9) a. John hammered the metal (?in an hour/for an hour).
 b. John wiped the table (?in an hour/for an hour).
- (10) 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 strong resultative construction². 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/wet³.

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

Green (1972) (his (6b-7b)).

(12) John hammered the metal flat/straight/*long/*expensive⁴.

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

(13) **Wechsler's Generalization:**

Only total AAs are licensed as strong resultative secondary predicates.

(13) is a robust generalization. As shown in (14) and (15), Italian (a Romance language that allows such constructions) and Dutch show the same distinctions as English.

(14) **Italian:**⁵

- a. *Gianni ha battuto il ferro piatto piatto.*
 Gianni has beaten the iron flat flat
 'Gianni beat the iron flat.' (Total absolute adjective)

²Note carefully that I am only talking about a certain subclass of resultative constructions: what are known as 'real', 'strong' Germanic-style complex resultative predicates. Other similar constructions (like 'fake' resultatives (ex. *John sang himself hoarse*) and Slavic-style 'weak' resultatives) with different aspectual and syntactic properties are not subject to the same restrictions. See Wechsler (2005a) and Beavers (2008) for discussion.

³See Dowty (1979) and Goldberg and Jackendoff (2004) for the conclusion that contrasts such as those in (11) show that resultative secondary predicates are idiomatic expressions that must be simply be memorized. See Wechsler (2005b) for a criticism of this approach.

⁴Imagine that the metal increases in value the more that it is worked.

⁵For 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.

- b. **Gianni ha battuto il ferro **lungo lungo***
Gianni has beaten the iron long long.
'*Gianni beat the iron long.' (Relative adjective)

(15) **Dutch:**

- a. *Jan heeft het metaal **plat** gehamerd.*
Jan has the metal flat hammered
'Jan hammered the metal flat.'
- b. **Jan heeft het metaal **lang** gehamerd.*
Jan has the metal long hammered.
'*Jan hammered the metal long.'

Why should (13) 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 (10)) requires both the presence of an incremental structure and an upper bound to this structure (cf. Krifka (1989), Krifka (1998), Rothstein (2004), Kratzer (2004), among very many others). The fact that the simple transitive VPs in (9) 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 that total AAs 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 total AAs and telicity in degree achievements (causative verbs formed from scalar adjectives: ex. *to lengthen*, *to straighten* etc.) 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 succinctly reproduce 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 total AAs and telic VP interpretations constitutes a strong empirical argument that these (and only) these adjectives are associated with scales that have maximal endpoints.

5.2.1.3 The Distribution of Modifiers

The third argument that all and only total AAs are associated with scales with a maximal element comes from the distribution of scalar modifiers. This argument is a little tricky because it involves using the distribution and interpretation of a wide variety of linguistic expressions whose syntax, semantics, and pragmatics are independently under investigation to draw conclusions about the semantic and pragmatic properties of scalar adjectives. This being said, in this subsection, we will see that many modificational elements group relative and partial AAs together against total AAs, and that proposing that all and only the total AAs are associated with maximal endpoint scales can help us account for this data.

The first modifier that is sensitive to the scale structure of AAs is *almost* in English. As observed by Cruse (1986), Rotstein and Winter (2004), and Kennedy and McNally (2005) (among others), while *almost* is perfectly fine with total AAs, it is strange with partial AAs⁶.

- (16)
- a. This towel is **almost** dry/*wet.
 - b. The stick is **almost** straight/*bent.
 - c. The table is **almost** clean/*dirty.
 - d. The metal is **almost** flat/*curved.
 - e. John is **almost** bald.

Furthermore, *almost* is generally much less acceptable with relative adjectives than with total adjectives.

- (17)
- a. John is **almost** *fat/*tall/*wide.
 - b. This watch is **almost** *expensive/*attractive/*fashionable.

⁶ Rotstein and Winter (2004) notice that, to the extent that *almost* with partial adjectives is ok for some speakers, *almost* with partial AAs requires a somewhat strange context (ex. their example (p.266)).

- (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.

Rotstein and Winter also show that partial and total AAs with *almost* give rise to different inferences; therefore, *almost* still makes an interpretative distinction between the two classes of AAs.

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 total AAs are possible with this modifier: only they have top endpoints.

Another modifier that shows a similar distribution (which, to my knowledge, has not been discussed before) is *loosely speaking*. While *loosely speaking* is acceptable with total AAs, it is impossible with RAs and partial AAs.

- (18) a. The towel is dry, loosely speaking.
b. The stick is straight, loosely speaking.
c. The room is empty, loosely speaking.
etc.
- (19) a. #John is tall/fat, loosely speaking.
b. #This watch is expensive, loosely speaking.
c. #Mary is attractive, loosely speaking.
- (20) a. #The towel is wet, loosely speaking.
b. #The stick is bent, loosely speaking.
c. #My dress is dirty, loosely speaking.

If we suppose that the function of *loosely speaking* is similar to *almost* in that it picks out just those individuals that are close to (but not at) the top endpoint of a scale, then the distribution of *loosely speaking* is expected if only total AAs have top endpoints.

Another set of modifiers in English and French actually appear to target this endpoint. For example, both English *absolutely* and French *strictement* ‘definitely/absolutely’⁷ only apply to total adjectives and appear to pick out the individuals at the top endpoint of the scale. Note that for *absolutely* we are interested in the maximal scalar use, not the emphatic use that can appear with all kinds of predicates (i.e. *Absolutely, John is tall!*).

⁷I thank Jérémy Zehr for discussion of the French data and the contrasts between *strictement* and *strictement parlant* (see below).

- (21) a. The room is **absolutely** empty (# but it could be emptier).
 b. The stick is **absolutely flat** (# but it could be flatter).
- (22) a. *La salle est strictement vide.*
 The room is strictly empty.
 ‘The room is absolutely empty.’
 b. *La branche est strictement droite.*
 The stick is strictly straight.
 ‘The stick is absolutely straight.’

We see that both maximal interpretations of *absolutely* and *strictement* are impossible with RAs and partial AAs.

- (23) a. John is absolutely tall/fat/happy. (only emphatic use)
 b. *Jean est strictement grand/gros/content.
- (24) a. John is absolutely wet/dirty. (only emphatic use)
 b. *Jean est strictement mouillé/sale.

If at least one of the uses of elements like *absolutely* and *strictement* can pick out the individuals who have the highest degree of a predicate, and only the scales associated with total AAs have such a degree, then it is understandable why these modifiers do not have maximal interpretations with RAs and partial AAs.

Finally, a third set of modifiers display a pattern of interpretative variation with total AAs, which would be expected if only these predicates were associated with scales that have maximal elements. As discussed by Kennedy and McNally (2005) and Sauerland and Stateva (2007) (among many others), when *completely* appears with total AAs, 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 (25).

- (25) a. John is completely bald \approx John has the highest degree of baldness.
 b. This room is completely empty \approx 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 (26a), 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*) (27).

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

- (27) a. John is **completely** happy.
b. Susan is **completely** red.
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)

We find the same pattern with partial AAs: in the examples in (28), only the mereological reading (not the maximal one) is available.

- (28) a. The cat is **completely** wet.
b. The cat is **completely** dirty.
c. Sarah is **completely** healthy.

Another modifier that shows a similar pattern is *tout* ‘all’ in French (and its cognates in the Romance languages, cf. Burnett (2011b)). When this element applies to total AAs, it receives a maximal interpretation.

- (29) a. *La salle est toute vide.*
The room is ALL empty.
‘The room is completely empty’ (lit. ‘The room is all empty’)
b. *Jean est tout chauve.*
Jean is ALL bald.
‘Jean is completely bald.’

However, no such interpretation is possible with other kinds of adjectives: with RAs and partial AAs, *tout* has an intensive interpretation (30) or both an intensive and a mereological interpretation (31).

- (30) a. *Marie est toute grande!*
 Marie is ALL big
 ‘Marie is all grown up!’
- b. *Marie est toute petite!*
 Marie is ALL small
 ‘Marie is really small!’
- (31) a. *Jean est tout content.*
 Jean is ALL happy
 ‘Jean is really happy’ or ‘Jean is happy in all aspects of his life’.
- b. *Le chat est tout mouillé.*
 The cat is ALL wet.
 ‘The cat is really wet’ or ‘The cat is wet in all of his parts’.
- c. *Le chat est tout sale.*
 The cat is ALL dirty
 ‘The cat is really dirty’ or ‘The cat is dirty in all of his parts’

Finally, a third modifier, *plumb* in Appalachian English, shows a maximal interpretation with total AAs that disappears with RAs and partial AAs. In her article on the syntax and semantics of this element, Abner (2011) shows that, while *plumb* can have an emphatic interpretation with an adjective of any scale structure class, only total AAs can give rise to maximal interpretations (32) (All examples with *plumb* are from Abner (2011)).

- (32) a. You make sure to get that fence **plumb** straight.
- b. My bank account’s **plumb** empty, ain’t got no money at all.
- Abner (2011)’s (28)

Note that *plumb* with total AAs can also have an intensive interpretation (33); indeed, Abner observes that sentences where *plumb* modifies total AAs are generally ambiguous between maximal/completive readings and intensive readings.

(33) I'm **plumb** clean-yis'dy wuz Sat'dy. Cou'se, mebbly not as clean as an Apostolic man; but purty clean fur a Prisbyterian.

(≈ pretty clean, The Atlantic Monthly Vol. 144 1929; Abner (3))

With relative adjectives, *plumb* can have an intensive interpretation (34) or an excessive interpretation (35), but never a maximal interpretation. Abner (p.c.) notes that the same pattern is found with partial AAs.

(34) a. Sist' Humphreys be'n an edicated lady; an' she is a **plumb** good cook.

(≈ very good, The Century Vol. 68; cited from Abner (2011))

b. That **plumb** old preacher from up in the holler come in here yesterday talkin' some business about fire and brimstone.

(≈ very old; Abner's (13))

(35) I s'pose I seem **plumb** old for sech foolishness.

(≈ too old; Baily 1915, cited in Abner (2011) (30))

In summary, I have shown (based on previous research) that many modifiers in many languages/dialects have maximal interpretations with total AAs, and that these interpretations disappear with other kinds of adjectives. A summary of the data associated with 'maximal' modifiers is shown in table 5.2.

Interpretation	<i>tall/short</i>	<i>familiar/happy</i>	Partial AAs	Total AAs
<i>strictement</i>	*	*	*	maximal
<i>absolutely</i>	emphatic	emphatic	emphatic	maximal/emphatic
<i>completely</i>	*	mereological	mereological	maximal
<i>tout</i>	intensive	intensive/mereological	intensive/mereological	maximal
<i>plumb</i>	intensive/excessive	intensive/excessive	intensive/excessive	intensive/maximal

Table 5.2: The fine-grained nature of modification patterns

We can make two observations about the patterns in table 5.2: firstly, we can see that various modifiers interact with members of the scale structure classes (and subclasses of these classes) in very different ways: some (like *absolutely*) have emphatic interpretations, some (like *tout* and *plumb*) have intensive interpretations, and others, like *completely* and *tout* have mereological in-

terpretations. Although the distribution of an item like *completely* is often presented as a simple and fundamental test for having a top endpoint (cf. Wechsler (2005a), Beavers (2008), McNally (2011), among others), the distribution and interpretation of these modifiers are not trivial matters. Nevertheless, at the same time, we can observe that, despite this variation, all these expressions get maximal interpretations with total AAs. Therefore, based on this data, it seems reasonable to conclude that the scales that total AAs are associated with have maximal elements, and it's this maximality that the modifiers are picking up on. In other words, even though the semantic and pragmatic contribution of these modifiers is still very much under investigation (cf. Moltmann (1997), Kennedy and McNally (2005), Sauerland and Stateva (2007) for *completely*, Moltmann (1997), Moltmann (2005), Burnett (2011b) for adjectival *tout* and its cognates, Wolfram (2004), Abner (2011) for *plumb* etc.), I believe that the sheer consistency of the maximal interpretations that we see with total AAs across modifiers and across languages constitutes a solid argument that these predicates are associated with scales with upper bounds.

5.2.1.4 Summary

In this section, I have presented a number of tests for the presence of a maximal element (i.e. an upper bound) on an adjective's scale. These tests are summarized in table 5.3.

Maximal Endpoint Test	Relative	Total Absolute	Partial Absolute
Accentuation test	×	✓	×
Strong resultative test	×	✓	×
Telic degree achievement	×	✓	×
<i>Almost</i>	×	✓	×
<i>Loosely speaking</i>	×	✓	×
<i>Absolutely</i>	×	✓	×
<i>Strictement</i>	×	✓	×
Maximal <i>completely</i>	×	✓	×
Maximal <i>tout</i>	×	✓	×
Maximal <i>plumb</i>	×	✓	×

Table 5.3: Tests for a Maximal Endpoint

From this table, we can see that only total AAs pass these tests, and, therefore, I propose (following many authors) that only total AAs have scales with maximal endpoints. So we can start filling in the 'scale structure' table as in table 5.4.

Pattern	Relative	Total	Partial
Maximal Element?	×	✓	×
Minimal Element?			

Table 5.4: Scale Structure of Scalar Adjectives

5.2.2 Tests for a Minimal Element

There are many fewer tests for the presence of a minimal element than for the presence of a maximal element, and both of them involve the distribution and interpretation of modifiers. In fact, I only know of two tests: the ‘existential slightly’ test and the ‘strictly speaking’ test.

5.2.2.1 Existential Slightly

The modifiers *slightly* and *a little* can combine with a scalar adjective of any class; however, with partial adjectives, they can receive an additional interpretation that is impossible with both total and relative adjectives (see Solt (2011) 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.

(36) 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.

(37) Total Adjectives

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

(38) 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 (39) 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.

- (39) 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)⁸.

- (40) 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 club is a little empty.
 Only ‘The club 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 partial AAs are the only adjectives with bottom endpoints, they are the only ones that can give *slightly/a little/un peu* an existential interpretation.

5.2.2.2 Strictly Speaking

A second way in which we can see the contribution of the minimal endpoint of a partial AA is through the distribution of the phrase *strictly speaking*. This expression can modify a partial AA,

⁸I thank Francis Corblin for bringing this to my attention.

in which case it seems to restrict the denotation of the adjective to just those individuals that lie close to the bottom endpoint of the scale.

- (41)
- a. Strictly speaking, my dress is dirty, but it's not too dirty for me to wear to work.
 - b. Strictly speaking, this towel is wet, but I can still use it to dry myself.
 - c. Strictly speaking, this stick is bent, but it will still work as an impromptu ruler.

On the other hand, sentences with both total AAs and RAs are incoherent with *strictly speaking*.

(42) **Total AAs**

- a. #Strictly speaking, the room is empty, but...?
- b. #Strictly speaking, my dress is clean, but...?
- c. #Strictly speaking, this stick is straight, but...?

(43) **Relative Adjectives**

- a. #Strictly speaking, John is tall/fat, but...?
- b. #Strictly speaking, Mary is short/beautiful, but...?
- c. #Strictly speaking, Phil is friendly/attractive, but...?

Contrary to the very similar expression *strictement*, the French *strictement parlant* 'strictly speaking' shows the same pattern as its English counterpart.

- (44)
- a. *Strictement parlant, ma robe est sale, mais elle n'est pas trop sale pour porter.*
Strictly speaking, my dress is dirty, but she not is not too dirty to wear.
'Strictly speaking, my dress is dirty, but it is not too dirty to wear.'
 - b. #*Strictement parlant, cette salle est vide.*
Strictly speaking, this room is empty.
'Strictly speaking, this room is empty.'
 - c. #*Strictement parlant, Jean est grand.*
Strictly speaking, Jean is tall.
'Strictly speaking, Jean is tall.'

If we propose that *strictly speaking* targets the bottom endpoint of a scale, as it intuitively seems to be doing, and that only partial AAs have bottom endpoints, then we can explain the distributional patterns shown above.

5.2.3 Summary of Scale Structure Patterns

In summary, I gave two linguistic tests for the presence of a minimal element. The results are summarized in table 5.5:

Minimal Endpoint Test	Relative	Total Absolute	Partial Absolute
Existential <i>slightly</i>	×	×	✓
<i>Strictly speaking</i>	×	×	✓

Table 5.5: Tests for a Minimal Endpoint

Thus, I propose (following many authors) that total AAs are associated with scales with maximal endpoints, partial AAs 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 5.6.

Pattern	Relative	Total	Partial
Maximal Element?	×	✓	×
Minimal Element?	×	×	✓

Table 5.6: Scale Structure of Scalar Adjectives

5.2.3.1 Fully Closed Scale Adjectives

Readers that are very familiar with the literature on scale structure will notice that my description of the data has proceeded along the total/universal and partial/existential lines. While, in many works (ex. Kamp and Rossdeutscher (1994), Yoon (1996), Rotstein and Winter (2004), Moltmann (2009), a.o.), a distinction is made between only these two classes of absolute adjectives, in many other works (ex. Kennedy and McNally (2005), Kennedy (2007), Rett (2008), Sassoon and Toledo (2011) a.o.), the class of total AAs is further proposed to be divided into *upper closed scale* adjectives and *fully closed scale* adjectives. In this section, I go through the tests that have been proposed to

distinguish ‘fully closed adjectives’ as a class of AAs. I apply them to the set of adjectives that have been proposed to be fully closed in the literature, and I argue that, for most of the predicates, the results do not support the claim that these adjectives have scales with both top and bottom elements. The list of proposed ‘fully-closed scale’ adjectives (compiled from Kennedy and McNally (2005), Kennedy (2007), and Rett (2008)) is given in (45).

(45) **Proposed ‘fully closed scale’ adjectives:**

empty/full, open/closed, visible/invisible, transparent/opaque, complete/incomplete, perfect/imperfect

In my opinion, by far the most reliable test for having a scale that has both a maximal and minimal element is simply whether or not you pass all the tests discussed in the previous section⁹. For example, if we consider the adjective pair *open* and *closed*, Kennedy and McNally (2005) observe that these adjectives are compatible with both *completely* and existential *slightly*.

