Intuition and Semantic Explanation

by

Charlie Ho Kin Siu

B.A. University of Hong Kong, 2006

M.A. Texas A&M University, 2008

M.A. Brown University, 2012

A Dissertation submitted to the Department of Philosophy

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Brown University

May 2018

This dissertation by Charlie Ho Kin Siu is accepted in its present form

by the Department of Philosophy as satisfying the

dissertation requirement for the degree of Doctor of Philosophy.

Date__________________ __________________________

Richard G. Heck, Advisor

Recommended to the Graduate Council

Date__________________ __________________________

Bernhard Nickel, Reader

Date__________________ __________________________

Pauline Jacobson, Reader

Approved by the Graduate Council

Date__________________ __________________________

Andrew G. Campbell,

Dean of the Graduate School

Thanks to everyone who has shown me goodwill

Abstract of "Intuition and Semantic Explanation", by Charlie Ho Kin Siu, Ph.D., Brown University, May 2018

This dissertation consists of three independent but related papers on the boundary between semantics and pragmatics. The overarching theme is that it is most fruitful to study how semantics interacts with pragmatics by identifying linguistic phenomena that call out for explanation, and by comparing the theoretical virtues of the semantics-pragmatics packages that explain or predict those phenomena. Contrary to what is often said about their role in semantic theorizing, ordinary intuitions are not the targets of semantic explanations, because they only play the role of bringing to light or justifying linguistic phenomena.

Chapter 1 argues that the imprecise interpretations of maximal standard absolute adjectives, such as 'clean' and 'certain', are the adjectives' semantic contents. My argument is based on the phenomenon that the result of embedding Rotstein and Winter's 'Both towels are clean, but the red one is cleaner than the blue one' inside the belief context 'Mary believes that' has three readings. I argue that my semantic account is preferable to the extant accounts because only it can deliver all three readings.

Chapter 2 argues that Jason Stanley's binding assumption is faced with two well-founded overgeneration worries, and that a feasible response to those worries is to draw from some parallels between quantifier domain restriction and adjectival domain restriction, and to adopt the variable-free approach to binding, which is compatible with Stanley's grammatical approach to quantifier domain restriction. This chapter as a whole illustrates that what makes Stanley's grammatical approach better than the pragmatic approach isn't its ability to respect ordinary intuitions, but its better explanatory and predictive power.

Chapter 3 objects to Rothschild and Segal's arguments for their account of color adjectives, on which color adjectives are fully-fledged indexicals with minimal semantic constraints on their possible extensions. I argue that, instead of choosing

between Rothschild and Segal's account and the rival accounts they argue against, we can defuse Travis cases by clarifying the relation between ordinary intuitions, linguistic phenomena, and explanation and prediction.

Acknowledgement

I am very grateful to Bernhard Nickel for all his help. He has been tremendously generous with his time. Without his prompt and detailed responses to my drafts and endless emails, I couldn't have made any progress. I also thank him for his patience with me as I made numerous (serious) mistakes in my drafts, for forcing me to become a more charitable reader, and for stopping me from taking the wrong turns at various junctures. I am deeply indebted to Polly Jacobson, who taught me semantics (her semantics class was one of the most enjoyable classes I took at Brown). She has rescued my dissertation and has given me insightful comments on my chapters (The remaining mistakes are all mine). I am grateful to Richard for his support throughout graduate school and for giving me the freedom to pursue the topics in this dissertation.

I would also like to thank my teachers and my colleagues at Texas A&M University, where I studied for my MA. Chris Menzel taught me logic, and I am deeply grateful for all his help and encouragement. I was fortunate to meet such a kind and supportive group of people: the late Scott Austin, Erik Berquist, Max Cresswell (who taught me modal logic and treated us ice cream), Mark Dondero, Michael Lebuffe, Muhammad Haris (who helped me numerous times), Dorothy Houston, Michael Long, the late Hugh McCann, Maggie McClean, Michael Oviedo, Philip Park, Marzena Plizga, Linda Radzik, Adriane Rini, Roger Sansom, Robin Smith, Cole Williams ... Mark Dondero deserves special mention. I am grateful for his friendship.

As someone who grew up in Hong Kong, I have learned a lot during my time in Rhode Island. I thank the coaches at the Ronald McDonald House of Providence Running Club and the friendly runners I met there for being so welcoming. I also thank Jim Dumont and his family for their friendship. I met Jim in one of the rides of the Narragansett Bay Wheelmen (another treasure of Rhode Island). He said he wanted a copy of my dissertation. This gave me an additional reason to finish.

Finally, I would like to thank my parents for their sacrifice, and for their being largely

supportive of the path I have chosen.