- (46) a. The door is **completely** open/closed.
b. The door is **slightly** open/closed.

Therefore, they conclude that both *open* and *closed* have scales with both a top and bottom endpoint. However, I argue that, if we consider the wide range of tests presented above, we see a

⁹ A second diagnostic that Kennedy and McNally (2005) give for fully-closed adjectives is compatibility with proportional modifiers like *half*. Thus, the fact that the examples in (i) are all good is claimed to indicate that the adjectives in (45) must be associated with scales with two endpoints.

- (i) a. The door is **half** open.
b. The door is **half** closed.
c. The room is **half** full.
d. The room is **half** empty.

However, it’s not clear that the ‘half’ test picks out exactly the adjectives that Kennedy and McNally (2005) claim to be fully closed (45) to the exclusion of top-closed ones: *half* is fine with properly total adjectives *bald*, *straight*, and *clean*. I therefore conclude that we need a different criterion to establish the existence of fully closed adjectives.

- (ii) a. John is **half** bald.
b. This stick is **half** straight.
c. The room is **half** clean.

different pattern. Consider for example the *almost* test: modifying *closed* with *almost* is perfectly fine; however, modifying *open* with this expression is bizarre and receives the ‘temporal’ kind of interpretation we see when partial adjectives appear with this modifier (see footnote 6).

- (47) a. The door is almost closed.
 b. ?The door is almost open.
 (only: ‘The lock is sticking, but...’)

Secondly, only *closed* passes the accentuation test:

- (48) a. #The door is CLOSED, but it could be more closed.
 b. The door is OPEN, but it could be more open.

Furthermore, while *closed* passes the tests for a top endpoint with modifiers *absolutely*, *strictement*, and *loosely speaking*, *open* does not.

- (49) a. The door is absolutely closed.
 (highest degree interpretation possible)
 b. *La porte est strictement fermée.*
 The door is strictly closed
 ‘The door is absolutely closed’
 c. The door is closed, loosely speaking.
- (50) a. The door is absolutely open! (emphatic interpretation only)
 b. **La porte est strictement ouverte.*
 c. #The door is open, loosely speaking, but...?

Finally, both *tout* and *plumb* only have intensive interpretations with *open* (although maximal interpretations are possible with *closed*).

- (51) a. *La porte est toute fermée*
 The door is ALL closed

- ‘The door is completely closed.’
- b. The door is plumb closed. (\approx completely closed; Abner (p.c.))
- (52) a. *La porte est toute ouverte.*
 The door is ALL open
 ‘The door is really open’ (but it could be more open)
- b. The door is plumb open. (\approx really open; Abner (p.c.))

This being said, when we look at the telicity test, the data are less clear. *Closed* clearly passes both the resultative test and the degree achievement test (it can be the secondary predicate to a strong resultative (53a) and its degree achievement verb is telic (53b)).

- (53) a. John pounded the door closed.
- b. John was closing the door \nrightarrow John closed the door.

For me, the resultative with *open* in (54a) is not great, but certainly better than *John wiped the table wet*, and (54b) shows that the degree achievement verb *to open* is clearly telic.

- (54) a. ?John pounded the door open.
- b. John was opening the door \nrightarrow John opened the door.

We can therefore summarize the results of the application of the tests for a maximal endpoint to *open/closed* as in table 5.7.

Now consider the tests for the presence of a minimal endpoint. We saw (following Kennedy and McNally (2005)) that both *open* and *closed* were compatible with existential *slightly*, but, when we look at the distribution of *strictly speaking*, we see a different pattern. As shown in (55), *strictly speaking* is compatible with *open* but not with *closed*.

- (55) a. #Strictly speaking, the door is closed, but...?
- b. Strictly speaking, the door is open, but we can’t get through.

Maximal Endpoint Test	<i>open</i>	<i>closed</i>
Accentuation test	×	✓
Strong resultative test	?	✓
Telic degree achievement	✓	✓
<i>Almost</i>	×	✓
<i>Loosely speaking</i>	×	✓
<i>Absolutely</i>	×	✓
<i>Strictement</i>	×	✓
Maximal <i>completely</i>	✓	✓
Maximal <i>tout</i>	×	✓
Maximal <i>plumb</i>	×	✓

Table 5.7: Tests for a Maximal Endpoint with *open/closed*

Thus, the results of the ‘minimal element’ diagnostics are shown in table 5.8.

Minimal Endpoint Test	<i>open</i>	<i>closed</i>
Existential <i>slightly</i>	✓	✓
<i>Strictly Speaking</i>	✓	×

Table 5.8: Tests for a Minimal Endpoint

Although we have seen some variation, given these results, I believe that the safest conclusion to make is that *open* is a partial adjective like *wet*, albeit one that interacts with modifiers like *completely* in somewhat of a different way. Furthermore, given the patterns in table 5.7 and 5.8, I believe that it is appropriate to conclude that *closed* is simply a total adjective like *dry*.

Concerning the rest of the adjectives in (45) (repeated below): both *empty* and *full* pass only the tests for having a maximal endpoint; therefore, I see no reason to think that they are associated with scales with minimal endpoints (but cf. footnote 9 for an argument based on proportional modifiers).

(56) **Proposed ‘fully closed scale’ adjectives:**

empty/full, open/closed, visible/invisible, transparent/opaque, complete/incomplete, perfect/imperfect

Indeed, someone who proposes that *empty* and *full* have scales that are closed at the bottom would have to explain they pattern like proper total AAs like *dry* and *clean* in the data in (57) and (58).

- (57) a. The room is slightly/a little empty. (Only excessive interpretation)
 b. The room is slightly/a little full. (Only excessive interpretation)
 c. The towel is slightly/a little dry. (Only excessive interpretation)
- (58) a. #The room is, strictly speaking, empty, but...?
 b. #The room is, strictly speaking, full, but...?
 c. #The towel is, strict speaking, dry, but...?

Both *perfect* and *imperfect* sound very bizarre to me in the comparative out of context; therefore, I will discuss these adjectives in the following section on coerced non-scalar adjectives. This leaves us *visible*, *invisible*, *transparent*, *opaque*, *complete*, and *incomplete*. Here, indeed, we see variation. Consider first the tests for a minimal endpoint.

(59) **Existential Slightly**

- a. Your application is slightly complete.
 b. Your application is slightly incomplete.
 c. The North star is slightly visible.
 d. *The North star is slightly invisible. (if ok, only excessive)

- (60) a. Your dress is slightly transparent. (Only excessive)
 b. This is a teapot that is slightly opaque. Not entirely opaque, because that would be pretentious.

From <http://manloveisthegreatest.deviantart.com/art/A-slightly-opaque-teapot-114739367>

(61) **Strictly Speaking**

- a. Your application is, strictly speaking, complete, but you didn't include any of the extra documents.
 b. Your application is, strictly speaking, incomplete, but the important information is filled in.
 c. The North star is, strictly speaking, visible from earth, but you need a powerful tele-

scope.

- d. #The North star is, strictly speaking, invisible, but... ?
- (62)
- a. #Your dress is, strictly speaking, transparent, but... ?
 - b. Your dress is, strictly speaking, opaque, but it's a very thin material...

Minimal Endpoint Test	<i>complete</i>	<i>incomplete</i>	<i>visible</i>	<i>invisible</i>	<i>transparent</i>	<i>opaque</i>
Existential <i>slightly</i>	✓	✓	✓	×	×	✓
Strictly Speaking	✓	✓	✓	×	×	✓

Table 5.9: Tests for a Minimal Endpoint

Now we can consider tests for maximal endpoints. For space considerations, I will only present a representative subset of these tests.

(63) **Accentuation Test**

- a. #Your application is COMPLETE, but it could be more complete.
- b. Your application is INCOMPLETE, but it could be more incomplete.
- c. The North star is VISIBLE, but it could be more visible.
- d. #The North star is INVISIBLE, but it could be more invisible.

(64) a. #Your dress is TRANSPARENT, but it could be more transparent.

- b. #Your dress is OPAQUE, but it could be more opaque.

(65) **Modifiers**

- a. Your application is *completely/almost/absolutely/loosely speaking* complete.
- b. Your application is *completely/*almost/# absolutely/*loosely speaking* incomplete.
- c. The North star is *completely/?almost/# absolutely/*loosely speaking* visible.
- d. The North star is *completely/almost/absolutely/loosely speaking* invisible.

(66) a. Your dress is *completely/almost/absolutely/loosely speaking* transparent.

- b. Your dress is *completely/almost/absolutely/loosely speaking* opaque.

Maximal Endpoint Test	<i>complete</i>	<i>incomplete</i>	<i>visible</i>	<i>invisible</i>	<i>transparent</i>	<i>opaque</i>
Accentuation Test	✓	×	×	✓	✓	✓
<i>Almost</i>	✓	×	?	✓	✓	✓
Maximal <i>absolutely</i>	✓	×	×	✓	✓	✓
<i>Loosely Speaking</i>	✓	×	×	✓	✓	✓
Maximal <i>completely</i>	✓	×	×	✓	✓	✓

Table 5.10: (Some) Tests for a Maximal Endpoint

In summary, only *complete* and *opaque* pass the tests for having both a top and bottom endpoint. Although there is some variation in the tests, the other adjectives in (45) display either the regular total or partial pattern. Thus, we can update our table of scale structure patterns as in table 5.11:

Pattern	Relative	Total	Partial	{<i>complete, opaque</i>}
Maximal Element?	×	✓	×	✓
Minimal Element?	×	×	✓	✓

Table 5.11: Scale Structure of Scalar Adjectives

In the next section, we will see a whole other class of adjectives that show the same pattern as *complete* and *opaque*: coerced non-scalar adjectives.

5.2.4 Non-Scalar Adjectives

When they are used precisely, non-scalar adjectives 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 coerce these predicates into scalar ones. So now we might investigate what the properties of these coerced scales are.

In what follows, note that the judgements presented below are either my own or taken from other sources whose judgements 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. (1983)'s experiment could associate a scale with the predicate *even*. I find scalar coercion of *even* extremely difficult to do and find a comparative like (67) (based on the results of the Armstrong et al. (1983) study) incoherent.

(67) 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 exist, I believe we can make some generalizations about how coerced NSs can differ from ‘true’ AAs.

For some adjectives, coercion creates a total AA: this is what we see with coerced *hexagonal*¹⁰.

(68) **Top endpoint tests**

- a. France is almost hexagonal.
- b. Loosely speaking, France is hexagonal.

(69) **Bottom endpoint tests**

- a. *This shape is slightly hexagonal.
- b. # Strictly speaking, this shape is hexagonal, but. . .

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

(70) **Top endpoint tests**

- a. *Mary is almost pregnant.
- b. #Loosely speaking, Mary is pregnant. . .

(71) **Bottom endpoint tests**

- a. Mary is slightly pregnant.
(She’s showing, but not very much).
- b. Strictly speaking, Mary is pregnant, but she’s not showing very much.

However, many coerced non-scalar adjectives seem to be able to be associated with both a scale

¹⁰Although (69b) is much stranger than (68b), I think it might be possible to say (i).

(i) Strictly speaking, this shape is hexagonal, but its sides are all uneven, so it looks weird.

So maybe it is possible to associate *hexagonal* with a scale with a minimal element after all.

with a maximal element and a scale with a minimal element. As shown in (72) and (73), *dead* is one of these adjectives.

- (72) 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.
- (73) a. Dead person is Actually Only **Slightly** Dead.
Headline from http://www.toplessrobot.com/2010/08/dead_person_caught_on_google_street_view_not_actua.php
- b. Strictly speaking, Zombie John is dead, but he’s still chasing after us to eat our brains.

Nationality terms like *Canadian* or *French* are also adjectives that permit both maximal endpoint and minimal endpoint scales (74)-(75).

- (74) 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 almost/completely gone through the citizenship process.)
- (75) a. She has lived in France for 20 years, so, loosely speaking, Mary is French.
- b. Strictly speaking, Marie is French, but she moved away when she was five years old.

Finally, we can see the same pattern with *illegal*:

- (76) a. Your accountant’s tax practices are **almost/completely** illegal.
 b. Your accountant’s tax practices are **slightly/a little** illegal.
 c. **Strictly speaking**, jaywalking is illegal, but no one ever gets fined for it in Montréal.

In summary, although there is some variation, with many non-scalar adjectives it is relatively easy to coerce them into both partial and total AAs. Thus, I propose that (at least most) coerced NSs are like *complete* and *opaque*¹¹ associated with scales that have both top endpoints and bottom endpoints.

5.2.5 Summary of Scale Structure Data

In this section, I have argued for the following scale structure patterns (shown in table 5.12): Relative adjectives are associated with scales with no endpoints, total AAs are associated with scales with a top endpoint, partial AAs are associated with scales with a bottom endpoint, and non-scalar adjectives, when they are coerced, are associated with scales with a top endpoint, a bottom endpoint or both.

Pattern	Relative	Total	Partial	Coerced Non-Scalar
Maximal Element?	×	✓	×	✓
Minimal Element?	×	×	✓	✓

Table 5.12: Scale Structure Patterns

In the next section, I show that the patterns in table 5.12 are exactly predicted by the analysis in parts 1 and 2 of the dissertation.

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 $\approx_p^{t/s}$ relations. Since, in this framework, degrees are equivalence classes of individuals ($[x]_{\approx_p s}$), if the domain D is finite, then the scales associated

¹¹Given this result, we might wonder whether there is really such a difference between an adjective like *complete* and one like *dead*. Maybe *complete* is simply a (frequently) coerced non-scalar adjective.

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, partial AAs, total AAs or coerced NSs) must have both a top element and a bottom element. So how can we account for the open, top-closed, bottom-closed and fully-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 individuals¹². 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 definitions¹³:

Definition 5.3.1 Top-closed scale. For a predicate P , $>_P$ is an top-closed scale iff for all models M and 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 .

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 , $>_P$ is an bottom-closed scale iff for all models M and all extensions of M, M' , there is no $x \in D_{M'} - D_M$ such that $d >_P^n x$ in M' , for $d : \neg \exists d' : d >_P d'$ in M .

¹²This strategy was suggested to me by Denis Bonnay.

¹³A more general statement of the top-closed (and, in fact, of all the scale structure properties) would be the existential statement in (i). The definitions are given for the $>_P$ relations, but the *top-closed/bottom-closed/open* properties can be defined from the $>_P^t$ and $>_P^s$ relations in a parallel way.

- (i) **Top-closed scale.** For a predicate P , $>_P$ is an *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 CCs in my system (i.e. the AAA), (i) 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 dissertation.

Thus, we'll 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 , $>_P$ is an open scale iff $>_P$ is neither top-closed nor bottom-closed.

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. In the next subsection, I present a series of results associated with these scale structure properties.

5.3.1 Scale Structure Results

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 5.13.

Adjective	$>_P$: non-trivial SWO?	$>^t_P$: non-trivial SWO?	$>^s_P$: non-trivial SWO?
Relative	✓	×	×
Total Absolute	×	✓	×
Partial Absolute	×	×	✓
Non-Scalar	×	×	×
Coerced Non-Scalar	×	✓	✓

Table 5.13: Scalarity Patterns

Using the definitions presented above, we can show a further series of results concerning the properties of those non-trivial scales in table 5.13. 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 5.3.1 *Total Top Endpoint.* For all $Q \in AA$, all models M , and $d, d' \in D$,

- If $d \in \llbracket Q \rrbracket_D$ then there is no $d' \in D$ such that $d' >^t_Q d$.

Proof Let $d \in \llbracket Q \rrbracket_D$ and suppose there is some d' such that $d' >^t_Q d$. Then, there is some $X \in CC$ such that $d' \in \llbracket Q \rrbracket^t_X$ and $d \notin \llbracket Q \rrbracket^t_X$. But $d \in \llbracket Q \rrbracket_D$, so by the AAA, $d \in \llbracket Q \rrbracket^t_X$. \perp □

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 show that, given the fact in lemma 5.3.1, an AA's tolerant scale ($>^t_Q$) is top closed (i.e. has a maximal element).

Theorem 5.3.2 *If $Q \in AA$, then $>^t_Q$ is a top-closed scale.*

Proof Let M be a CC t-model and let $x \in \llbracket Q \rrbracket_D$. Therefore, by lemma 5.3.1, there is no $y \in D_M$ such that $y >^t_Q x$. Now consider the CC t-model M' such that $D_{M'} = D_M \cup \{z\}$. Show $z \not>^t_Q x$. Suppose for a contradiction that $z >^t_Q x$. Then there is some $X \in CC_{M'}$ such that $z \in \llbracket Q \rrbracket_X^t$ and $x \notin \llbracket Q \rrbracket_X^t$. But, since M' is an extension of M , $x \in \llbracket Q \rrbracket_D$. So, by the AAA, $x \in \llbracket Q \rrbracket_X$ and $x \in \llbracket Q \rrbracket_X^t$. \perp \square

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

Lemma 5.3.3 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 $d >^s_Q d'$.*

Proof Let $d \notin \llbracket Q \rrbracket_D$ and suppose there is some $d' \in D$ such that $d >^s_Q d'$. Then, there is some $X \in CC$ such that $d \in \llbracket Q \rrbracket_X^s$ and $d' \notin \llbracket Q \rrbracket_X^s$. But $d \notin \llbracket Q \rrbracket_D$, so by the AAA, $d \notin \llbracket Q \rrbracket_X^s$. \perp \square

In other words, the analysis correctly predicts that the minimal element of the non-trivial scale associated with a partial AA 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 AA's strict scale is a bottom closed scale:

Theorem 5.3.4 *If $Q \in AA$, then $>^s_Q$ is a bottom closed scale.*

Proof Let M be a CC t-model and let $x \notin \llbracket Q \rrbracket_D$. Therefore, by lemma 5.3.3, there is no $y \in D_M$ such that $x >^s_Q y$. Now consider the CC t-model M' such that $D_{M'} = D_M \cup \{z\}$. Show $x \not>^s_Q z$. Suppose for a contradiction that $x >^s_Q z$. Then there is some $X \in CC_{M'}$ such that $x \in \llbracket Q \rrbracket_X^s$. 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, $x \notin \llbracket Q \rrbracket_X^s$. \perp \square

In other words, based on theorems 5.3.1 and 5.3.3, 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.

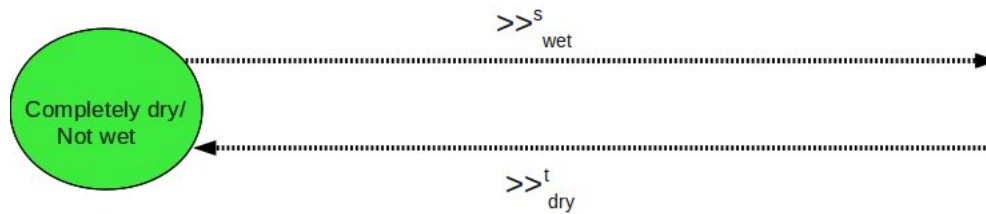


Figure 5.1: The Degree Scales ($\gg^{t/s}$) of *dry* and *wet*

This result replicates exactly a proposal made by Rotstein and Winter (2004) (p.260) to account for scale structure data like that presented in the first part of this chapter. We can note however that, while this coincidence between the endpoints of partial and total scales is part of Rotstein and Winter (2004)’s main proposal, it follows from the analysis of context-sensitivity and potential vagueness patterns developed in this work. Thus, we predict the patterns in table 5.14.

Pattern	Relative	Total	Partial	Coerced Non-Scalar
Maximal Element?		✓	×	
Minimal Element?		×	✓	

Table 5.14: Absolute Adjective Scale Structure Patterns

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. Therefore, we can fill in table 5.14 as in table 5.15.

Pattern	Relative	Total	Partial	Coerced Non-Scalar
Maximal Element?		✓	×	✓
Minimal Element?		×	✓	✓

Table 5.15: Absolute/Non-Scalar Scale Structure Patterns

While AAs are subject to the absolute adjective axiom, which imposes very strong constraints on the semantic denotations of absolute constituents, I proposed in the first part of the dissertation that RAs were only subject to van Benthem’s axioms (No Reversal, Upward Difference, and Downward Difference). These conditions are very weak (they just ensure an 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 5.3.5 *If $P \in RA$, then $>_Q$ is an open scale.*

Proof (Not Top Closed:) Let M be a CC-t-model and let $x \in D_M$ such that there is no $d \in D_M$ such that $d >_Q x$. Now consider the proper extension of M , M' , such that $D_{M'} = D_M \cup \{y\}$. Suppose that $y \in \llbracket P \rrbracket_{\{x,y\}}$ and $x \notin \llbracket P \rrbracket_{\{x,y\}}$. This is permitted (provided P still satisfies NR, UD, and DD) because $\llbracket P \rrbracket$ can vary across CCs. So $y >_Q x$. ✓

(Not Bottom Closed:) Let M be a CC-t-model and let $x \in D_M$ such that there is no $d \in D_M$ such that $x >_Q d$. Now consider the proper extension of M , M' , such that $D_{M'} = D_M \cup \{y\}$. Suppose $x \in \llbracket P \rrbracket_{\{x,y\}}$ and $y \notin \llbracket P \rrbracket_{\{x,y\}}$. So $x >_Q y$. ✓ □

We therefore correctly predict that RAs should pass neither the tests for having a maximal element nor the tests for having a minimal element, and we can complete table 5.15 as in table 5.16.

Pattern	Relative	Total	Partial	Coerced Non-Scalar
Maximal Element?	×	✓	×	✓
Minimal Element?	×	×	✓	✓

Table 5.16: Absolute/Non-Scalar Scale Structure Patterns

5.3.2 Summary

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 vagueness of scalar and non-scalar adjectives that was presented in the previous chapters. In other words, given the analysis presented earlier, 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 an top-closed scale within a delineation framework.

In the next section, I explore some further consequences of the framework, namely those dealing with antonymic relations between scalar adjectives.

5.4 Consequences: Antonymy

In this final subsection, I show that, with some simple assumptions about antonymic relations between scalar predicates, we can immediately account for certain important entailment relations between scalar antonyms.

5.4.0.1 Antonymic Relations

What kind of antonymic relations exist in the adjectival domain? As discussed in Cruse (1986), it is useful to adopt two notions from Aristotelian logic to describe such negative relations between adjectival predicates: *contrariness* and *contradictoriness*.

Definition 5.4.1 *Contraries*. Two predicates P, Q are contraries iff, for all $a \in D$,

(77) If a is P , then a is not Q .

Definition 5.4.2 *Contradictories*. Two predicates P, Q are contradictories iff, for all $a \in D$,

(78) a is P iff a is not Q .

Firstly, we can observe that the relative adjectives *tall* and *short* are contraries: John cannot be both tall and short at the same time.

(79) If John is tall then John is not short.

However, *tall* and *short* are not contradictories because, if John is not tall, then it does not necessarily follow that he is short.

(80) **False:** If John is not short, then he is tall.

As Cruse (1986) points out, contrariness but not contradictoriness seems to be a general property of pairs of relative adjectives:

(81) a. If this watch is expensive, then it is not cheap.
b. **False:** If this watch is not cheap, then it is expensive.

(82) a. If this doorway is wide, then it is not narrow.
b. **False:** If this doorway is not narrow, then it is wide.

(83) a. If John is intelligent, then he is not stupid.
b. **False:** If John is not stupid, then he is intelligent.

And so on...

We can also observe that some absolute adjectives, particularly total/universal AAs, are also in contrary relationships. Consider, for example, *empty* and *full*.

(84) a. If room A is *full* then it is not *empty*.
b. **False:** If room A is not *empty*, then it is *full*.

Thus, both RAs and AAs can be in contrary relationships with other predicates in their class. This being said, note that contrariness is a very weak property: if we consider the adjectives *frequent* and *impossible*, we can observe (as do Rotstein and Winter (2004)) that they are also contraries.

(85) a. If event A is *frequent*, then it is not *impossible*.
b. **False:** If event A is not *impossible*, then it is *frequent*.

However, certain relative contraries, like *tall/short* and *wide/narrow*, satisfy a stronger condition, one that concerns their comparative relations. In particular, the equivalences in (86) hold.

- (86) a. John is *taller* than Mary iff Mary is *shorter* than John.
b. Doorway A is *wider* than doorway B iff Doorway B is *narrower* than doorway A.

In other words, the scale associated with *short/narrow* is the converse/inverse of the scale associated with *tall/wide*. In some analyses (cf. Kennedy and McNally (2005), Rett (2008), Lassiter (2011) among others), this observation is cashed out by saying that *tall* and *short* as on the **same** scale, but their comparative relations are reversed.

At the same time, the equivalences in (86) distinguish RA pairs like *tall/short* and *wide/narrow* from AA contrary pairs like *fully/empty* and RA/AA pairs like *frequent/impossible*. First consider the *frequent/impossible* pair: unlike with *tall/short*, the equivalence in (87). Thus, we have found a test that distinguishes contraries whose scales are intimately related (i.e. *tall/short*) from those whose scales are not so closely related (i.e. *frequent/impossible*).

- (87) **False:** Event a is more *frequent* than event b iff event b is more *impossible* than event a.

Now consider the *empty/full* pair. We saw in chapter 4 that total AA comparatives are, in my words, ‘pseudo-evaluative’; that is, total comparatives are only felicitous if their subject is somewhat close to the top endpoint of their associated scale. Thus, the sentence in (88) is strange.

- (88) # This very full room is emptier than that very full room.

The ‘pseudo-evaluativity’ of *empty* and *full* is precisely what makes the equivalence in (89) fail.

- (89) **False:** This room is emptier than that room iff that room is fuller than this room.
(Countermodel: Both this room and that room are close to being completely full)

Furthermore, I argue that because (89) does fail, we have no argument that the scale associated with *empty* is the inverse of the scale associated with *full*. I therefore conclude, contrary to much of the literature on the topic (cf. Kennedy and McNally (2005), Rett (2008) etc.), that *empty* and *full* bear the same relationship to each other that *frequent* and *impossible* do: they are contraries, but they are not associated with the ‘same’ scale.

In summary, following Cruse and others, we can observe that relative adjectives have contrary antonyms (but not contradictory antonyms), and some of these contraries appear to be associated with the same scale¹⁴. Some AAs, like *empty* and *full*, also have contrary antonyms, but these antonyms do not appear to share a scale.

However, as observed by Cruse (1986), Yoon (1996), Rotstein and Winter (2004) i.a., there are some AAs that bear the stronger *contradictory* relation to each other. In particular, these are the total (i.e. universal)/partial (i.e. existential) pairs like *dry/wet*, *clean/dirty*, and *bent/straight*. For these pairs of adjectives, the following biconditionals hold:

- (90)
- a. This towel is dry iff it is not wet.
 - b. This towel is clean iff it is not dirty.
 - c. This stick is straight iff it is not bent.

Furthermore, the contradictory pairs also satisfy the equivalence in (91). Thus, we have an argument for locating them on the ‘same’ scales.

- (91)
- a. This towel is drier than that towel iff that towel is wetter than this towel.
 - b. This towel is cleaner than that towel iff that towel is dirtier than this towel.
 - c. This stick is straighter than that stick iff that stick is more bent than this stick.

¹⁴An example of an RA contrary pair that does not share a scale is *brilliant/stupid*. Although these adjectives both talk about intelligence, the equivalence in (i) does not hold:

- (i) **False:** John is more brilliant than Mary iff Mary is stupider than John.

In summary, relative adjectives can have contrary antonyms whose scales are inverses, and absolute adjectives can have contradictory antonyms whose scales are inverses¹⁵. In the next section, I will give the main lines of an analysis of the relation between contrary and contradictory antonyms within delineation TCS.

5.4.0.2 Analysis of Antonym Patterns

The framework proposed so far gives us a straightforward way of capturing contradictory antonymic relations between total and partial adjectives. Simply, we could propose that, from a total/partial pair, one of the two adjectives is actually lexically decomposed into the negation of the other. For the sake of illustration, let us suppose that the partial adjectives are the ones that are defined in terms of negation and the total AAs. Thus, the word *wet* is actually interpreted as *not dry*, the word *dirty* is actually interpreted as *not clean*, and so on.

- (92) a. $[wet] \Rightarrow [not [dry]]$.
 b. $[dirty] \Rightarrow [not [clean]]$.
 c. $[bent] \Rightarrow [not [straight]]$.

Furthermore, we might assume that indifference relations for a predicate and its negation are simply inverses:

(93) For all predicates P and $X \in CC$, $\sim_{not P}^X = (\sim_P^X)^{-1}$

If we adopt such an analysis, it is easy to show that 1) the total/partial pair will satisfy the condition for being contradictories and 2) the strict scale associated with the partial adjective will be exactly the inverse of the tolerant scale associated with the total adjective. Thus, we capture the intuition that adjectives like *wet* and *dry* are on the ‘same’ scale.

¹⁵Note that not all adjectives (relative or absolute) have antonyms: consider the total AA *bald*. To my knowledge, there is no scalar adjective in English that means “haired”. Thus, there is always a certain amount of lexical idiosyncrasy in antonymic relations between words.

Let Q_1, Q_2 be predicates such that $Q_1 \in AA^T$ and $Q_2 \in AA^P$. Suppose furthermore that Q_2 can be rewritten as *not* Q_1 . Then:

Theorem 5.4.1 For all $X \in CC$, and all $a \in X$, $a \in \llbracket Q_1 \rrbracket_X$ iff $a \in \llbracket \text{not } Q_2 \rrbracket_X$.

Proof \Rightarrow Let $X \in CC$ and $a \in \llbracket Q_1 \rrbracket_X$. So, by the definition of $\llbracket \cdot \rrbracket$, $a \notin \llbracket Q_1 \rrbracket_X$, and, by the definition of (classical) negation, $a \notin \llbracket \text{not } Q_1 \rrbracket_X$. Since *not* Q_1 can be rewritten as Q_2 , $a \notin \llbracket Q_2 \rrbracket_X$, and, by the definition of negation, $a \in \llbracket \text{not } Q_2 \rrbracket_X$. \Leftarrow By similar reasoning to \Rightarrow . \square

Furthermore:

Theorem 5.4.2 For all models M , $>_{Q_1}^t = (>_{Q_2}^s)^{-1}$.

Proof \subseteq Let M be a CC t-model and let $a, b \in D$ such that $a >_{Q_1}^t b$. Show $b >_{Q_2}^s a$. Since $a >_{Q_1}^t b$, there is some $X \in CC$ such that $a \in \llbracket Q_1 \rrbracket_X^t$ and $b \notin \llbracket Q_1 \rrbracket_X^t$. **Case 1:** $a \in \llbracket Q_1 \rrbracket_X$. Since $b \notin \llbracket Q_1 \rrbracket_X^t$, $b \notin \llbracket Q_1 \rrbracket_X$. By the definition of negation, $b \in \llbracket \text{not } Q_1 \rrbracket_X$. Since Q_2 can be rewritten as *not* Q_1 , $b \in \llbracket Q_2 \rrbracket_X$. Furthermore, since Q_2 can be rewritten as *not* Q_1 and $a \in \llbracket \text{not not } Q_1 \rrbracket_X$, $a \in \llbracket \text{not } Q_2 \rrbracket_X$, so the by the definition of negation, $a \notin \llbracket Q_2 \rrbracket_X$. So $b >_{Q_2} a$ and by theorem 4.6.4, $b >_{Q_2}^s a$. \checkmark

Case 2: $a \notin \llbracket Q_1 \rrbracket_X$. So there is some $d \in X$ such that $d \sim_{Q_1}^X a$ and $d \in \llbracket Q_1 \rrbracket_X$. Since $b \notin \llbracket Q_1 \rrbracket_X^t$, there is no $d' \in X : d' \sim_{Q_1}^X b$. Furthermore, $b \notin \llbracket Q_1 \rrbracket_X$. By the definition of negation, $b \in \llbracket \text{not } Q_1 \rrbracket_X$. Since Q_2 can be rewritten as *not* Q_1 , $b \in \llbracket Q_2 \rrbracket_X$. Furthermore, since $b \notin \llbracket Q_1 \rrbracket_X^t$, by (93), there is no $d' \in X : b \sim_{\text{not } Q_2}^X d'$. So $b \in \llbracket Q_2 \rrbracket_X^s$. Since $d \in \llbracket Q_1 \rrbracket_X$, $d \in \llbracket \text{not } Q_2 \rrbracket_X$. Since $d \sim_{Q_1}^X a$, and Q_2 can be rewritten as *not* Q_1 (so Q_1 can be rewritten as *not* Q_2), $d \sim_{\text{not } Q_2}^X a$. So by the definition of $\llbracket \cdot \rrbracket^s$, $a \notin \llbracket Q_2 \rrbracket_X^s$. So $b >_{Q_2}^s a$. \checkmark

\supseteq Let $a >_{Q_2}^s b$ to show $b >_{Q_1}^t a$. Since $a >_{Q_2}^s b$, there is some $X \in CC$ such that $a \in \llbracket Q_2 \rrbracket_X^s$ and $b \notin \llbracket Q_2 \rrbracket_X^s$. By the definition of $\llbracket \cdot \rrbracket^s$, $a \in \llbracket Q_2 \rrbracket_X$. Since Q_2 can be rewritten as *not* Q_1 , $a \in \llbracket \text{not } Q_1 \rrbracket_X$, so $a \notin \llbracket Q_1 \rrbracket_X$. Since $b \notin \llbracket Q_2 \rrbracket_X^s$, there is some $d \in X$ such that $d \sim_{Q_2}^X b$ and $d \in \llbracket \text{not } Q_2 \rrbracket_X$. Since Q_2 can be rewritten as *not* Q_1 , by double negation, $d \in \llbracket Q_1 \rrbracket_X$ and $d \sim_{Q_1}^X b$. So $b \in \llbracket Q_1 \rrbracket_X^t$. Furthermore, since $a \in \llbracket Q_2 \rrbracket_X^s$, there is no $d' \in X$ such that $d' \notin \llbracket Q_2 \rrbracket_X$ and $d' \sim_{Q_2}^X a$. Since Q_2 can be rewritten as *not* Q_1 , by double negation and (93), there is no $d' \in X$ such that $d' \in \llbracket Q_1 \rrbracket_X$ and $d' \sim_{Q_1}^X a$. Therefore, $a \notin \llbracket Q_1 \rrbracket_X^t$ and $b >_{Q_1}^t a$. \checkmark \square

I now turn to contrary antonyms, which are more complicated. Indeed, how to capture the relationship between *tall* and *short* is still very much an open question (cf. Krifka (2007), Rett (2008), Sassoon (2010) and references cited within these works). Thus, I will give a simplistic analysis of how the contrary antonymy of these adjectives could be treated within delineation TCS. I highlight that this proposal is meant to only give an illustration of the style of analysis that could be pursued within the framework, and it is not meant to be a comprehensive account of how *tall* and *short* are similar/different.

The first step is to give an account for contrary relations between predicates. I suggest that these relations (between *tall/short*, *empty/full*, and *frequent/impossible*) could be encoded through imposing constraints on the co-application of contrary predicates across comparison classes. One constraint that would have the desired effect is the following:

(94) Let P, Q be predicates that show contrariness.

For all models M , comparison classes X and individuals $x \in X$:

a. If $x \in \llbracket P \rrbracket_X$, then $x \notin \llbracket Q \rrbracket_X$.

The condition in (94) will rule out individuals being both *tall* and *short* or both *empty* and *full* in the same comparison class, without us having to define *short* as *not tall*. This is a welcome result since, on the face of it, *short* does not appear to be synonymous with *not tall*. For example, as discussed in chapter 3, while the sentence in (95a) may be acceptable for some speakers when considering borderline cases of vague predicates, replacing *not tall* by *short* in (95b) is a clear contradiction¹⁶.

(95) a. Mary is both tall and not tall.

b. #Mary is both tall and short.