Contents

Chapter 1

'Absolute' Adjectives in Belief Contexts

1.1 Introduction

What do gradable adjectives such as 'flat', 'certain' and 'clean' mean? According to the absolutists (Unger 1978, Lasersohn 1999, Kennedy and McNally 2005, Kennedy 2007), they are true of an object just in case it has the properties expressed by them to the maximal degree. So, on their view, to be clean just is to be perfectly or absolutely clean. They are not deterred by the consequence that the assertions we make using these adjectives are hardly true of any object. They reassure us that since there are no important differences between saying that something is clean and saying that something is nearly clean, our assertions, while false, are close enough to true to be assertable.

According to the relativists, the assertions we make using these adjectives are very often true. They are true because to say that something is clean just is to say that it is clean enough relative to a certain factor. There are different views on what that factor is. On one view, it is the standard of precision salient in a discourse (Lewis 1979b). On others, it is the comparison class (Toledo and Sassoon 2011), or the rule according to which the adjective is used (McNally 2009).

In this paper, I am concerned with defending a relativist-friendly modification to Kennedy's account of maximal standard absolute adjectives, which is perhaps the most well-developed version of the absolutist view. I argue that to say that something is clean is to say that it is clean enough relative to discourse participants' standard of precision. I will focus on the first example below, given by Rotstein and Winter (2004), and the result of embedding it inside belief contexts:

(1) Both towels are clean, but the red one is cleaner than the blue one.

(2) Mary believes that both towels are clean but that the red one is cleaner than the blue one.

Let me explain the significance of these examples. It follows from Kennedy's absolutist view that the first sentence is a contradiction, because both towels must have the maximal degree of cleanness in order to be clean, but if so, the red one can't have a higher degree of cleanness than the blue one. Like any good absolutist, Kennedy need not be deterred by this result, because he could say that the sentence, while false, is still close enough to true for practical purposes and is hence assertable. This is where the second example comes in. While it can be read as an attribution of a contradictory belief to Mary, it need not be: Suppose we saw Mary, an animal lover, wipe Porky's face with the blue towel but save the red towel for herself. We can explain her behavior by uttering the sentence, because instead of conveying that she believes a contradiction, it naturally conveys that while she thinks that both towels are clean, she thinks that the red towel is cleaner, which is why she saves it for herself.

So one of the problems with the pragmatic, false-but-true-enough style, explanation for the assertability of (1) is that it doesn't explain why we can attribute coherent or even true beliefs to someone by embedding (1) in belief contexts. A related problem is that it doesn't explain the assertability of the result of embedding (1) under knowledge contexts: Since knowledge is factive, if (1) is a contradiction, the result of embedding it under 'know' ought to be unacceptable.

We observe a similar problem with the pragmatic explanation for the assertability of affirmative sentences containing number terms. For example, the pragmatic explanation holds that the following sentence is false, because the figures in the sentences are merely approximations of the maximally exact figures:

(3) Mozambique has a population of 23.4 million with a birth rate of 38/1000 and a death

rate of 15/1000 .

However, if Mary asks us the question (4-a), we can felicitously attribute a true belief to her by uttering (4-b):[1]

(4) a. Do you know that Mozambique has a population of 23.4 million with a birth rate of 38/1000 and a death rate of 15/1000?[2]

b. She knows that Mozambique has a population of 23.4 million with a birth rate of 38/1000 and a death rate of 15/1000.

So our topic has broader implications to how imprecision should be treated in general, which in turn has implications to the much-discussed semantics-pragmatics distinction.[3] But to keep our discussion sufficiently focused, we will focus on the imprecision of maximal standard absolute adjectives. I will discuss the extent to which my account can be extended to the imprecision of number terms in §1.5, but leave its implications to the debates on the semantics-pragmatics distinction for another occasion.

This rest of our discussion is structured as follows. In the next section (§1.2), I discuss why Kennedy is committed to the absolutist view, and examine some data that purport to support it. After that, I expand on the problem for the absolutist view, discuss how one might attempt to address it using the resources in Lasersohn (1999)'s theory of pragmatic halos and Toledo and Sassoon (2011)'s recent account of gradable adjectives, and explain why those attempts are likely to fail (§1.3). I will then defend my own account (§1.4), and discuss the extent to which it can be extended to number terms (§1.5). I conclude our discussion in the last section (§1.6).

[1] According to the pragmatic explanation, such as Lasersohn (1999)'s theory of pragmatic halos, the false sentence (3) is close enough to true because the precise numbers denoted by the number terms are good approximations of the exact numbers, and the sentence that is just like (3) except that its number terms are replaced by those that denote the exact numbers is literally true. This explanation, however, doesn't apply the result of embedding (3) in a knowledge context. Consider the result of embedding (3) under the context 'Mary knows that':

(i) Mary knows that Mozambique has a population of 23.4 million with a birth rate of 38/1000 and a death rate of 15/1000

Since Mary is in no position to know the exact numbers approximated by the number terms in the sentence, the sentence that is just like (i) except that the number terms are replaced by those that denote the exact numbers is still false. So the pragmatic explanation fails to extend to knowledge contexts.

[2] https://twitter.com/unicef_moz/status/329959791420112896

[3] Loose talk and imprecision are often regarded as pragmatic phenomena. See Bach (1994), Carston (2002), Sperber and Wilson (2005), Camp (2006).

1.2 Kennedy on Gradable Adjectives

According to Kennedy (2007), the meaning of a gradable adjective is a measure function that maps the objects in its domain onto a set of degrees. The set of degrees is naturally thought of as a scale, which is essentially a totally ordered set of points along a certain dimension, such as height and cleanness. For example, the meaning of 'tall' is a function that maps Porky and Esther to their heights on the tall scale. This meaning gives us a very straightforward truth condition for the comparative 'Porky is taller than Esther': The sentence is true just in case Porky's height is higher than Esther's height on the tall scale.

According to Kennedy and McNally (2005), there are four main types of adjectival meanings, each of which is characterized by its unique scale structure. As we can see below, their classification is supported by the distributions of the modifiers 'slightly' and 'perfectly'. We should also notice that maximal standard absolute adjectives, such as 'flat' and 'clean', belong to the second class below:

(5) a. Totally open-scales ( ) :
{??perfectly, ??slightly} tall/ short, expensive/ inexpensive
b. Partially-closed scales with maximal endpoints ( ] :
{✓perfectly, ??slightly} flat, clean, dry, certain
c. Partially-closed scales with minimal endpoints [ ) :
{??perfectly, ✓slightly} bumpy, dirty, wet, uncertain
d. Totally-closed scales [ ] :
{✓perfectly, ✓slightly} transparent/ opaque

No theory of gradable adjectives is complete without an account of the truth conditions of the positive form such as 'Porky is tall'. So crucial to Kennedy's theory is his account of how the four scale structures above determine the truth conditions of the positive form. There are two elements in his account. The first element is that each gradable adjective g is assigned a contextual threshold $s(g)$, such that a sentence such as 'Porky is g' is true in a context just in case Porky's degree as measured by g is at least as high as the contextual threshold. The following formula sums up this element:

(6) $pos(g) = \lambda x[g(x) \geq \mathbf{s}(g)]$, where g is an adjective's degree function, $pos(g)$ is the adjective's meaning in its positive form, $\mathbf{s}$ is a contextually given function that maps each gradable adjective to its contextual standard, and $\mathbf{s}(g)$ is the contextual standard for the adjective.

The second element is a principle that determines how the contextual standard is selected based on the adjective's scale structure. The idea behind the principle is that participants in a discourse ought to maximize the role of the conventional meanings of the words in the sentences they use to communicate with each other, and minimize the role of the context in computing the truth conditions of the sentences. Here is Kennedy's own statement of the principle:

(7) *The Principle of Interpretive Economy*
Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions (2007, p.36).

This is how this principle is supposed to determine the truth conditions of the positive form. Since the last three classes of adjectives above encode closed scales, participants in a discourse ought to make full use of those scale structures, selecting contextual standards that either coincide or depart minimally from the scales' endpoints. So, the thought goes, since an adjective such as 'clean' encodes a partially-closed scale with a maximal endpoint, its contextual standard just is the scale's maximal endpoint. But since 'dirty' encodes a partially-closed scale with a minimal endpoint, and since that minimal endpoint signifies the complete absence of dirt, its contextual standard can't be that minimal endpoint, or discourse participants arrive at the absurd interpretation that an object that's free of dirt counts as dirty. So the contextual standard ought to be minimally above that minimal point.[4] Putting two and two together, the contextual

[4]Kennedy doesn't explain why it follows from the economy principle that, while the contextual standard of a maximal standard absolute adjective (e.g. clean) coincides with the maximal endpoint of its scale, the contextual standard of a minimal standard absolute adjective (e.g. dirty) is minimally above the minimal endpoint of its scale.

standard for an adjective with totally closed scale such as 'opaque' can either be the scale's maximal endpoint or the point that is minimally above the scale's minimal endpoint. As to adjectives with totally open scales such as 'tall', since no points on those scales are special, their contextual standards can be anywhere on those scales, and they vary from context to context.[5] According to Kennedy, it is this contextual variability that sets this class of adjectives apart from the other three classes. He calls them 'relative adjectives' and the adjectives in the other three classes 'absolute adjectives'. The main characters of our discussion (e.g. 'clean', 'full', 'dry', 'certain') have a partially-closed scale with a maximal endpoint, so they are aptly called 'maximal standard absolute adjectives'.