Finally, we saw in the previous section that, unlike the *empty/full* and *frequent/impossible* pairs, *tall* and *short* seem to ‘share’ a scale. To capture this closer relationship within delineation TCS,

¹⁶A similar point is brought up as a challenge by von Stechow (1984) for the delineation approach to the semantics of gradable expressions.

we can propose an additional constraint that co-scalar adjectives are subject to. In particular, we propose that the application of *tall* within a comparison class partially determines the application of *short* in that class, in a way that the application of *frequent* does not determine the application of *impossible*. A condition that will do the trick within the assumptions made in this dissertation is the following:

(96) For all models M and all $X \in CC$, if there are $x, y \in X$ such that $x \in \llbracket \text{tall} \rrbracket_X$ and $y \notin \llbracket \text{tall} \rrbracket_X$, then there are $w, z \in X$ such that $w \in \llbracket \text{short} \rrbracket_X$ and $z \notin \llbracket \text{short} \rrbracket_X$.

(96) says that, if, in a comparison class, there is a tall/not tall contrast, then there also has to be a short/not short in that CC.

If we incorporate this new assumption into the framework, we can show that the scales associated with the semantic denotations of *short* will be inverses of the semantic scales associated with *tall*¹⁷.

Theorem 5.4.3 For all CC t-models M , $>_{\text{tall}} = (>_{\text{short}})^{-1}$.

Proof \subseteq Let M be a CC t-model and let $a, b \in D$. Suppose $a >_{\text{tall}} b$ to show that $b >_{\text{short}} a$. Since $a >_{\text{tall}} b$, by theorem 2.4.9 (two-element reducibility), $a \in \llbracket \text{tall} \rrbracket_{a,b}$ and $b \notin \llbracket \text{tall} \rrbracket_{a,b}$. Since $a \in \llbracket \text{tall} \rrbracket_{a,b}$, by (94), $a \notin \llbracket \text{short} \rrbracket_{a,b}$. By (96), there must be a short/not short contrast in $\{a, b\}$, so $b \in \llbracket \text{short} \rrbracket_{a,b}$. So $b >_{\text{short}} a$. $\checkmark \supseteq$ By parallel reasoning to \subseteq . \square

In summary, I conclude that a treatment of both contrary and contradictory scalar antonyms is possible within delineation TCS.

¹⁷Note that the proof of theorem 5.4.3 crucially makes use of the two-element reducibility property of RAs (theorem 2.4.9). I argued (following Kennedy and van Rooij) in chapter 2 that this property may be too strong to reflect the relationship between the positive form and comparative form of RAs. Thus, if we get rid of two-element reducibility, the condition in (96) would need to be strengthened.

5.5 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 dissertation. 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 chapters 2 and 4 concern how to properly analyze the context-dependence and the vagueness/imprecision 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. The next chapter sets out the framework in succinct and formal way and gives a summary of the empirical patterns analyzed in this work.

CHAPTER 6

Delineation TCS

6.1 Introduction

This chapter presents the system *Delineation TCS* that was developed in this dissertation to model the observed dependencies between (potential) vagueness, context-sensitivity, and scale structure. I first give the syntax and semantics of the basic system, and then I list the proposed axioms that characterize adjectival predicates of different classes. Finally, I give a summary of the main empirical ways in which RAs, total AAs, partial AAs, and NSs were argued to be distinguished.

6.2 Vocabulary

The vocabulary consists of the following expressions:

1. A series of individual constants: $a_1, a_2, a_3 \dots$
2. A series of individual variables: $x_1, x_2, x_3 \dots$
3. Four series of unary predicate symbols:
 - Relative scalar adjectives: $P_1, P_2, P_3 \dots$
 - Total absolute scalar adjectives: $Q_1^T, Q_2^T, Q_3^T \dots$
 - Partial absolute scalar adjectives: $Q_1^P, Q_2^P, Q_3^P \dots$
 - Non-Scalar adjectives: $S_1, S_2, S_3 \dots$
4. For every unary predicate symbol P , there is a binary predicate $>_P$.

5. Quantifiers \forall and \exists , and connectives \wedge, \vee, \neg and \rightarrow , plus parentheses.

When appropriate, I will use AA to refer to $AA^T \cup AA^P$, and I will use SA to refer to $RA \cup AA$.

6.3 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 t_1 and t_2 are terms and P is a predicate symbol, then $t_1 >_P t_2$ is a wff.
4. For any variable x , if ϕ and ψ are wffs, then $\neg\phi$, $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \rightarrow \psi$, $\forall x\phi$, and $\exists x\phi$ are wffs.
5. Nothing else is a wff.

6.4 Semantics

Definition 6.4.1 C-model. A *c-model* is a tuple $M = \langle D, m \rangle$ where D is a non-empty domain of individuals, and m is a function from pairs consisting of a member of the non-logical vocabulary and a comparison class (a subset of the domain) satisfying:

- For each individual constant a_1 , $m(a_1) \in D$.
- For each $X \in \mathcal{P}(D)$ and for each predicate P , $m(P, X) \subseteq X$.

Definition 6.4.2 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/comparison class pairs such that:

- For all P and all $X \in \mathcal{P}(D)$, \sim_P^X is a binary relation on X .

Definition 6.4.3 Assignment. An assignment for a *c/t-model* M is a function $g : \{x_n : n \in \mathbb{N}\} \rightarrow D$ (from the set of variables to the domain D).

Definition 6.4.4 Interpretation. An interpretation $\llbracket \cdot \rrbracket_{M,g}$ is a pair $\langle M, g \rangle$, where M is a t -model, and g is an assignment.

Definition 6.4.5 Interpretation of terms ($\llbracket \cdot \rrbracket_{M,g}$). For a model M , an assignment g ,

1. If x_1 is a variable, $\llbracket x_1 \rrbracket_{M,g} = g(x_1)$.
2. If a_1 is a constant, $\llbracket a_1 \rrbracket_{M,g} = m(a_1)$.

In what follows, for an interpretation $\llbracket \cdot \rrbracket_{M,g}$, a variable x_1 , and a constant a_1 , let $g[a_1/x_1]$ be the assignment for M which maps x_1 to a_1 , but agrees with g on all variables that are distinct from x_1 .

Definition 6.4.6 Classical Satisfaction ($\llbracket \cdot \rrbracket^c$). For all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$, all formulas ϕ, ψ , all predicates P , and all terms t_1, t_2 ,

1. $\llbracket P(t_1) \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in m(P, X) \\ 0 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in X - m(P, X) \\ i & \text{otherwise} \end{cases}$
2. $\llbracket t_1 >_P t_2 \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if there is some } X' \subseteq D : \llbracket P(t_1) \rrbracket_{M,g,X'}^c = 1 \text{ and } \llbracket P(t_2) \rrbracket_{M,g,X'}^c = 0 \\ 0 & \text{otherwise} \end{cases}$
3. $\llbracket \neg \phi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 1 \\ i & \text{otherwise} \end{cases}$
4. $\llbracket \phi \wedge \psi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 1 \text{ and } \llbracket \psi \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \{ \llbracket \phi \rrbracket_{M,g,X}^c, \llbracket \psi \rrbracket_{M,g,X}^c \} = \{1, 0\} \text{ or } \{0\} \\ i & \text{otherwise} \end{cases}$

$$5. \llbracket \phi \vee \psi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 1 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = \llbracket \psi \rrbracket_{M,g,X}^c = 0 \\ i & \text{otherwise} \end{cases}$$

$$6. \llbracket \phi \rightarrow \psi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 0 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^c = 1 \text{ and } \llbracket \psi \rrbracket_{M,g,X}^c = 0 \\ i & \text{otherwise} \end{cases}$$

$$7. \llbracket \forall x_1 \phi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^c = 1 \\ 0 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^c = 0 \\ i & \text{otherwise} \end{cases}$$

$$8. \llbracket \exists x_1 \phi \rrbracket_{M,g,X}^c = \begin{cases} 1 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^c = 1 \\ 0 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^c = 0 \\ i & \text{otherwise} \end{cases}$$

Tolerant and strict satisfaction are inter-defined.

Definition 6.4.7 Tolerant Satisfaction($\llbracket \cdot \rrbracket^t$). For all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$, all formulas ϕ, ψ , all predicates P , and all terms t_1, t_2 ,

$$1. \llbracket P(t_1) \rrbracket_{M,g,X}^t = \begin{cases} 1 & \text{if there is some } a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} : \llbracket P(a_1) \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in X, \text{ and there is no } a_1 \in X : a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} \\ i & \text{otherwise} \end{cases}$$

$$2. \llbracket t_1 >_P t_2 \rrbracket_{M,g,X}^t = \begin{cases} 1 & \text{if there is some } X' \subseteq D : \llbracket P(t_1) \rrbracket_{M,g,X'}^t = 1 \text{ and } \llbracket P(t_2) \rrbracket_{M,g,X'}^t = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned}
3. \llbracket \neg \phi \rrbracket_{M,g,X}^t &= \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 1 \\ i & \text{otherwise} \end{cases} \\
4. \llbracket \phi \wedge \psi \rrbracket_{M,g,X}^t &= \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 1 \text{ and } \llbracket \psi \rrbracket_{M,g,X}^t = 1 \\ 0 & \text{if } \{ \llbracket \phi \rrbracket_{M,g,X}^t, \llbracket \psi \rrbracket_{M,g,X}^t \} = \{1,0\} \text{ or } \{0\} \\ i & \text{otherwise} \end{cases} \\
5. \llbracket \phi \vee \psi \rrbracket_{M,g,X}^t &= \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 1 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^t = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = \llbracket \psi \rrbracket_{M,g,X}^t = 0 \\ i & \text{otherwise} \end{cases} \\
6. \llbracket \phi \rightarrow \psi \rrbracket_{M,g,X}^t &= \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 0 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^t = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 1 \text{ and } \llbracket \psi \rrbracket_{M,g,X}^t = 0 \\ i & \text{otherwise} \end{cases} \\
7. \llbracket \forall x_1 \phi \rrbracket_{M,g,X}^t &= \begin{cases} 1 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^t = 1 \\ 0 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^t = 0 \\ i & \text{otherwise} \end{cases} \\
8. \llbracket \exists x_1 \phi \rrbracket_{M,g,X}^t &= \begin{cases} 1 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^t = 1 \\ 0 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^t = 0 \\ i & \text{otherwise} \end{cases}
\end{aligned}$$

Definition 6.4.8 *Strict Satisfaction* ($\llbracket \cdot \rrbracket^s$). For all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$, all formulas ϕ, ψ , all predicates P , and all terms t_1, t_2 ,

$$1. \llbracket P(t_1) \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if for all } a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} : \llbracket P(a_1) \rrbracket_{M,g,X}^c = 1 \\ 0 & \text{if } \llbracket t_1 \rrbracket_{M,g} \in X, \text{ and there is no } a_1 \in X : a_1 \sim_P^X \llbracket t_1 \rrbracket_{M,g} \\ i & \text{otherwise} \end{cases}$$

$$2. \llbracket t_1 >_P t_2 \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if there is some } X' \subseteq D : \llbracket P(t_1) \rrbracket_{M,g,X'}^s = 1 \text{ and } \llbracket P(t_2) \rrbracket_{M,g,X'}^s = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$3. \llbracket \neg \phi \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = 1 \\ i & \text{otherwise} \end{cases}$$

$$4. \llbracket \phi \wedge \psi \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 1 \text{ and } \llbracket \psi \rrbracket_{M,g,X}^s = 1 \\ 0 & \text{if } \{ \llbracket \phi \rrbracket_{M,g,X}^s, \llbracket \psi \rrbracket_{M,g,X}^s \} = \{1, 0\} \text{ or } \{0\} \\ i & \text{otherwise} \end{cases}$$

$$5. \llbracket \phi \vee \psi \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 1 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^s = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^t = \llbracket \psi \rrbracket_{M,g,X}^t = 0 \\ i & \text{otherwise} \end{cases}$$

$$6. \llbracket \phi \rightarrow \psi \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 0 \text{ or } \llbracket \psi \rrbracket_{M,g,X}^s = 1 \\ 0 & \text{if } \llbracket \phi \rrbracket_{M,g,X}^s = 1 \text{ and } \llbracket \psi \rrbracket_{M,g,X}^s = 0 \\ i & \text{otherwise} \end{cases}$$

$$7. \llbracket \forall x_1 \phi \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^s = 1 \\ 0 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^s = 0 \\ i & \text{otherwise} \end{cases}$$

$$8. \llbracket \exists x_1 \phi \rrbracket_{M,g,X}^s = \begin{cases} 1 & \text{if for some } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^s = 1 \\ 0 & \text{if for every } a_1 \in X, \llbracket \phi \rrbracket_{M,g[a_1/x_1],X}^s = 0 \\ i & \text{otherwise} \end{cases}$$

6.5 Axiom Schemata

6.5.1 Some Definitions

We first define an equivalence relation \approx_P :

Definition 6.5.1 Equivalent. (\approx) For an interpretation $\llbracket \cdot \rrbracket_{M,g,X}$, a predicate P , $a_1, a_2 \in D$:

1. $a_1 \approx_P^c a_2$ iff $\llbracket a_1 >_P a_2 \rrbracket_{M,g,X}^c = 0$ and $\llbracket a_2 >_P a_1 \rrbracket_{M,g,X}^c = 0$.
2. $a_1 \approx_P^t a_2$ iff $\llbracket a_1 >_P a_2 \rrbracket_{M,g,X}^t = 0$ and $\llbracket a_2 >_P a_1 \rrbracket_{M,g,X}^t = 0$.
3. $a_1 \approx_P^s a_2$ iff $\llbracket a_1 >_P a_2 \rrbracket_{M,g,X}^s = 0$ and $\llbracket a_2 >_P a_1 \rrbracket_{M,g,X}^s = 0$.

Now we define \geq_P :

Definition 6.5.2 Greater than or equal. (\geq) For an interpretation $\llbracket \cdot \rrbracket_{M,g,X}$, a predicate P , $a_1, a_2 \in D$:

1. $a_1 \geq_P^c a_2$ iff $\llbracket a_1 >_P a_2 \rrbracket_{M,g,X}^c = 1$ or $a_1 \approx_P^c a_2$.
2. $a_1 \geq_P^t a_2$ iff $\llbracket a_1 >_P a_2 \rrbracket_{M,g,X}^t = 1$ or $a_1 \approx_P^t a_2$.
3. $a_1 \geq_P^s a_2$ iff $\llbracket a_1 >_P a_2 \rrbracket_{M,g,X}^s = 1$ or $a_1 \approx_P^s a_2$.

6.5.2 Axioms governing $\llbracket \cdot \rrbracket$

6.5.2.1 Relative Adjectives

For all $P_1 \in SA \cup NS$, all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$ and $a_1, a_2 \in X$ such that $\llbracket P_1(a_1) \rrbracket_{M,g,X}^c = 1$ and $\llbracket P_1(a_2) \rrbracket_{M,g,X}^c = 0$,

1. **No Reversal (NR:)** There is no $X' \in \mathcal{P}(D)$ such that $\llbracket P_1(a_2) \rrbracket_{M,g,X'}^c = 1$ and $\llbracket P(a_1) \rrbracket_{M,g,X'}^c = 0$.
2. **Upward difference (UD):** For all $X' \in \mathcal{P}(D)$, if $X \subseteq X'$, then there are some a_3, a_4 : $\llbracket P_1(a_3) \rrbracket_{M,g,X'}^c = 1$ and $\llbracket P_1(a_4) \rrbracket_{M,g,X'}^c = 0$.
3. **Downward difference (DD):** For all $X' \in \mathcal{P}(D)$, if $X' \subseteq X$ and $a_1, a_2 \in X'$, then there are some a_3, a_4 : $\llbracket P_1(a_3) \rrbracket_{M,g,X'}^c = 1$ and $\llbracket P_1(a_4) \rrbracket_{M,g,X'}^c = 0$.

6.5.2.2 Absolute/Non-Scalar Adjectives

1. **Absolute Adjective Axiom (AAA):** For all $Q_1 \in AA \cup NS$, all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$ and $a_1 \in X$, $\llbracket Q_1(a_1) \rrbracket_{M,g,X}^c = 1$ iff $\llbracket Q_1(a_1) \rrbracket_{M,g,D}^c = 1$.

6.5.3 Axioms governing \sim

6.5.3.1 Axioms characterizing indifference

For all $P \in SA \cup NS$, all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$,

1. **Reflexivity (R):** For all $a_1 \in X$, $a_1 \sim_{P_1}^X a_1$.
2. **Tolerant No Skipping (T-NS):** For all $a_1, a_2 \in X$, if $a_1 \sim_{P_1}^X a_2$ and there is some $a_3 \in X$ such that $a_1 \geq_{P_1}^t a_3 \geq_{P_1}^t a_2$, then $a_1 \sim_{P_1}^X a_3$.
3. **Strict No Skipping (S-NS):** For all $a_1, a_2 \in X$, if $a_1 \sim_{P_1}^X a_2$ and there is some $a_3 \in X$ such that $a_1 \geq_{P_1}^s a_3 \geq_{P_1}^s a_2$, then $a_3 \sim_{P_1}^X a_2$.
4. **Granularity (G):** For all $a_1, a_2 \in X$, if $a_1 \sim_{P_1}^X a_2$, then for all $X' \in \mathcal{P}(D) : X \subseteq X'$, $a_1 \sim_{P_1}^{X'} a_2$.
5. **Contrast Preservation (CP):** For all $X' \in \mathcal{P}(D)$, and $a_1, a_2 \in X$, if $X \subset X'$ and $a_1 \not\sim_{P_1}^{X'} a_2$ and $a_1 \sim_{P_1}^{X'} a_2$, then $\exists a_3 \in X' - X : a_1 \not\sim_{P_1}^{X'} a_3$.
6. **Minimal Difference (MD):** For all $a_1, a_2 \in D$, if $\llbracket a_1 >_{P_1} a_2 \rrbracket_{M,g,X}^c = 1$, then $a_1 \not\sim_{P_1}^{\{x,y\}} a_2$.