Before we continue, let me make explicit why the principle is the source of Kennedy's absolutist commitment: Due to the principle, the contextual standard for 'clean' is always identical to the maximal endpoint of its scale. So it follows from the relation between its basic meaning and its positive form meaning (stated in (6)) that an object is clean in a context if and only if it is maximally clean.

Let's call Kennedy's overall account of gradable adjectives *the enhanced account*, and his account without the economy principle *the basic account*. Since the basic account is sufficiently well-established, we now focus on examining some crucial data that purport to support the enhanced account.

Syrett et al. (2006, 2009) have found that, when presented with a request like the following sentence and two objects with different lengths, which have been judged to be either both long or both not long in an independent task, their adult subjects systematically interpret it as a felicitous request for the longer of the two objects.

(8) Please give me the long one.

However, when the subjects are presented with a request like the following sentence and two partially-filled jars, they find the request infelicitous. They only accept the request when one jar is full and the other is about 2/3 full.

Here I hazard an explanation on his behalf. Without such an explanation, since 'dirty' and 'clean' share the same scale, the economy principle should make the wrong prediction that they have the same contextual standard.

[5]This reasoning rests on the assumption that the mid-point of the scale isn't different from other points.

(9) Please give me the full one.

Since ‘long’ has a totally open scale and ‘full’ has a partially-closed scale with a maximal endpoint, Kennedy’s economy principle predicts that only the contextual standard for ‘long’ can shift along the adjective’s scale. So the enhanced account predicts that, while the subjects shift the contextual standard for ‘long’ so that only one of the test items stands out, they don’t shift the contextual standard for ‘full’ in a similar fashion.

While Syrett et al.’s findings do seem to support the enhanced account, it is not clear we need a principle as strong as the economy principle to account for the findings. That is, it is not clear the reason why the adult subjects are unwilling to shift the contextual standard for ‘full’ is that the standard is always maximal. Syrett et al. (2009, p.27) acknowledge the possibility that, had the fuller jar been closer to full without being noticeably full, the subjects would have been willing to adopt a non-maximal contextual standard. Of course, this acknowledgement doesn’t undermine the enhanced account’s position that the contextual standards for maximal standard absolute adjectives are always maximal, because the subjects’ willingness to adopt a non-maximal contextual standard may be due to imprecision, which according to both Sryrett et al. and Kennedy is a pragmatic phenomenon. But it seems to suggest that the findings don’t conclusively show that the contextual standards for maximal standard absolute adjectives always coincide with the scales’ endpoints. So the findings don’t provide conclusive evidence for the economy principle.

Here is an alternative explanation for the findings that doesn’t depend on the principle. General principles of rational communication, such as the principle that discourse participants ought to make sure that their utterances are truthful and are as informative and as relevant as the situation demands, require that the discourse participants interpret maximal standard absolute adjectives at a high enough degree of precision, rather than the maximal degree of precision, which typically results in false utterances. A consideration in favor of this explanation is that general principles of rational communication can not only apply to maximal standard absolute adjectives, but also to number terms and non-gradable adjectives. For example, if I utter ‘we’ll meet at 3pm’ and ‘Italy is boot-shaped’, there is a presumption that ‘3pm’ and 'boot-shaped' ought to be interpreted at suitably high degrees of precision, so that my utterances end up true. So the general principles of rational communication are serious contenders against the economy principle.

We now turn to some consistency and entailment claims that purport to support the enhanced account. The followings are predicted by the account, because if the glass is full, it is maximally full and cannot be fuller, and if the countertop is less dry than the floor, then it isn't maximally dry and is hence not dry:

(10) #My glass is FULL, but it could be fuller (Kennedy 2007, 45a)

(11) a. The floor is drier than the countertop (Kennedy 2007, 50a) $\vDash$

b. The countertop is not dry. (Kennedy 2007, 50b)

However, these patterns don't seem to hold uniformly. Rotstein and Winter's example is a clear counterexample to both: The less clean towel in their example could be cleaner; despite the fact that the red towel is cleaner than the blue towel, the blue towel is clean. Other counterexamples include rice bowls, which are conventionally regarded to be full if they are filled roughly up to the rims but which can certainly be fuller, thanks to the stickiness of rice.[6][7]

The following entailment claim has also been said to lend support to the enhanced account:

(12) The table is not wet $\vDash$ The table is dry. (Kennedy 2007, 47b)

This is why the enhanced account predicts this entailment. 'Dry' and 'wet' share the same partially-closed scale, with the maximal endpoint representing the maximal degree of dryness, which amounts to the absence of wetness. So if the table is not wet, its degree of wetness must coincide with the maximal endpoint on the scale, which means that it is dry, and maximally so. However, our judgment about this entailment may vary with the object referred to in the example and the context. For example, Rotstein and Winter (2004) observe that: 'in some

[6] See also McNally (2009)'s example of wine glasses. Wine glasses are said to be full when they are filled up to the fill-line, but they can be fuller.

[7] One may attempt to block these counterexamples by placing focal stress on the adjectives in their positive forms, because, as Kennedy and Unger observe, it blocks the imprecise interpretations of the adjectives. However, since we can't antecedently assume that the blocked imprecise interpretations are pragmatic and not semantic, we can't restrict the data to sentences with focal stress placed on the adjectives in their positive forms.

contexts a moist towel may be considered neither wet nor dry'. Similarly, a rag we use for general cleaning may be considered neither dirty nor clean. So this entailment doesn't seem to hold generally.

Before we leave this section, let me forestall an attempt to dismiss our counterexamples to the consistency and entailment claims above. One may say that the reason why a full bowl of rice can be fuller, a clean towel can be cleaner than another clean towel, and a moist towel can be neither wet nor dry, is that the maximal standard absolute adjectives in their positive forms are given a relative interpretation that is characteristic of relative adjectives such as 'tall'. But this attempt doesn't work: If 'full', 'clean' and 'dry' do have a relative, open-scale, meaning on top of their absolute, closed-scale, meanings, they should figure in patterns that are characteristic of relative adjectives. For example, there should be cases where 'perfectly full/clean/dry' is infelicitous. But such cases don't seem to exist.

We have traced Kennedy's absolutist view to the economy principle, and examined the extent to which his enhanced account is supported by linguistic data. We now proceed to our main objection to the absolutist view.

1.3 The Challenge from Belief Contexts

As we discussed in the introduction, our main objection to the absolutist view is a response to the typical pragmatic, false-but-true-enough style, explanation for the assertability of sentences with maximal standard absolute adjectives. Our plan in this section is to expand on the objection. After that, I will argue that attempts to answer it by modifying Lasersohn's theory of pragmatic halos are likely to fail, and that, although Toledo and Sassoon's recent account of gradable adjectives is sensitive to the challenge Rotstein and Winter's example poses to the absolutist view, it doesn't yet provide a satisfactory analysis of the sentence on which our objection is based.

This is our objection to the absolutist view in a nutshell: The following belief sentence has three readings. The problem for that view is that it fails to deliver all of them.

(13) Mary believes that both towels are clean but that the red one is cleaner than the blue one.

The first reading, which is the only reading predicted by Kennedy's enhanced account, says that Mary believes a contradiction. The second reading, which can be delivered by Kennedy's basic account, says that Mary believes that both towels are clean up to a contextually given non-maximal degree but that the red one is cleaner.[8] We can bring out this reading of the positive form more clearly by using 'know' instead of 'believe' in our example (We omit the second conjunct for easier parsing):

(14) Mary not only believes but knows that both towels are clean.

Let me supply the context for this sentence. Suppose Mary and I are employees of a towel cleaning company, which specifies strict protocol on what counts as a clean towel. To keep our jobs, we know the protocol by heart. You are new to the company. So I recite the protocol for you: If a towel has been boiled and disinfected 5 times using the company's patented procedure, it counts as clean. If it is boiled and disinfected 10 times or more, it is perfectly clean, and we are entitled to charge the customers more for cleaning it. Mary has cleaned her red towel 6 times and her blue towel 5 times, and she knows it. It seems that I can report her information state by uttering (14).

On this reading, the contextual standard for the maximal standard absolute adjective in positive form is either in the common ground of the discourse participants, or can readily be added to it. Let's call this reading *the public standard reading*, or the wide scope reading, because it seems as if the maximal standard absolute adjective in positive form takes wide scope over the belief context. The first reading is a special case of this reading when the contextual standard is maximal. For example, if Mary has cleaned both towels 10 times, her belief that the red towel is cleaner than the blue towel must be a contradiction.

Let me make the third reading salient with a context like the following: The discourse

[8] In saying that the degree is contextually given, I don't suggest it is literally given by the context. It is plausible that discourse participants coordinate on that degree or on a set of possible values for that degree.

participants see Mary from afar and call her 'Mary' and the pig 'Porky', but Mary isn't one of the discourse participants. Mary wipes Porky's face with the blue towel, apparently with loving care, but she wipes her own face with the red towel. The speaker doesn't see the towels clearly enough to tell how clean they are. She utters the sentence in order to proffer an explanation for Mary's behavior.