6.5.3.2 Axioms characterizing total/partial asymmetries

1. **Symmetry (S):** For $P_1 \in RA \cup NS$, an interpretation $\llbracket \cdot \rrbracket_{M,g}$, and $a_1, a_2 \in D$, if $a_1 \sim_{P_1}^X a_2$, then $a_2 \sim_{P_1}^X a_1$.
2. **Total Axiom (TA):** For a total AA Q_1^T , an interpretation $\llbracket \cdot \rrbracket_{M,g}$, and $a_1, a_2 \in D$, if $\llbracket Q_1^T(a_1) \rrbracket_{M,g,D}^c = 1$ and $\llbracket Q_1^T(a_2) \rrbracket_{M,g,D}^c = 0$, then $a_2 \not\sim_{Q_1^T}^X a_1$, for all $X \in \mathcal{P}(D)$.
3. **Partial Axiom (PA):** For a partial AA Q_1^P , an interpretation $\llbracket \cdot \rrbracket_{M,g}$, $X \in \mathcal{P}(D)$, and $a_1, a_2 \in D$, if $\llbracket Q_1^P(a_1) \rrbracket_{M,g,D}^c = 1$ and $\llbracket Q_1^P(a_2) \rrbracket_{M,g,D}^c = 0$, then $a_1 \not\sim_{Q_1^P}^X a_2$, for all $X \in \mathcal{P}(D)$.

6.5.3.3 Conversational Principles

1. **Be Precise (BP):** For all $S_1 \in NS$, all interpretations $\llbracket \cdot \rrbracket_{M,g}$, all $X \in \mathcal{P}(D)$ and $a_1, a_2 \in X$, if $a_1 \sim_{S_1}^X a_2$ and there is some $a_3 \in X$ such that $a_1 \geq_{S_1}^c a_3 \geq_{S_1}^c a_2$, then $a_1 \sim_{S_1}^X a_3$.

6.5.4 Summary

The semantic and pragmatic analysis of (non)scalar adjectives is summarized in tables 6.1 and 6.2.

Axiom	RA	AA	NS
No Reversal (NR)	✓	(✓)	(✓)
Upward Difference (UD)	✓	(✓)	(✓)
Downward Difference (DD)	✓	(✓)	(✓)
Absolute Adjective Axiom (AAA)	×	✓	✓

Table 6.1: Axioms governing the semantic denotation of adjectives

6.6 Empirical Patterns

This section presents a list of the empirical patterns discussed in the dissertation that distinguish the four principle ‘scale structure’ classes of adjectives. The goal of this list is to be a handy catalogue of the differences that have been identified in the literature between RAs, AAs, and NSs. For each data point, I also indicate where in the thesis its discussion can be found. Finally, in what follows, by *non-scalar* adjectives, I mean *non-coerced non-scalar* adjectives.

Axiom	Relative	Total AA	Partial AA	Non-Scalar
Reflexivity (R)	✓	✓	✓	✓
Tolerant No Skipping (T-NS)	✓	✓	✓	✓
Strict No Skipping (S-NS)	✓	✓	✓	✓
Granularity (G)	✓	✓	✓	✓
Minimal Difference (MD)	✓	✓	✓	✓
Contrast Preservation (CP)	✓	✓	✓	✓
Symmetry (S)	✓	×	×	✓
Total Axiom (TA)	×	✓	×	×
Partial Axiom (PA)	×	×	✓	×
Be Precise (BP)	×	×	×	✓

Table 6.2: Axioms governing \sim with (Non)Scalar Adjectives

Chapter	Construction	Relative	Total AA	Partial AA	Non-Scalar
2	Comparative construction	✓	✓	✓	×
	Degree morphology	✓	✓	✓	×
	Definite description test	✓	×	×	×
	Productive <i>for</i> phrases	✓	×	×	×
4	Potentially vague <i>P</i>	✓	✓	×	×
	Pseudo-evaluative comparative	×	✓	×	×
	Potentially vague <i>not P</i>	✓	×	✓	×
	Evaluative comparative	×	×	✓	×
5	Accentuation test	×	✓	×	×
	Strong resultative test	×	✓	×	×
	Telic degree achievement	×	✓	×	×
	<i>Almost</i>	×	✓	×	×
	<i>Loosely speaking</i>	×	✓	×	×
	<i>Absolutely</i>	×	✓	×	×
	<i>Strictement</i>	×	✓	×	×
	Maximal <i>completely</i>	×	✓	×	×
	Maximal <i>tout</i>	×	✓	×	×
	Maximal <i>plumb</i>	×	✓	×	×
	Proportional modifiers	×	✓	×	×
	Existential <i>slightly</i>	×	×	✓	×
	<i>Strictly speaking</i>	×	×	✓	×

Table 6.3: Scale Structure Distinctions

CHAPTER 7

Comparison with Other Approaches

7.1 Introduction

This chapter presents a comparison between the theory developed in this work and other current approaches to scale structure distinctions. In particular, I compare the predictions made by my account with those made by three influential proposals within the degree semantics framework: Kennedy (2007), Kennedy and McNally (2005), and Rotstein and Winter (2004). I present the main lines of each of the degree analyses and discuss how they could be extended to account for the full range of data discussed in chapters 2-5. However, I argue that, while the analyses within degree semantics can all be appropriately extended to capture the context-sensitivity, vagueness, and scale-structure patterns that we have seen, in all cases, the analysis proposed in chapters 2 and 4 accounts for the same data in a much simpler and more parsimonious way. I therefore conclude that the account given in this dissertation should be preferred over competing approaches within degree semantics.

The chapter is laid out as follows: firstly, in section 7.2, I summarize the main empirical patterns that a theory of adjectival scale structure distinctions aims to account for. Then, in section 7.3, I present a brief introduction to the degree semantics framework and discuss certain basic ontological and grammatical assumptions that are shared across all analyses within this approach. I then compare these assumptions to the corresponding ontological and grammatical assumptions shared by delineation frameworks. In this section, I will focus primarily on the treatment of relative adjectives like *tall* and *short* and the treatment of the scalar/non-scalar distinction. Finally, in sections 7.4, 7.5.1, and 7.5.2, I discuss the more specific proposals of Kennedy (2007), Kennedy and McNally (2005), and Rotstein and Winter (2004), respectively. I compare the analysis of the

absolute/relative distinction given in each of these papers to the analysis given in this dissertation.

7.2 Review of the Core Data

In this section, I give a review of the principle contrasts in the adjectival domain that were studied throughout the previous chapters of this work.

If we consider the set of adjectives as a whole, we saw in chapter two that we can make a first distinction between those adjectives, called *scalar* adjectives (i.e. *tall*, *wet*, *dry*...) that can naturally appear in a comparative or degree construction, and those that cannot, which are called *non-scalar* adjectives (*prime*).

- (1) a. John is taller than Mary.
b. This towel is wetter than that towel.
c. This towel is drier than that towel.
- (2) a. John is very tall.
b. This towel is very wet.
c. This towel is very dry.
- (3) a. ?This number is primer than that number.
b. ?This number is very prime.

Also in chapter 2, we saw that we can make a second distinction within the set of scalar adjectives between *relative adjectives* (like *tall*), which can shift their extension in the definite description test to distinguish between individuals at the middle of their associated scale, and *absolute adjectives* (like *dry* and *wet*), which can only be applied to individuals that are at (or close to) the endpoint of their associated scale.

- (4) Pass me the tall one.
Ok: even if neither/both are tall.

- (5) Pass me the dry/wet one.
 (But both/neither are dry/wet!)

Finally, we saw in chapters 4 and 5 that we can make a further distinction between two classes of AAs. We first identified the *total* AAs (like *dry*), which are associated with scales with maximal endpoints (6) and have potentially vague positive forms, and then we identified the *partial* AAs (like *wet*), which are associated with scales with minimal endpoints (7) and have potentially vague negative forms. We can further note that relative adjectives are associated with scales that have no endpoints and have both positive and negative forms that are potentially vague.

- (6) a. This towel is almost dry.
 b. *This towel is almost wet.
 c. *John is almost tall.
- (7) a. #This towel is slightly dry.
 b. This towel is slightly wet.
 (Existential interpretation)
 c. #John is slightly tall.

In sum, in chapters 2-5, I argued in favour of the following scale structure-based typology of adjectival predicates (figure 7.1):

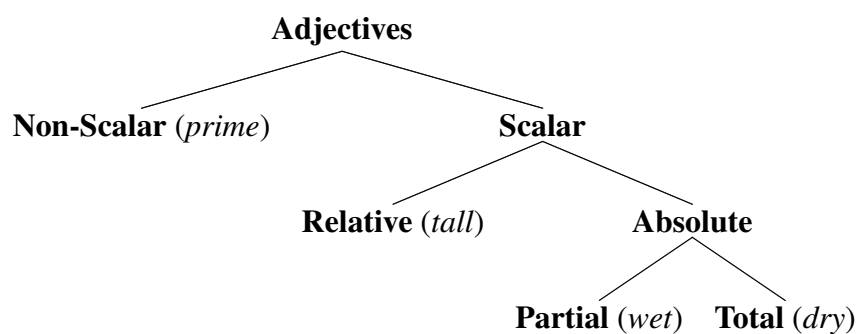


Figure 7.1: Adjectival Scale Structure Typology

In the next section, I introduce the broad framework in which the distinctions in figure 7.1 have been previously analyzed: degree semantics.

7.3 Basic Features of Degree Semantics

In this section, I present a basic overview of certain assumptions that are shared by analyses within the degree semantics (DegS) framework. Giving a complete yet succinct description of the framework and its various incarnations is difficult, since there are currently active debates in the literature about even the most fundamental parts of the theory. Fortunately, many of the empirical and conceptual concerns that drive these debates are orthogonal to the RA/AA/NS distinctions. Therefore, at the risk of glossing over certain finer distinctions in the literature, I will present a version of the DegS framework that, I believe, represents the most commonly used version in the field.

7.3.1 Degrees in the Ontology

The first main characterizing feature of a DegS analysis is the proposal that universe/domain of interpretation is sorted; that is, it contains (at least) two distinct kinds of objects: individuals and *degrees*¹. Furthermore, the individuals that are degrees are proposed to be ordered in a certain way; in particular, the set of degrees form a relational structure known as a *scale*.

Definition 7.3.1 *Scale*. A scale is a triple $\langle D_d, >, \phi \rangle$, where D is a set, $>$ is an ordering on D_d , and ϕ is a dimension (i.e. height, baldness etc.).

The nature and the cardinality of the set D_d in definition 7.3.1 is one of the more controversial aspects of the framework; however, in many analyses (ex. Kennedy and McNally (2005), Fox and Hackl (2006), among others) the scales associated with adjectival predicates are proposed to be dense; therefore, for these authors, the domain of degrees has uncountably many members. Although this is not logically necessary², the $>$ relations over degrees are generally assumed to be linear (irreflexive, transitive, and total) relations with an addition or multiplication operation on them (called *ratio* scales in measure-theoretic terms- von Stechow (1984); van Rooij (2011a)). That is to say, in addition to the axioms associated with an addition and multiplication operation,

¹Note that, while the vast majority of degree semantic analyses postulate degrees directly in the ontology, this ontological commitment is not necessary. For example, while Cresswell (1977) is generally categorized as a ‘degree’ analysis of gradable adjectives since he proposes that these predicates are transitive, he constructs (like I do) degrees from equivalence classes of individuals.

²See Solt (forthcoming) for a degree analysis of *most*, but that makes use of weaker orderings than ratio scales.

the following axioms governing the ontological relations between degree individuals are adopted (let D_d be the set of degrees in the domain):

(8) **Scalar Axioms**

- a. **Irreflexivity (I):** For all $d_1 \in D_d$, $d_1 \not> d_1$.
- b. **Transitivity (Tr):** For all $d_1, d_2, d_3 \in D_d$, if $d_1 > d_2$ and $d_2 > d_3$, then $d_1 > d_3$.
- c. **Totality (Tl):** For all distinct $d_1, d_2 \in D_d$, either $d_1 > d_2$ or $d_2 > d_1$.

In summary, I have highlighted two important features of degree analyses of gradable predicates:

1. Degrees are primitive individuals in the ontology.
2. The ontological relations between degrees satisfy certain strong ordering constraints (i.e. those that define ratio scales).

7.3.2 Scalarity as Argument Structure

The second feature that (to my knowledge) all analyses within degree semantics share is the claim that the scalarity of a predicate is determined by its syntactic/semantic argument structure. In particular, while a non-scalar adjective like *prime* is generally proposed to be an intransitive predicate (9), scalar predicates like *tall*, *wet*, and *dry* are taken to be transitive predicates: binary relations between individuals and degrees (10)³.

(9) $\llbracket \text{prime} \rrbracket = \lambda x. x \text{ is prime.}$

(10) a. $\llbracket \text{tall} \rrbracket = \lambda d \lambda x. x \text{ is tall to degree } d.$

b. $\llbracket \text{wet} \rrbracket = \lambda d \lambda x. x \text{ is wet to degree } d.$

c. $\llbracket \text{dry} \rrbracket = \lambda d \lambda x. x \text{ is dry to degree } d.$

³In some analyses, scalar adjectives denote measure functions: functions from individuals to degrees (Bartsch and Vennemann (1972), Kennedy (1997), a.o.), but even in this case, they are transitive predicates, unlike non-scalar adjectives.

When the positive form is used in a sentence, the degree argument of a scalar predicate is bound by an implicit operator (called either POS ((Bartsch and Vennemann 1972), Kennedy (1997), Kennedy (2007), a.o.) or EVAL (Rett (2008))) (11a). After the application of POS, the sentence is true just in case the degree to which the predicate holds of the subject (significantly) exceeds a contextually given standard degree, call it d_s .

- (11) John is tall.
- a. John is [POS [tall]].
 - b. $\llbracket \text{John is tall} \rrbracket = 1$ iff John's height $> d_s$.

7.3.3 Comparatives as Quantifiers

Having scalar predicates take a degree argument leads naturally to an analysis of comparatives as quantifiers that bind degrees in the syntax (cf. Seuren (1973), von Stechow (1984), Heim (1985), Heim (2000), Schwarzschild (2008) a.o.). Indeed, another key feature that is common to most (if not all) analyses in DegS is the treatment of the comparative morpheme *-er* as a quantificational determiner (in the generalized quantifier theory sense of the term⁴) that applies to sets of degrees as in (12).

- (12) $\llbracket \text{John is taller than Mary} \rrbracket = 1$ iff
- a. $\neg \text{ER}(\{d : \text{John is tall to degree } d\}, \{d : \text{Mary is tall to degree } d\}) = 1$ iff
 - b. $\{d : \text{Mary is tall to degree } d\} \subset \{d : \text{John is tall to degree } d\}$ iff
 - c. $\exists d : \text{John is } d \text{ tall and Mary is not } d \text{ tall.}$

One argument in favour of the quantifier approach to the comparative construction is that, when we look at sentences that have comparatives and other quantifiers in them, we sometimes see patterns that are reminiscent of DP quantifier scope interaction patterns. The idea is that, if comparatives like *taller than Mary* denote generalized quantifiers (over degrees) in the same way that

⁴I will take it for granted that the reader is familiar with basic generalized quantifier theory. For the classic introductory texts, see Barwise and Cooper (1981) and Keenan and Stavi (1986). For more recent surveys of GQ theory in both logic and natural language, see Peters and Westerstahl (2006) and Szabolcsi (2010).

DPs like *every student* denote generalized quantifiers (over objects), we expect degree quantifiers to be able to under go the syntactic rule of *quantifier raising* (QR) and create scope ambiguities, just like DPs do. And with certain kinds of quantifiers, this prediction is borne out. For example, Heim (2000) (based on observations by Stateva (2000)) shows that sentences with comparatives and the intensional predicate *to be required*, which is generally analyzed as a universal quantifier over possible worlds, are ambiguous in the way that we would expect if there was a scope interaction between these two quantificational expressions. For instance, the sentence in (13) can mean that, in every acceptable world, the paper is 15 pages long, i.e. it is not allowed to be longer than 15 pages. This reading of (13) corresponds to the logical form in (13a). The sentence in (13) could also be interpreted as meaning that, in every acceptable world, it is at least 15 pages long, but could possibly be longer. This reading corresponds to the logical form in (13b).

(13) (This draft is 10 pages.) *The paper is required to be exactly 5 pages longer than that.*

a. required [[exactly 5 pp -er than that] the paper be *t* long]

$$\forall w \in Acc : \max\{d : \text{long}_w(p,d)\} = 15pp$$

b. [exactly 5 pp -er than that] [required [the paper be *t* long]]

$$\max\{d : \forall w \in Acc : \text{long}_w(p,d)\} = 15pp$$

Heim (2000) (p. 224)

We find the same pattern with the possibility modal verb *allow* (14): on one reading (14a), (14) says that there is some acceptable world in which the paper is 15 pages long, and it is left open whether the paper could have been other lengths. On the second reading (14b), the sentence says that it is exactly 15 pages long in the acceptable worlds where it is longest, which means it is not allowed to be longer than 15 pages. Other examples of what seems to be raising of a degree quantifier over an intensional verb are given in Heim (2000), Stateva (2000) and many other subsequent papers addressing these observations.

(14) (This draft is 10 pages.) *The paper is allowed to be exactly 5 pages longer than that.*

a. allowed [[exactly 5 pp -er than that] the paper be *t* long]

$$\exists w \in Acc : \max\{d : \text{long}_w(p,d)\} = 15pp$$

b. [exactly 5 pp -er than that] [required [the paper be t long]]

$$\max\{d : \exists w \in Acc : \text{long}_w(p,d)\} = 15pp$$

Heim (2000) (p. 224)

In sum, in most theories within the degree semantics tradition, the comparative construction is analyzed as a quantifier that binds a syntactic degree argument, and we have seen that there are some empirical arguments from what appear to be quantifier scope ambiguities in favour of doing so.

7.3.3.1 Summary

In this subsection, I highlighted certain key features of the degree semantic framework. They are the following:

1. Degrees are primitive individuals in the ontology.
2. The ontological relations between degrees satisfy certain strong ordering constraints (i.e. those that define ratio scales).
3. The scalar/non-scalar distinction is an argument structure distinction.
4. Comparatives are quantifiers over degrees.