Given the context above, the following continuation of the sentence is odd (let's assume that the speaker does have a towel that isn't noticeably dirty):

(15) #... my towel may not be so/that clean, so I am not going to wipe a pig's face with it, as much as an animal lover as I am.

Here is a possible explanation for its oddness. Since Mary isn't one of the discourse participants, the speaker is in no position to know how clean is clean for Mary. Further, since the speaker observes the towels from afar, she is in no position to know how clean the towels are either. The continuation (15) is odd, because 'so/ that clean' appears to make reference to a standard of cleanness that is either already in the common ground of the discourse participants or can readily be added to it, but no such standard is available given the context.

So the third reading is possible without there being a contextually salient standard of cleanness or a standard that can readily be made salient. It conveys that Mary believes that the towels are clean by her own standard but that the red one is cleaner. We call this reading *the private standard reading*, or the narrow scope reading, because it seems as if the maximal standard absolute adjective in positive form takes narrow scope under the belief context.

The distinction between the public standard reading and the private standard reading is related to a potential problem with Kennedy's basic account. However, since our primary target is the absolutist commitment of his enhanced account, I will mention what the problem is, but leave a more detailed discussion of it for future work. Let's recall from our discussion in the last section that crucial to Kennedy's basic account is his proposal of how the basic meanings of gradable adjectives are converted into their positive form meanings. He achieves the conversion by invoking a contextually given function, which determines the minimal degree it takes for an object to satisfy a gradable adjective in its positive form. As a memory aid, I repeat here the formula (6), which summarizes the core idea of the conversion:

(16) $pos(g) = \lambda x[g(x) \geq \boldsymbol{s}(g)]$, where g is an adjective's degree function, $pos(g)$ is the adjective's meaning in its positive form, $\boldsymbol{s}$ is a contextually given function that maps each gradable adjective to its contextual standard, and $\boldsymbol{s}(g)$ is the contextual standard for the adjective.

This is the potential problem. The context is supposed to supply a specific standard $\boldsymbol{s}(g)$ for the adjective g. But this is in conflict with the private standard reading, because that reading doesn't seem to depend on any specific contextually given standard. This issue concerns Kennedy's basic account, because the private standard reading is not only available for maximal standard absolute adjectives, but for all gradable adjectives more generally. For example, suppose we know that Mary loves competition and that she only races with people she thinks are tall. We saw from afar that Mary is having a race with a person, but we are unable to tell how tall that person is. It is quite natural to explain Mary's behavior by uttering that 'Mary believes that that person is tall', without meaning that Mary believes that that person's height is at least as high as a certain contextually salient standard.

In the following, we will set aside this broader issue with Kennedy's basic account and focus on the public standard and the private standard readings of maximal standard absolute adjectives. One fortuitous feature of our purposed analysis, which we will get to in §1.4, is that it can readily be applied to other kinds of gradable adjectives.

We are done presenting the three readings of the belief sentence on which our objection to the absolutist view is based. Let's now consider the extent to which one can answer it by drawing from Lasersohn's account of pragmatic halos and Toledo and Sassoon's recent account of gradable adjectives.

1.3.1 Lasersohn's theory of pragmatic halos

Lasersohn's theory of pragmatic halos is perhaps the most sophisticated pragmatic, false-but-true-enough style, explanation for the assertability of sentences with maximal standard absolute adjectives. As I said, our objection to the absolutist is crafted with this explanation in mind. We now consider how it fares against our objection.

The core idea of his theory is that each expression has both an actual denotation and a set of denotations that are of the same type as the actual denotation but differs from it in 'pragmatically ignorable' ways. The precise size of the halo, which correlates with what counts as pragmatically ignorable differences in meanings in a context, depend on various features of the context, such as the interests and the expectations of the interlocutors, and the standard of precision of their discourse. For example, the adjective 'clean' has both its actual denotation *clean* and the pragmatic halo {*clean*, *clean-*, *clean–*, ...}, which is a context-dependent set containing denotations that differ from the actual denotation to different degrees. The composition of a complex expression's regular denotation is carried out in the usual way. But the computation of a complex expression's pragmatic halo is analogous to the computation of focus values (Rooth 1985). For example, to obtain the pragmatic halo of an expression such as 'Porky is clean', we simply apply each halo value of 'is clean' to the halo value of 'Porky' (which is Porky himself), and collect the results in a set. If the pragmatic halo of 'clean' is {*clean*, *clean-*}, the resulting pragmatic halo of the sentence is {Porky is *clean*, Porky is *clean-*}. A sentence is said to be 'close enough to true' if one of its halo values is true. And it is literally true if its regular denotation is true.

This theory provides an elegant explanation for the assertability of Rotstein and Winter's example: Although its first conjunct is literally false according to the absolutist, since a halo value of 'clean' is true of the two towels, the first conjunct, and the sentence as a whole, are close enough to true in some context and are hence assertable.

However, it appears to run into trouble when we embed Rotstein and Winter's example in belief contexts. Since the theory preserves the regular meaning of Rotstein and Winter's example, which we should recall is a necessarily false proposition, it predicts that the resulting sentence is an attribution of a contradictory belief. So it doesn't make it any easier for the absolutist to get the readings on which the sentence is a attribution of coherent beliefs to Mary.

In order to get second reading, one might modify Lasersohn's theory and have the meaning of 'believe' be sensitive to the halo values of the embedded sentence. The idea is that the sentence 'Mary believe that the towels are clean but that the red one is cleaner than the blue one' is true in a context just in case Mary bears the belief relation to a non-regular halo proposition of the embedded sentence, instead of its necessarily false regular proposition. This allows for the possibility that Mary bears the belief relation to a true but imprecise content.

However, it is not clear how this strategy can be extended to deliver the third, private-standard, reading, because the pragmatic halo of the embedded sentence need not correspond to Mary's own standard of cleanness. For example, let's suppose you and I agree on the same standard for towels' cleanness. So, naturally, when you and I are the only discourse participants, the size of the pragmatic halo of 'clean' corresponds to our own standard of cleanness. Suppose we now see Mary from afar wiping a pig's face with a blue towel and hers with a red towel. Since we don't know whether she has the same standard for towels' cleanness as ours, we can't be certain that one of the halo values of 'clean' is suitable for attributing to her the belief that the towels are clean by her own standard: The halo of 'clean' would be too large if Mary's standard were higher than ours; too small if her standard were lower than ours. Since we can never be certain that the halo is of the right size, the modified account fails to explain why we can truthfully attribute to Mary the belief that the towels are clean by her own standard.

So unless there is a new halo manipulation strategy that yields the third, private-standard, reading, it seems safe to conclude that Lasersohn's theory doesn't provide the absolutist with a satisfactory response to our objection.

1.3.2 Granularity shifts and local absolutism

Toledo and Sassoon are well aware of the challenge Rotstein and Winter's example poses to the absolutist. As they show with the following examples, the absolutist commitment appears overly strong.

(17) a. This kitchen knife is clean (Cruse 1980).

b. This surgical instrument is clean (Cruse 1980).

(18) The gas tank is full, but you can still top it off. It's not completely full yet (2011, 12b).

According to their judgment, the standard of cleanness relevant to the interpretation of (17-a) is lower than that of (17-b); the sentence (18) is perfectly natural, which it can't be if to be full is to be completely full.

Their goal, as I understand it, is to weaken the absolutist commitment while keeping much of Kennedy's account of gradable adjectives intact — including the economy principle. Their account is of particular interest to us because it provides clear answers to several key questions about the proper treatment of maximal standard absolute adjectives: Should their imprecision be treated semantically? Can we achieve theoretical gain by having the interpretation of maximal standard absolute adjectives depend on comparison classes? Can some form of absolutism be salvaged from various counterexamples? Their account answers these questions in the positive.

Instead of presenting their account in full, we will focus on two crucial ideas in their account, and consider their potential in meeting our objection against the absolutist. The first idea concerns treating the imprecision of maximal standard absolute adjectives semantically. On their account, a gradable adjective's degree function varies with the contextually salient level of precision, and that the precision of an adjective's degree function can increase in a discourse. It is important to notice that, on their view, what varies with the contextually salient level of precision isn't the contextual standard of the adjective, but the very degree function it denotes.

Toledo and Sassoon argue convincingly that a gradable adjective can be associated with degree functions of different *granularities* or fineness of grain. Due to the limits of our perceptual power, we normally treat glasses of water which look indistinguishable to us as equally full, even though, with appropriate measurement tools and enough time, we are able to make finer distinctions among the same glasses. The relevant granularity can also be determined by purely pragmatic reasons. For example, a bowl that is filled with rice up to its rim is already full. A bowl with even more rice is usually considered full just the same. Toledo and Sassoon refer to this as the 'ceiling effect'.

They suggest that we can analyze Rotstein and Winter's example — which I repeat here as (19) — as expressing a literal truth by invoking two degree functions with different granularities.

(19) Both towels are clean, but the red one is cleaner than the blue one.