In the next section, I compare these proposals to their counterparts within delineation semantics.

7.3.4 Basic Features of Delineation Semantics

In this section, I summarize the basic features of the delineation semantic framework (DeIS), particularly as it is presented in this work. I then compare the empirical and theoretical predictions made by these frameworks.

Firstly, we can observe that DelS makes use of only one kind of object in the domains of interpretation: regular individuals, and, although, as discussed in chapter 2, it is possible to derive degree-like scales in this framework, the ‘degrees’ that make up the scales are simply equivalence classes of individuals.

Secondly, contrary to what is sometimes assumed about the DelS framework (see Kennedy (1997) and Moltmann (2009)), in this approach, no basic ontological relations are proposed to hold among individuals. This being said, we saw in chapter 2 that it was necessary to propose certain ‘coherence’ constraints on the definition of the $\llbracket \cdot \rrbracket$ function across comparison classes: for relative adjectives, we adopted van Benthem’s axiom set (*No Reversal, Upward Difference, Downward Difference*). We then saw that, from these constraints and the semantic/pragmatic analysis that we gave to the comparative construction, we could prove that the comparative relations (the $>_{ps}$ and \gg_{ps}) have the same kinds of properties that are stipulated to hold of the $>_s$ in degree semantics.

Thirdly, in DelS, the gradability of a predicate is a direct result of its context-sensitivity: the scales associated with adjectives are constructed from looking at how their denotations vary across comparison classes. Thus, as discussed in chapters 2 and 4, the scalar/non-scalar distinction can be reduced to a context-sensitive/context-independent distinction.

Finally, we can note that the comparative morpheme in DelS is also a quantifier. Instead of quantifying over degrees, however, it quantifies over comparison classes (as shown in (15)), which are analyzed as parameters of evaluation of adjectival predicates.

$$(15) \quad a >_p b \text{ iff there is some } X \subseteq D: a \in \llbracket P \rrbracket_X \text{ and } b \notin \llbracket P \rrbracket_X.$$

Thus, the main lines of degree approaches and delineation approaches can be summarized and compared as in the table 7.1.

Feature	Degree Semantics	Delineation Semantics
Degrees are:	primitive	derived
Constraints imposed on:	$>$	$\llbracket \cdot \rrbracket$
Gradability is due to:	argument structure	context-sensitivity
Comparatives are Qs over:	syntactic degree arguments	parametric comparison classes

Table 7.1: Degree Semantics vs Delineation Semantics

Are these accounts empirically and conceptually equivalent? Or are there some reasons to prefer DegS over DelS (or vice versa)? In the rest of this section, I go through each of the four points in table 7.1 where DegS and DelS differ, and discuss to what extent it is possible to tease the two frameworks apart.

7.3.4.1 Primitive Degrees vs Ontological Minimalism

The delineation framework has an immediate advantage over the degree framework: it does not require proposing new ontological primitives. DelS needs comparison classes, but these objects are simply subsets of the unsorted domain; therefore, the delineation framework can do without a very costly additional assumption that is integral to DegS. Of course, degrees have proven to be extremely useful in the analysis of a wide range of empirical phenomena in natural language, and, as discussed in Kennedy (1997), it is not immediately obvious how to analyze certain constructions involving gradable adjectives within a Klein-ian framework. For example, we can observe that, in most cases, comparing individuals using two adjectives of different ‘polarities’, like *tall* and *short* (16) is anomalous.

(16) Cross-polar anomalies

- a. ?John is shorter than Mary is tall.
- b. *John is taller than Mary is short.

However, under certain conditions (discussed in, for example, Faller (1998) and Büring (2007)), similar sentences can be felicitous (17). Although there exist analyses of these patterns in DegS (cf. the works cited above), these constructions are not accounted for in the system in Klein (1980).

(17) Cross-polar ‘nomalies’

- a. Unfortunately, the ladder was shorter than the house was high.
- b. My yacht is shorter than yours is wide.
- c. Your dinghy should be shorter than your boat is wide (otherwise you’ll bump into the bulkhead all the time).

Büring (2007), his (2)

In other words, if we just take Klein's original 1980 paper, we can see that there are certain gaps in the its empirical coverage. This dissertation, which modifies the traditional Klein-ian system, makes a contribution to filling some of these gaps. For example, Kennedy (2007) (p.41) raises the following challenge for degree-free approaches dealing with the absolute/relative distinction. He says,

In particular, an analysis that derives gradability from a general, non-scalar semantics for vague predicates must explain the empirical phenomena that have been the focus of this paper: the semantic properties of relative and absolute gradable adjectives in the positive form. While it may be difficult but not impossible to explain some of these features, I do not see how such an approach can account for the basic facts of the relative/absolute distinction in a non-stipulative way. . . the challenge for a non-degree-based analysis is to explain why only relative adjectives are vague in the positive form, while absolute adjectives have fixed positive and negative extensions, but remain fully gradable.

I believe that this challenge has been answered in this work, and, furthermore, there has been a recent rebirth of interest in DelS and its extension to deal with cross-polar (a)nomalies and other troublesome constructions (cf. van Rooij (2011a), Doetjes (2010), van Rooij (2011b), Doetjes et al. (2011), and Sassoon and Toledo (2011), among others). However, only time will tell whether DelS can be extended enough to achieve the same impressive level of empirical coverage as DegS. I therefore conclude that, to the extent that the DelS framework can be developed to analyze the same range of data that DegS can, DelS should be preferred to DegS on conceptual grounds. In other words, I agree with the following sentiments by von Stechow (1984) (p.50):

Perhaps a more concrete analysis in Klein's style (without the defect mentioned⁵)

⁵HB: von Stechow's main criticism of Klein (1980) is that he sees no way in which measure phrases (like the constructions in (i)) can be analyzed in his 1980 approach.

- (i) a. John is six inches taller than Mary.

should be upheld as long as possible. Only when this is impossible will we switch over to degrees.

7.3.4.2 Ratio Scale Axioms vs ‘Coherence’ Axioms

As discussed above, both frameworks make important use of constraints that contribute to yielding the scales associated with adjectival predicates. In DegS, these constraints are placed on the relations between degrees in the ontology (i.e. irreflexivity, transitivity, and totality). In DelS, the constraints are placed on the definition of the interpretation function to ensure that the application of a scalar predicate across comparison classes reasonably approximates how we apply these predicates in natural language. It is then shown that relations denoted by the comparative have the desired ordering properties (irreflexivity, transitivity, and almost-connectedness).

Again, on one level, the analyses look very similar: what is stipulated about the structure of the domain in DegS is stipulated about the interpretation of predicates in DelS. However, it is important to see that the degree semantics analysis makes stronger ontological claims than the delineation semantics analysis. The degree semanticist is committed to the existence not only of degrees but also of the relations between degrees in the ontology (either in the actual world or in the ontology of our cognitive models); however, all the delineation semanticist is committed to is the existence of certain guidelines for the application of particular kinds of context-sensitive linguistic expressions. Note that DelS is still consistent with the view that scales are part of the ontology. Indeed, if we were to find independent non-linguistic arguments for their existence, the presence of these abstract objects might help explain the metaphysical or cognitive source of these ‘coherence’ axioms. However, DelS is also consistent with the view that the scales that we associate with adjectival predicates are themselves products of how we use these linguistic expressions. I

-
- b. Ede is twice as fat as Angelika.
 - c. Ede is more tall than broad.
- von Stechow (1984) (p.50)

However, given that I propose adopting the axioms of van Benthem and deriving linear orders from strict weak orders in the style of Bale, the account presented in this dissertation does not suffer from this ‘defect’. The ‘comparison of deviation’ construction in (ic) might also be able to be analyzed within my framework in a similar way as in Bale (2008) or van Rooij (2011a).

therefore conclude again that, because it is consistent with a wider range of range of ontological claims, all other things being equal, DeIS should be preferred over DegS.

What kind of data could help us decide between these two hypotheses about the source of scale structure? This is a difficult question. Nevertheless, I would like to suggest that a promising line of enquiry involves the close examination of apparent mismatches between what we would intuitively expect the scale that encodes degrees of a property to be and what our linguistic tests for scale structure tell us that it is. For example, if we consider the relative predicate *tall*, we see such a mismatch. It is generally proposed that the comparative *taller than* lexicalizes the HEIGHT scale, and, naively, we would think that the height scale (and our mental representation of it) ought to have a bottom element, namely the property of having no height (i.e. measuring 0 millimetres). And yet, our linguistic tests tell us that *tall* is associated with a scale that has no endpoints, as shown in (18).

- (18) a. ??John is slightly tall.
(Only excessive interpretation)
b. ??John is completely tall.

Thus, we might wonder how it is that we learn that *tall* is associated with an open scale, given that the HEIGHT scale has a lower bound. A similar observation is made by Kennedy (2007) about the difference between relative adjectives *expensive* and *inexpensive*, and absolute adjectives *dirty* and *clean*. He says (pp.34-35),

For example, naive intuition suggests that the COST scale should have a minimal value representing complete lack of cost, just as the DIRT scale has a minimal value representing complete lack of dirt. However, the unacceptability of ??*slightly/partially expensive* and ??*perfectly/completely/absolutely inexpensive* (cf. *slightly/partially dirty* and *perfectly/completely/absolutely clean*) indicates that, as far as the gradable adjective pair *expensive/inexpensive* is concerned, this is not the case: the scale used by these adjectives to represent measures of cost does not have a minimal element.

Kennedy continues (p.35),

The structure of a scale is presumably determined mainly by the nature of the property that it is used to measure, but the different behaviour of eg. *expensive/inexpensive* vs. *dirty/clean* suggests that this aspect of linguistic representation may diverge from what naive intuitions suggest.

The question of why adjectival scale structure sometimes diverges from our intuitions does not arise for the delineation approach developed in this dissertation, like it does for the degree semantics analysis. In my analysis, by virtue of the fact that both *tall* and *expensive* pass the definite description test, they are predicted to be associated with open scales, and by virtue of the fact that *dirty* and *clean* fail the definite description test, they are predicted to be associated with scales with endpoints. Thus, an approach in which the scales that are grammatically important are constructed from the way that linguistic predicates are used can explain how these scales may diverge from our cognitive representations of concepts like height or cost. Of course, these remarks are only preliminary. It may turn out that the way we actually perceive height and cost is different from the ‘naive intuitions’ that I have just discussed, and furthermore, we might later find other reasons for preferring to have scales in the ontology. Generally speaking, the questions of what our mental representations of abstract concepts like height are and how they influence our language are both very broad and very difficult. As such, I leave additional consideration of these issues to future research.

7.3.4.3 Scalarity as Context-Sensitivity vs Argument Structure

A third way in which DegS and DelS differ is in where they locate the source of a predicate’s gradability. As mentioned, in DegS, the gradability of a predicate is determined by its argument structure: scalar adjectives have a degree argument, while non-scalar adjectives have no degree argument. In DelS, on the other hand, scalar adjectives have semantic (and, in the version presented here, pragmatic) denotations that can vary depending on a comparison class; while non-scalar adjectives have semantic and pragmatic denotations that are constant across CCs. What predictions do these two analyses make? As discussed at length in chapter 2, DelS makes the very strong prediction that gradability should coincide with context-sensitivity. Furthermore, I argued in that

chapter that this prediction is borne out in the data. Namely, I argued that the following generalization holds (cf. chapter 2's (28)):

(19) **Scalarity Generalization:**

An adjective is scalar iff it is type 2 context-sensitive.

By proposing that scalarity is a matter of argument structure, DegS also makes a prediction: unless we have reason to think otherwise, the null hypothesis is that degree arguments should bear the same kind of relationship to their predicate that other kinds of arguments (such as individual arguments) do. What kind of relationships hold between individual arguments and the predicate that selects them?

Although this is not an uncontroversial claim, in many frameworks, it is proposed that the relation between a syntactic argument and its selecting predicate is one of semantic arbitrariness. Empirical arguments for this claim come from observations about the existence of variability in the valency of a predicate that does not seem to be attributable to any semantic property. For example, as discussed in Fodor and Lepore (1998) (among others, cf. also Fillmore (1986) (among others) for discussions of optional transitivity), the sentences in (20) all describe roughly the same action involving the same participants (the eater and the thing being eaten), yet they all involve verbs with different argument structure patterns: *devour* is obligatorily transitive, and *eat* is optionally transitive, and *dine* is obligatorily intransitive. Thus, the argument structure of a predicate that takes individual arguments does seem to be predictable from its semantic properties.

- (20) a. John devoured *(the turkey).
b. John ate (the turkey).
c. John dined (*the turkey).

I have argued in this dissertation that whether or not an adjective is gradable (i.e. in the terminology of DegS, whether or not it takes a degree argument) is, in fact, predictable from its semantic properties, namely its context-sensitivity (cf. (19) above). Thus, if (19) is correct, and it is correct to think that the presence of a syntactic argument is not necessarily correlated with some semantic

property, it would seem that DegS makes wrong predictions when it comes to the kind of semantic dependencies that ‘degree arguments’ are subject to.

Of course this is hardly a knock down argument against degree arguments. For example, we might want to deny that syntactic argument structure is disconnected from semantics, and that, in fact, there is some (as yet undiscovered) difference in meaning between the predicates in (20) that can explain their different subcategorization frames. Or we might want to deny the claim that degree arguments should be expected to show the same argument-structure-based properties as individual arguments. I therefore conclude that, at this point, we can simply note that DeIS has an advantage over DegS in that, without any stipulation, the former framework accurately captures the relationship between context-sensitivity and gradability in natural language in a way that the latter framework does not.

7.3.4.4 Comparatives as Quantifiers over Degrees vs CCs

When comparing the DeIS and DegS accounts of the comparative construction, the first point to stress is that on one level, the two analyses are extremely similar, particularly when we consider what Schwarzschild (2008) calls the ‘A-NOT-A’ degree analysis of comparatives (ex. (21), where θ below is a threshold degree).

(21) A is more expensive than B is.

- a. There is some expense-threshold: A meets or exceeds it and B does not meet or exceed it.
- b. $\exists \theta \text{ expensive}(a, \theta) \wedge \neg \text{expensive}(b, \theta)$

Schwarzschild (2008) (p.2)

If we look at the delineation version of (21), we see that it also involves existential quantification and negation (22).

(22) A is more expensive than B is.

- a. $\exists X \subseteq D: a \in \llbracket \text{expensive} \rrbracket_X$ and $b \in \llbracket \neg \text{expensive} \rrbracket_X$.

On the other hand, although the quantificational force of the quantifiers is the same in DegS and DeIS, these theories make very different predictions with respect to possible scope interactions with other quantificational expressions in the syntax. In particular, since the quantification in DeIS is over values for a parameter, in the form given in Klein (1980) and in this dissertation, DeIS predicts that scope interactions between comparatives and other operators should not be possible; whereas, such interactions are predicted to occur in DegS. We saw in the previous section that comparatives do give rise to ambiguities when they appear in sentences with certain intensional verbs. Thus, we have a first argument in favour of DegS over DeIS.

However, it is well known that comparatives do not display exactly the scopal properties that we might expect if they were simply generalized quantifiers over degrees. As discussed in Heim (2000) and Kennedy (1997), with quantifier phrases that are not intensional verbs, comparatives obligatorily take the lowest possible scope. For example, the examples in (23) and (24) with both non-monotonic quantifier phrases (23) and degree quantifier phrases (24) lack the (b.) readings that are predicted within a straightforward DegS analysis.

(23) Exactly two girls are taller than 5 feet.

- a. $[\text{exactly two girls}]_1 [\text{-er than 5'}]_2 t_1$ are t_2 tall
 $|\{x: \text{girl}(x) \ \& \ \max\{d: \text{tall}(x,d)\} > 5'\}| = 2$
- b. $[\text{-er than 5'}]_2 [\text{exactly two girls}]_1 t_1$ are t_2 tall
 $\max\{d: |\{x: \text{girl}(x) \ \& \ \text{tall}(x,d)\}| = 2\} > 5'$ (**not available**)

Heim (2000) (p.221), cited from Kennedy (1997).

(24) John is 4' tall. *Every girl is exactly 1" taller than that.*

- a. $\forall x[\text{girl}(x) \rightarrow \max\{d: \text{tall}(x,d)\} = 4' + 1'']$
- b. $\max\{d: \forall x[\text{girl}(x) \rightarrow \text{tall}(x,d)]\} = 4' + 1''$ (**not available**)

Heim (2000) (p.222)

I therefore conclude that the question of whether comparatives are best analyzed as syntactic quantifiers is still open.

Fortunately, it is easy for the delineation semanticist to be agnostic as to whether we need comparatives to bind variables *in the syntax*: there is a straightforward way in which we could incorporate Heim’s insights about the scopal nature of comparatives into the framework developed in the rest of the dissertation. Simply, following authors such as Stanley (2000), we could propose that, instead of being parameters, comparison classes are covert variables in the syntax and then the comparative morpheme could bind these CC variables⁶. I therefore conclude that DelS is still compatible with the view that comparatives can interact with other quantificational expressions in the way that true QPs do.

7.3.5 Summary

In this section, I compared the analyses of relative adjectives and the scalar/non-scalar distinction within DegS and DelS. In the rest of the chapter, I will examine three influential analyses of the AAs and the source of the absolute/relative distinction within the degree semantics framework, and I will compare the proposals made in these works to the one developed in this dissertation.

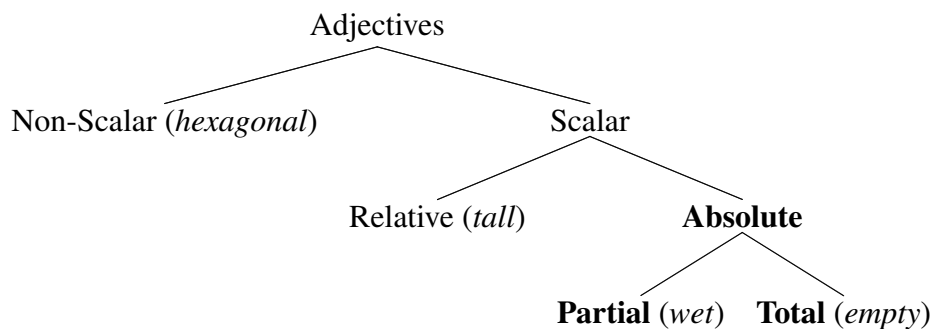


Figure 7.2: Adjectival Scale Structure Typology

⁶There is a parallel to be made here with the literature on indexical expressions: in Kaplan (1989)’s original theory, the context, which is necessary for determining the value of an indexical, is a parameter. However, it was later argued by Schlenker (2003) and Anand and Nevins (2004) that operators in the syntax could enter into scopal relations with the context, and so these authors propose to introduce context variables into the syntax.

7.4 Kennedy (2007)

For Kennedy (2007), relative adjectives take degree argument, and the positive form is proposed to have context-sensitive semantics: the implicit operator POS binds the RA's degree argument and returns the set of individuals who satisfy the predicate to a degree that is (significantly) higher than a contextually determined standard degree and, moreover, 'stands out' in the context (cf. Kennedy (2007) (pp.28-32) for a discussion of what it means for a degree to 'stand out' in context). This is consistent with the version of the degree semantics framework that I outlined above, and it straightforwardly explains why RAs are both type 1 context-sensitive (they have context-sensitive semantic denotations) and why these predicates are scalar (they take degree arguments). Furthermore, as is common in degree semantics, Kennedy (2007) stipulates that the scales associated with RAs are open. Although there is no formal link in this analysis between having a context-sensitive semantic denotation and being associated with an open scale, Kennedy hypothesizes that an RA's lack of a scalar endpoint make it such that there is no *natural transition* (in the sense of Williamson (1992)) which can be used as the cut-off point for its semantic denotation, and this lack of natural transition forces the RA to have a context-dependent denotation. He says (p.35),

More generally, there is nothing inherent to the structure of an open scale that results in natural transitions: open scales represent infinitely increasing or decreasing measures. As a result, there is nothing about the meaning of an open scale adjective alone that provides a basis for determining whether an object stands out relative to the kind of measure that it encodes. In order to make such a judgement, we need to invoke distributions over domains relative to which a standard can be established; i.e., we need a comparison class, which can be provided either by the context or by the adjective via domain restriction, as we have seen.

Kennedy's analysis of absolute adjectives is a little bit more complicated. He proposes that AAs also take degree arguments, and, likewise, that the context-sensitive operator POS binds these arguments. Thus, with these proposals, we can immediately account for why AAs are scalar (they take degree arguments) and how they can be type 2 context-sensitive (they have context-sensitive

semantic denotations)⁷. As briefly discussed in chapter 2, the challenge for this analysis is now to explain why AAs are not type 1 context-sensitive. Kennedy proposes that the key to the analysis of the contrast in context-sensitivity between RAs and AAs is the structure of the scales associated with these different kinds of predicates. He proposes that total AAs and partial AAs are associated with scales with top and bottom endpoints respectively, and he suggests that the presence of a scalar endpoint allows for the natural transitions that are necessary for a context-independent interpretation. This being said, Kennedy notes (p.36):

This can only be part of the story, however. Having a closed scale is a sufficient condition for absolute truth conditions, but not a necessary one. That is, even if we accept the reasoning just articulated, we don't rule out the possibility that something less than the maximal degree would be enough to stand out, or that something more than a (merely) non-minimal one would be required. That is, there is no inherent semantic incompatibility between (totally or partially) closed scales and a relative interpretation of the positive form, so we might expect that relative truth conditions are always an option for absolute adjectives, contrary to fact.

To eliminate the possibility of a relative use of an AA, Kennedy proposes the meta-grammatical principle called *Interpretative Economy* (25) (see also the discussion in chapter 2), which is designed to force the interpretation of an AA to be as close to the endpoint of the scale as possible.

(25) **Interpretative Economy:**

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

(Kennedy 2007) (p.36)

⁷Kennedy also invokes the phenomenon of imprecision to analyze examples like (i). Thus, in this proposal, type 2 context-sensitivity could be due either to the context-sensitive semantics of AAs or to the pragmatic phenomenon of imprecision.

- (i) The theater is empty tonight.
(When there are a couple of people there)

Thus, the different context-sensitivity patterns displayed by RAs and AAs are the result of different scale structures and the intervention of the principle in (25).

Finally, with respect to vagueness, Kennedy adopts a nuanced version of Williamson (1994)'s *epistemicism* theory for the appearance of the puzzling properties of vague language with RAs and, as discussed in chapter 4, attributes the context-dependent presence of fuzzy boundaries etc. with AAs to an independent phenomenon: *imprecision*.

To sum up, the analysis of scale structure distinctions in Kennedy (2007) has the following basic components:

1. The proposal that degrees and ordering relations between degrees are ontological primitives.
2. **Scalar/Non-Scalar Distinction:** The proposal that scalar predicates take degree arguments, while non-scalar predicates take only individual arguments.
3. **RA/AA Distinction:** The proposal that RAs and AAs are associated with scales with different properties to explain the scale structure differences between these predicates.
4. An analysis of RAs as having context-sensitive semantic denotations (the same is also proposed for AAs).
5. **RA/AA Distinction:** A pragmatic principle such as *Interpretative Economy* to explain the contrast in context-sensitivity between RAs and AAs.
6. An account of the vagueness of RAs (as a semantic phenomenon).
7. An independent theory of imprecision with AAs as a general pragmatic phenomenon (cf. Kennedy (2007) (p.43)).

7.4.1 Comparison with Delineation TCS Analysis

To review, the analysis of scale structure distinctions presented within Delineation TCS has the following basic components:

1. A unified theory of vagueness and imprecision in the adjectival domain.

2. The proposal that the interpretation of adjectival predicates is relativized to a comparison class (a subset of the domain).
3. The proposal that the semantic denotations of relative adjectives are subject to certain basic ‘coherence’ constraints.
4. **The RA/AA Distinction:** An analysis of RAs as having context-sensitive denotations and AAs as having non-context-sensitive denotations.
5. **The Scalar/Non-Scalar Distinction:** A pragmatic principle like *Be Precise* that enforces a precise use of non-coerced non-scalar adjectives.

In other words, in the delineation TCS analysis, all that we need to account for the differences between *prime*, *tall*, *dry*, and *wet* is 1) an independently necessary theory of vagueness/imprecision, 2) an analysis of the contrast in context-sensitivity between RAs and AAs as the result of differences in their lexical semantics, and 3) an analysis of the contrast in scalarity between AAs and NSs as the result of differences in the level of precision with they are usually used. Unlike in the DegS proposal discussed above, there is no need for primitive degrees and scales in the domain: degrees and scales of different types can be derived from different patterns of context-sensitivity and potential vagueness. Furthermore, there is no need for a pragmatic principle like *Interpretative Economy*: the fact that AAs are not type 1 context-sensitive is directly encoded in their lexical semantics. Finally, there is no need for null degree arguments in the syntax: the association of a predicate with a scale is straightforwardly predictable from its context-sensitivity properties. I therefore conclude that the approach developed in this dissertation constitutes a major simplification of the proposal given in Kennedy (2007) with no loss of empirical coverage, and, as such, I believe that it has significant conceptual advantages over this influential proposal.

7.5 Other Analyses of the RA/AA/NS Distinctions

In this section, I present a very brief discussion of two other very similar analyses of the RA/AA distinction in the literature: Kennedy and McNally (2005) and Rotstein and Winter (2004).

7.5.1 Kennedy and McNally (2005)

As in Kennedy (2007), in Kennedy and McNally (2005), relative adjectives are proposed to be context-sensitive, have degree arguments, and be associated with open scales. The analysis of the semantics and pragmatics of AAs in Kennedy and McNally (2005) is also very similar to the analysis of these predicates in Kennedy (2007). For example, these authors propose that AAs take degree arguments and that they associated with scales with maximal or minimal endpoints. However, there is one key difference in this analysis: to account for the observation that AAs are not type 1 context-sensitive, Kennedy and McNally (2005) adopt a similar proposal to the one in this dissertation, namely, that AAs have non-context-sensitive semantic denotations. Furthermore, like I do, they analyze type 2 context-sensitivity as imprecision. Thus, there is no need for a principle like *Interpretative Economy* (34) to explain the contrast in context-sensitivity between RAs and AAs.

7.5.1.1 Comparison with Delineation TCS Analysis

Although the analysis in Kennedy and McNally (2005) bears more similarities to the one presented in this work, I still argue that the delineation TCS account of the relative/absolute distinction should be preferred. Principally, although Kennedy and McNally's proposal accurately captures the context-sensitivity and scale structure data, it shows a certain amount of redundancy. Firstly, since it is necessary in this system to state independently in the lexical entry of a scalar adjective 1) whether or not it has a context-sensitive semantics and 2) what the structure of its scale it like, there is no formal account of the empirical dependencies in (26).

(26) Context-Sensitivity/Scale Structure Connection:

- a. Type 1 context-sensitive adjectives are associated with open scales.
- b. Type 2 context-sensitive adjectives are associated with closed scales.

Kennedy and McNally (2005) acknowledge the link between lack of context-sensitivity and the presence of endpoint and give a functional explanation in which the 'scale structure' parameter

influences the ‘context-sensitivity’ parameter. They say (p. 360),

The endpoints of the scale provide a fixed value as a potential standard, which in turn makes it possible to assign context-independent truth conditions to the predicate. . . The alternative-and the only option available to adjectives with open scales- is to compute the standard based on some context-dependent property of degrees. . . If we assume that interpretations that minimize context-dependence are in general preferred, then closed-scale adjectives should favor an absolute interpretation.

However, there are reasons to think that this explanation is less than satisfying. Firstly, stating the relationship between context-sensitivity and scale structure as a preference rather than as an absolute raises the question of why open scale AAs and closed scale RAs are absent rather than infrequent. Secondly, and perhaps more importantly, it is not clear that the principle that ‘interpretations that minimize context-dependence are preferred’ by natural languages is sound. Kennedy and McNally give no arguments in favour of the proposal that such a principle is active in the grammar, and, in fact, many (if not all) studies examining the role of context-sensitivity in language have concluded the opposite: context-dependence is ubiquitous natural language and has often even been argued to be beneficial for communication (see, for example, Barwise and Perry (1983), Sperber and Wilson (1985), Récanati (2004) among many others). On the other hand, in the delineation TCS analysis that I gave in the first two parts of the dissertation, there is no need to stipulate the scale structure of an adjective in its lexical entry, and the dependencies in (26) are consequences of the system (cf. theorems 5.3.2, 5.3.4, and 5.3.5).

Secondly, Kennedy and McNally (2005) miss another empirical generalization that I argued for in chapter 2: the observation that (non-coerced) non-scalar adjectives are not context-sensitive:

(27) **Context-Sensitivity/Scalarity Connection:**

Adjectives that are neither type 1 nor type 2 context-sensitive are not associated with scales.

Of course, both the context-independence and the non-gradability of a non-scalar adjective could always be stipulated in its lexical entry in Kennedy and McNally (2005)’s system; however, we

can recall that the context-sensitivity/scalarity connection in (27) is a theorem in the delineation TCS system (cf. theorem 4.6.30). Thus, while Kennedy and McNally (2005)'s system could be extended to capture empirical observations like (27), what are theorems in the account presented in this dissertation would need to be axioms in the degree account.

7.5.2 Rotstein and Winter (2004)

The proposal presented in Rotstein and Winter (2004) can be seen, in a certain sense, as a hybrid between Kennedy and McNally (2005) and Kennedy (2007). In many of the ways that concern us in this dissertation, Rotstein and Winter (2004) make the same proposals as Kennedy and McNally (2005): RAs have context-sensitive semantic denotations and are associated with open scales; total AAs have non-context-sensitive denotations and are associated with scales with top endpoints. Thus, Rotstein and Winter (2004) have the same account for the contrast in context-sensitivity as both Kennedy and McNally (2005) and this work. Furthermore, Rotstein and Winter (2004) analyze type 2 context-sensitivity as imprecision.

The main difference (relevant to this dissertation) between Rotstein and Winter (2004)'s analysis and the two other analyses that I have discussed concerns their treatment of partial AAs like *wet*, *bent*, and *dirty*. Like Kennedy and McNally (2005) and Kennedy (2007), they propose that partial AAs are associated with scales with bottom endpoints; however, like Kennedy (2007), they propose that partial adjectives have context-sensitive semantic denotations, like relative adjectives. However, unlike Kennedy (2007), Rotstein and Winter (2004) do not propose the existence of a principle like *Interpretative Economy*. Thus, their analysis predicts that the generalization in (26b) should not hold, at least for the partial subclass of AAs. In other words, they predict that partial AAs should show the type 1 context-sensitivity pattern. But, as shown in Syrett et al. (2010) (and discussed in the context of Rotstein and Winter (2004)'s proposal by Sassoon and Toledo (2011)), partial AAs fail the definite description test just as well as total AAs do, and Rotstein and Winter (2004)'s predictions are incorrect.

- (28) Give me the bumpy one.
(But neither/both is bumpy!)

I therefore conclude that the analysis presented in this dissertation should be preferred to Rotstein and Winter (2004) on empirical grounds.

7.6 Conclusion

In this chapter, I compared the account that I have given for the ‘scale structure’ distinctions in the adjectival domain with a series of rival accounts given within the degree semantics framework. I argued that my account of the semantic and pragmatic differences between RAs, AAs, and NSs is superior to some accounts on empirical grounds and superior to all accounts on conceptual grounds. I therefore conclude (contra Kennedy (2007)) that not only is it possible to develop a simple and empirically satisfying account of adjectival scale structure distinctions within delineation semantics, but also that such an account is both simpler and more empirically adequate than current accounts in degree semantics.

CHAPTER 8

Conclusion

In this work, I have presented a theory of the interaction between context-sensitivity, vagueness, and scalarity in the adjectival domain. 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 TCS). 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 dissertation.

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 (tables 8.1 and 8.2):

Pattern	Relative	Total	Partial	Non-Scalar
Type 1 CS	✓	×	×	×
Type 2 CS	(✓)	✓	✓	×

Table 8.1: Context-Sensitivity Patterns

Pattern	Relative	Total	Partial	Non-Scalar
P. vague \negP	✓	×	✓	×
P. vague P	✓	✓	×	×

Table 8.2: Potential Vagueness Patterns

I gave an analysis of these patterns within the delineation TCS 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 (figures 8.3 and 8.4).

Adjective	$>_p$: non-trivial SWO?	$>_p^t$: non-trivial SWO?	$>_p^s$: non-trivial SWO?
Relative	✓	×	×
Total Absolute	×	✓	×
Partial Absolute	×	×	✓
Non-Scalar	×	×	×
Coerced Non-Scalar	×	✓	✓

Table 8.3: Scalarity Patterns

Pattern	Relative	Total	Partial	Coerced Non-Scalar
Maximal Element?	×	✓	×	✓
Minimal Element?	×	×	✓	✓

Table 8.4: Absolute/Non-Scalar Scale Structure Patterns

Finally, in these past chapters, 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 a tolerance-driven comparison-class-based analysis of the meaning of gradable constituents has significant advantages over existing theories within degree semantics, and, thus, the success of my proposal provides an argument in favour of viewing context-sensitivity and general cognitive indifference relations as the driving forces behind scalarity in natural language.

BIBLIOGRAPHY

- Natasha Abner. Got me a plumb, gonna get me a scale. In *Proceedings of the 47th Meeting of the Chicago linguistics society*, page ??, Chicago, 2011. Chicago linguistics society.
- Sam Alxatib and Jeffery Pelletier. The psychology of vagueness: borderline cases and contradictions. *Mind & Language*, (forthcoming), 2010.
- Pranav Anand and Andrew Nevins. Shifty indexicals in changing contexts. In , editor, *Proceedings of SALT 14*, pages –. CLC Publications, 2004.
- S Armstrong, L Gleitman, and H Gleitman. What some concepts might not be. *Cognition*, 13: 263–308, 1983.
- John Langshaw Austin. *How to do things with words*. Clarendon, Oxford, 1962.
- Alan Bale. A universal scale of comparison. *Linguistics and Philosophy*, 31:1–55, 2008.
- Alan Bale. Scales and comparison classes. *Natural Language Semantics*, 19:169–190, 2011.
- Yeoshua Bar-Hillel. Indexical expressions. *Mind*, 63:359–79, 1954.
- R Bartsch and Theo Vennemann. The grammar of relative adjectives and comparison. *Linguistische Berichte*, 20:19–32, 1972.
- John Barwise and Robin Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219, 1981.
- John Barwise and John Perry. *Situations and Attitudes*. MIT Press, Cambridge, 1983.
- John Beavers. Scalar complexity and the structure of events. In Johannes Dölling and Tatjana Heyde-Zybatow, editors, *Event Structures in Linguistic Form and Interpretation*. Mouton de Gruyter, Berlin, 2008.
- Manfred Bierwisch. The semantics of gradation. In Manfred Bierwisch and E Lang, editors, *Dimensional Adjectives*, pages 71–261. Springer, Berlin, 1989.

- Hans Boas. *A Constructional Approach to Resultatives*. CSLI Publications, Stanford, 2003.
- Daniel Büring. Cross-polar nomalies. In UConn, editor, *Proceedings of SALT17*, pages 1–16. University of Conneticut, 2007.
- Heather Burnett. Vagueness and plural predication. In Daniel Lassiter, editor, *Proceedings of the 2011 ESSLLI Student Session*, pages 34–45, 2011a.
- Heather Burnett. From quantification and intensification to slack regulation: Adjectival ALL. In *Paper presented at NELS 42*, pages –, 2011b.
- Heather Burnett. Vague determiner phrases and distributive predication. In Marija Slavkovic and Dan Lassiter, editors, *New Directions in logic, language and interaction*, page ?? Springer LNCS, Berlin, 2012a.
- Heather Burnett. A multi-valued delineation semantics for absolute adjectives. In Paul Égré and David Ripley, editors, *Proceedings of the workshop on the application of three-valued logics to natural language*, page ??, Opole, Poland, 2012b. ESSLLI 2012.
- Heather Burnett. The puzzle(s) of absolute adjectives. In Denis Paperno, editor, *UCLA Working Papers in Linguistics 16: Special issue on semantics*, pages 1–50. University of California, Los Angeles, 2012c.
- Elizaveta Bylinina. Functional standards. In Dan Lassiter, editor, *Proceedings of the 2011 ESSLLI Student session*, pages 1–14. ESSLLI, Ljubljana, 2011.
- Gennaro Chierchia and Sally McConnell-Ginet. *Meaning and grammar: an introduction to semantics*. MIT Press, Cambridge, 2000.
- Pablo Cobreros, Paul Égré, David Ripley, and Robert van Rooij. Tolerant, classical, strict. *Journal of Philosophical Logic*, (forthcoming), 2011a.
- Pablo Cobreros, Paul Égré, David Ripley, and Robert van Rooij. Tolerance and mixed consequence in the s'valuationist setting. *Studia Logica*, page ??, 2011b.

- Max Cresswell. The semantics of degree. In Barbara Partee, editor, *Montague Grammar*, pages 261–292. Academic Press, New York, 1977.
- DA Cruse. Antonyms and gradable complementaries. In D Kastovsky, editor, *Perspectiven der Lexikalischen Semantik: Beiträge zum Wuppertaler Semantikkolloquium vom 2-3, Dec. 1977*, pages 14–25. Bouvier, Bonn, 1980.
- DA Cruse. *Lexical Semantics*. Cambridge University Press, Cambridge, UK, 1986.
- Richard Dietz and Sebastiano Moruzzi. *Cuts and Clouds*. Oxford University Press, Oxford, 2010.
- Jenny Doetjes. Incommensurability. In Maria Aloni, H Bastiaanse, T Jager, and Katrin Schultz, editors, *Logic, Language, and Meaning: Proceedings of the 17th Amsterdam Colloquium*, pages 254–263, Berlin, 2010. Springer.
- Jenny Doetjes, C Constantinescu, and K Soucková. A neo-klein-ian approach to comparatives. In S Ito and E Cormanu, editors, *Proceedings of Semantics and Linguistic Theory 19*, page forthcoming. UMass, 2011.
- David Dowty. *Word Meaning and Montague Grammar*. Reidel, Dordrecht, 1979.
- Paul Égré and Denis Bonnay. Vagueness, uncertainty, and degrees of clarity. *synthese*, 154:–, 2010.
- Paul Égré and Nathan Klinedinst. Introduction. In Paul Égré and Nathan Klinedinst, editors, *Vagueness and Language Use*. Palgrave MacMillan, 2011.
- Matti Eklund. What vagueness consists in. *Philosophical Studies*, 125:27–60, 2005.
- Martina Faller. A vector space semantics for dimensional adjectives, 1998.
- Delia Fara. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, 28, 2000.
- Charles Fillmore. Pragmatically controlled null anaphora. In , editor, *Proceedings of the 12th meeting of the Berkeley Linguistics Society*, pages 97–107. Berkeley linguistics society, 1986.
- Kit Fine. Vagueness, truth, and logic. *Synthese*, 30:265–300, 1975.

- Jerry Fodor and Ernie Lepore. The emptiness of the lexicon. *Linguistic Inquiry*, 29:269–288, 1998.
- Rafaella Folli and Gillian Ramchand. Prepositions and results in Italian and English. In Henk Verkyul, Henriette de Swart, and Angelika van Hout, editors, *Perspectives on Aspect*, pages 81–105. Kluwer, Dordrecht, 2005.
- Francesca Foppolo and Francesca Panzeri. Do children know when their room counts as “clean”? In GLSA, editor, *Proceedings of NELS42*, pages –, Amherst, 2011. GLSA Publications.
- Danny Fox and Martin Hackl. The universal density of measurement. *Linguistics and Philosophy*, 29:537–586, 2006.
- Gottlob Frege. *Grundgesetze der Arithmetik (Band II)*. Verlag Hermann Pohle, Jena, 1904.
- Brendan Gillon. Ambiguity, generality, and indeterminacy: Tests and definitions. *synthese*, 85: 391–416, 1990.
- Brendan Gillon. Ambiguity, indeterminacy, deixis, and vagueness: Evidence and theory. In Stephen Davis and Brendan Gillon, editors, *Semantics: A reader*, pages 157–190. Oxford University Press, Oxford, 2004.
- Adele Goldberg and Ray Jackendoff. The English resultative as a family of constructions. *Language*, 80:532–567, 2004.
- G Green. Some observations on the syntax and semantics of instrumental verbs. In *Proceedings of the Chicago Linguistics Society*, pages 83–97. Chicago University Press, 1972.
- U Hahn and N Chater. Similarity and rules: distinct? exhaustive? empirically distinguishable? *Cognition*, 65:197–230, 1998.
- Jennifer Hay, Christopher Kennedy, and Beth Levin. Scalar structure underlies telicity in degree achievements. In *Proceedings of SALT IX*, pages 127–144, 1999.
- Irene Heim. Notes on comparatives and related matters. 1985.

- Irene Heim. Degree operators and scope. In *In Semantics and Linguistic Theory*, pages 40–64. CLC Publications, 2000.
- Irene Heim and Angelika Kratzer. *Semantics in Generative Grammar*. Cambridge: Blackwell, 1998.
- J Hobbes. Granularity. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 432–435, 1985.
- Matthew Husband. Severing scale structure from the adjective. *LSA 2011 Extended Abstracts*, 2011.
- Dominic Hyde. The prospects of a paraconsistent approach to vagueness. In Richard Dietz and Sebastiano Moruzzi, editors, *Cuts and Clouds*, pages 385–405. Oxford University Press, Oxford, 2010.
- Hans Kamp. Two theories about adjectives. In Edward Keenan, editor, *Formal Semantics of Natural Language*, pages –. Cambridge University Press, Cambridge, 1975.
- Hans Kamp. The paradox of the heap. In Uwe Mönnich, editor, *Aspects of philosophical logic*. Riedel, Dordrecht, 1981.
- Hans Kamp and Barbara Partee. Prototype theory and compositionality. *Cognition*, 57:129–191, 1995.
- Hans Kamp and Antje Rossdeutscher. DRS-construction and lexically driven inferences. *Theoretical Linguistics*, 20:165–235, 1994.
- David Kaplan. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes from Kaplan*, pages 481–563. Oxford University Press, Oxford, 1989.
- Rosanna Keefe. *Theories of vagueness*. Cambridge University Press, Cambridge, 2000.
- Edward Keenan and Leonard Faltz. *Boolean Semantics for Natural Language*. Reidel, Dordrecht, 1985.

- Edward Keenan and Jonathan Stavi. A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9:253–326, 1986.
- Christopher Kennedy. *Projecting the Adjective*. PhD thesis, University of California, Santa Cruz, 1997.
- Christopher Kennedy. Vagueness and grammar: The study of relative and absolute gradable predicates. *Linguistics and Philosophy*, 30:1–45, 2007.
- Christopher Kennedy. Vagueness and comparison. In Paul Égré and Nathan Klinedinst, editors, *Vagueness and Language Use*, pages 1–24. Palgrave Press, 2011.
- Christopher Kennedy and Beth Levin. Measures of change: the adjectival core of degree achievements. In Louise McNally and Christopher Kennedy, editors, *Adjectives and Adverbs: Syntax, Semantics, and Discourse*, pages 156–182. Oxford University Press, Oxford, 2008.
- Christopher Kennedy and Louise McNally. Scale structure and the semantic typology of gradable predicates. *Language*, 81:345–381, 2005.
- Ewan Klein. A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4: 1–45, 1980.
- Ewan Klein. Comparatives. In Arnim von Stechow and Dieter Wunderlich, editors, *Semantics: An International Handbook of Contemporary Research*, pages 673–691. de Gruyter, Berlin, 1991.
- Angelika Kratzer. Telicity and the meaning of objective case. In Jacqueline Guéron and Jacqueline Lecarme, editors, *The syntax of time*, pages 398–423. MIT Press, Cambridge, 2004.
- Manfred Krifka. Nominal reference, temporal constitution and thematic relations. In Anna Szabolcsi and Ivan Sag, editors, *Lexical Matters*, pages 29–53. CSLI Publications, Stanford, 1989.
- Manfred Krifka. The origins of telicity. In Susan Rothstein, editor, *Events and Grammar*, pages 197–235. Kluwer, Dordrecht, 1998.

- Manfred Krifka. Negated antonyms: creating and filling the gap. In Uli Sauerland and Penka Stateva, editors, *Presupposition and implicature in compositional semantics*, pages 163–177. 2007.
- A Kyburg and M Morreau. Fitting words: Vague language in context. *Linguistics and Philosophy*, 23:577–597, 2000.
- George Lakoff. *Women, Fire and Dangerous Things*. University of Chicago Press, Chicago, 1987.
- Richard Larson. Scope and comparatives. *Linguistics and Philosophy*, 11:1–26, 1988.
- Peter Lasersohn. Pragmatic halos. *Linguistics and Philosophy*, 75:522–571, 1999.
- Daniel Lassiter. *Measurement and Modality: The scalar basis of modal semantics*. PhD thesis, New York University, 2011.
- David Lewis. Score-keeping in a language game. *Journal of Philosophical Logic*, 8:339–359, 1979.
- Godehard Link. The logical analysis of plurals and mass nouns: A lattice-theoretic approach. In Rainer Bauerle, Cristophe Schwartz, and Arnim von Stechow, editors, *Meaning, Use and the interpretation of language*, pages 302–322. Mouton de Gruyter, The Hague, 1983.
- R Duncan Luce. Semi-orders and a theory of utility discrimination. *Econometrica*, 24:178–191, 1956.
- Robert Martin. Analyse sémantique du mot ‘peu’. *Langue française*, 4:75–87, 1969.
- Louise McNally. The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In Rick Nouwen, Robert van Rooij, and Uli Sauerland, editors, *Vagueness in Communication*, pages 151–168. Springer, 2011.
- Friederike Moltmann. *Parts and Wholes in Semantics*. Oxford University Press, Oxford, 1997.
- Friederike Moltmann. Part structures in situations: The semantics of individual and whole. *Linguistics and Philosophy*, 28:599–641, 2005.

- Friederike Moltmann. Degree structure as trope structure. *Linguistics and Philosophy*, 31:51–94, 2009.
- Richard Montague. Pragmatics. In R Kibansky, editor, *Contemporary Philosophy-la philosophie contemporaine*, pages 102–122. La Nuova Italia Editrice, Florence, 1968.
- C Morris. *Foundations of the theory of signs*. Chicago University Press, Chicago, 1938.
- Marcin Moryzcki. Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics*, 19:39–86, 2011.
- Villem Novak, Irina Perfilieva, and Jiri Mockor. *Mathematical Principles of Fuzzy Logic*. Springer, 1999.
- Andrew Ortony, Richard Vonduska, Mark Foss, and Lawrence Jones. Saliency, similes, and the asymmetry of similarity. *Journal of Memory and Language*, 24:569–594, 1985.
- Daniel Osherson and Edward Smith. On prototypicality and vagueness. *Cognition*, 64:189–206, 1997.
- Charles Sanders Peirce. Vague. In JM Baldwin, editor, *Dictionary of Philosophy and Psychology*, page 748. Macmillan, New York, 1901.
- Stanley Peters and Dag Westerstaahl. *Quantifiers in Language and Logic*. Oxford University Press, Oxford, 2006.
- Manfred Pinkal. *Logic and Lexicon*. Kluwer Academic Publishers, Dordrecht, 1995.
- J. Pogonowski. *Tolerance Spaces with Applications in Linguistics*. Poznan University Press, Poznan, 1981.
- Graham Priest. Logic of paradox. *Journal of Philosophical Logic*, 8:219–241, 1979.
- François Récanati. *Literal Meaning*. Cambridge University Press, Cambridge, 2004.
- François Récanati. *Truth-Conditional Pragmatics*. Oxford University Press, Oxford, 2010.

- Jessica Rett. *Degree modification in natural language*. PhD thesis, Rutgers University, 2008.
- Jessica Rett. Similatives and the degree arguments of verbs. In , editor, *Manuscript*, pages 1–36. UCLA, Los Angeles, 2012.
- David Ripley. Contradictions at the borders. In Rick Nouwen, , Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, editors, *Vagueness in Communication*, page forthcoming. Springer, 2011.
- Eleanor Rosch. Natural categories. *Cognitive Psychology*, 4:328–350, 1973.
- Elenor Rosch. Principles of categorization. In Elenor Rosch and B Loyd, editors, *Cognition and categorization*, pages –. Erlbaum, Hillsdale, 1978.
- Susan Rothstein. *Structuring Events*. Blackwell, Oxford, 2004.
- Carmen Rotstein and Yoad Winter. Total vs partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics*, 12:259–288, 2004.
- Bertrand Russell. Vagueness. *Australasian Journal of Philosophy*, 1:84–92, 1923.
- Edward Sapir. Grading. A study in semantics. *Philosophy of Science*, 11:93–116, 1944.
- Galit Sassoon. Measurement theory in linguistics. *Synthese*, 174:151–180, 2010.
- Galit Sassoon and Asaf Toledo. Absolute and relative adjectives and their comparison classes. In , editor, *Manuscript*, pages 1–43. ILLC and Utrecht University, Amsterdam, 2011.
- Uli Sauerland and Pentka Stateva. Scalar vs epistemic vagueness. In *Proceedings of Semantics and Linguistic Theory 17*, Cornell, 2007. CLC Publications.
- Philippe Schlenker. A plea for monsters. *Linguistics and philosophy*, 23:29–120, 2003.
- Roger Schwarzschild. The semantics of comparatives and other degree constructions. *Language and Linguistics Compass*, 22:308–331, 2008.
- Pieter Seuren. The comparative. In F Keifer and Nicolas Ruwet, editors, *Generative Grammar in Europe*, pages 528–564. Riedel, 1973.

- Muffy Siegel. Measure adjectives in Montague grammar. In S Davis and M Mithun, editors, *Linguistics, philosophy, and Montague grammar*, pages –. University of Texas Press, Austin, 1979.
- Nicolas Smith. *Vagueness and degrees of truth*. Oxford University Press, Oxford, 2008.
- Scott Soames. *Understanding Truth*. Oxford University Press, New York, 1999.
- Stephanie Solt. Comparison to arbitrary standards. In *Sinn und Bedeutung 16*, Utrecht University, 2011.
- Stephanie Solt. On quantification and measurement: the case of most vs more than half. *Language*, forthcoming.
- Dan Sperber and Deirdre Wilson. Loose talk. *Proceedings of the Aristotelian Society*, 86:153–171, 1985.
- Jason Stanley. Context and logical form. *Linguistics and Philosophy*, 23:391–434, 2000.
- Penka Stateva. In defense of the movement theory of superlatives. In R Daly and A Riehl, editors, *Proceedings of ESCOL*, pages 219–226, Cornell, 2000. Cornell linguistics club.
- Kristen Syrett, Chris Kennedy, and Jeff Lidz. Meaning and context in children’s understanding of gradable adjectives. *Journal of Semantics*, 27:1–35, 2010.
- Anna Szabolcsi. *Quantification*. Cambridge University Press, Cambridge, 2010.
- Jamie Tappenden. The liar and sorites paradoxes: towards a unified treatment. *Journal of Philosophy*, 90:551–577, 1993.
- Amos Tversky. Features of similarity. *Psychological Review*, 84:327–352, 1977.
- Amos Tversky and I Gati. Studies of similarity. pages 79–98, 1978.
- Peter Unger. *Ignorance*. Clarendon Press, Oxford, 1975.
- Johan van Benthem. Later than late: On the logical origin of the temporal order. *Pacific Philosophical Quarterly*, 63:193–203, 1982.

- Johan van Benthem. *The logic of time*. Reidel, Dordrecht, 1990.
- Kees van Deemter. The sorites fallacy and the context-dependence of vague predicates. In Makoto Kanazawa, Christopher Pinon, and Henriette de Swart, editors, *Quantifiers, Deduction, and context*, pages 59–86. CSLI Publications, Stanford, 1995.
- Robert van Rooij. Vagueness, tolerance and non-transitive entailment. 2010.
- Robert van Rooij. Measurement and interadjective comparisons. *Journal of Semantics*, 28:335–358, 2011a.
- Robert van Rooij. Implicit vs explicit comparatives. In Paul Égré and Nathan Klinedinst, editors, *Vagueness and Language Use*, pages –. Palgrave Macmillan, 2011b.
- Robert van Rooij. Vagueness and linguistics. In G Ronzitti, editor, *The vagueness handbook*, page forthcoming. Springer, Dordrecht, 2011c.
- Arnim von Stechow. Comparing semantic theories of comparison. *Journal of semantics*, 3:1–77, 1984.
- Stephen Wechsler. Resultatives under the ‘event-argument homomorphism’ model of telicity. In N. Erteschik-Shir and T. Rapoport, editors, *The syntax of aspect*, pages 255–273. Oxford University Press, Oxford, 2005a.
- Stephen Wechsler. Weighing in on scales: A reply to Goldberg and Jackendoff. *Language*, 81: 465–473, 2005b.
- Timothy Williamson. Vagueness and ignorance. *Proceedings of the Aristotelian Society*, 66:145–162, 1992.
- Timothy Williamson. *Vagueness*. Routledge, London, 1994.
- Yoad Winter. *Flexibility Principles in Boolean Semantics*. MIT Press, Cambridge, 2001.
- Walt Wolfram. The grammar of rural and ethnic varieties in the southeast. In B Kortmann and E Schneider, editors, *Handbook of Varieties of English*, pages 74–94. Mouton de Gruyter, Berlin, 2004.

Crispin Wright. On the coherence of vague predicates. *Synthese*, 30:325–365, 1975.

Y Yoon. Total and partial predicates and the weak and strong interpretations. *Natural Language Semantics*, 4:217–236, 1996.