The thought is that, since a shift from a less fine-grained degree function to a more fine-grained one is a 'licensed discourse move', the first instance of 'clean' denotes a degree function that maps both towels to the same maximal degree, but the second instance of 'clean' denotes a more discriminating degree function that maps the towels to different non-maximal degrees.

However, their proposal predicts that the following contradiction is consistent:

(20) #Both towels are clean, but the red one is cleaner than the blue one and the blue one is not clean.

The reason is that, since the granularity of the clean function increases as we go from the first conjunct to the second conjunct, and since only an increase in granularity is a licensed discourse move, the second instance of 'clean' must denote different function than the first instance. This means that 'both towels are clean' and 'the blue one is not clean' can be true at the same time. So while the idea of granularity shift may explain the assertability of (19), it deprives us of the most straightforward explanation for the unacceptability of (20) (i.e. that it is a contradiction).

Let's now consider the second crucial idea in their account. It concerns folding comparison classes in the semantics of maximal standard absolute adjectives and using them to salvage the economy principle and some form of the absolutist commitment. They draw their inspiration from Bierwisch's judgment about the contrast between the relative adjective 'tall' and the adjective 'industrious':

(21) All the pupils at this school are tall.

(22) All the pupils at this school are industrious.

> In the interpretation of (21) other people must be taken into account, but to interpret (22) they need not be. Put differently, for some people to be tall there must be short people too, but for some to be industrious there do not need to be any lazy ones. (Bierwisch 1989:89)

They take Bierwisch's intuitions about the truth conditions of 'industrious' to be the model of maximal standard absolute adjectives (in their positive form) in general. They attempt to implement this core idea by invoking two different kinds of comparison classes for relative adjectives 'tall' and maximal standard absolute adjectives such as 'clean': When deciding whether someone is tall, we are supposed to invoke a contextually salient comparison class of which the person is a member, such as people of her age or profession, and ask whether she stands out against other members in terms of her height. This move is intended to respect the intuition that 'for some people to be tall there must be short people too'. But when deciding whether someone is 'industrious', we are supposed to invoke a contextually salient set of possible selves of that person, and ask whether that person is at least as industrious as each of her possible selves. This is how they attempt to respect the intuition that 'for some to be industrious there do not need to be any lazy ones'. Being industrious, as they say, concerns *within-individual comparison*, while being tall concerns *between-individual comparison*.

We will postpone evaluating their motivating idea until the end of this section, because we want to focus on how they attempt to salvage a weakened form of absolutism by folding comparison classes into the semantics. They follow Kennedy in assuming that the meaning of a maximal standard absolute adjective like 'clean' is a degree function, but they now require that the truth conditions of its positive form be sensitive to a contextually salient comparison class, which is always a set of an individual's counterparts (or possible temporal stages) who stand for the individual's possible selves. The following formulas make their idea precise and sum up how their proposal about the positive form differs from Kennedy's:

(23) *Kennedy on the positive form:*

$pos(g) = \lambda x[g(x) \geq \boldsymbol{s}(g)]$, where g is an adjective's degree function, $pos(g)$ is the adjective's meaning in its positive form, and $\boldsymbol{s}(g)$ is the contextual standard for the adjective.

(24) *Toledo and Sassoon on the positive form:*

$pos(g) = \lambda x[\forall y \in \boldsymbol{c}(g,x)(g(x) \geq g(y))]$, where $\boldsymbol{c}$ is a two-place contextually given function that maps the degree function g of a maximal standard absolute adjective and an individual x to a contextually salient set $\boldsymbol{c}(g,x)$ of possible individuals that stands in the counterpart relation to individual x.

Here is an example showing how their idea works. In order to evaluate the truth value of 'Porky is clean' in a context, discourse participants are supposed to pick out a contextually salient set of Porky's counterparts who stand for his possible selves, and determine whether Porky's degree of cleanness is at least as high as the degree of cleanness of every counterpart of his. If it is, the sentence comes out true. It it isn't, the sentence comes out false. Their proposal can be illustrated by a diagram like the following:

(25) (◯]

The brackets represent the scale encoded by the adjective 'clean'. The circle in the middle represents the image of the comparison class under the degree function of 'clean'. Only the part of the scale covered by the circle is relevant to the truth conditions of the positive form of 'clean', because an object only needs to occupy a point that is at least as high as every point inside the circle in order to be clean. So Toledo and Sassoon are proposing a local and context-dependent form of the absolutist view. Their view is also meant to be compatible with the economy principle because they suggest that the principle is responsible for the locally maximal truth conditions of maximal standard absolute adjectives.[9]

This proposal has some initial promise for answering our objection to the absolutist view. Consider the belief sentence on which our objection is based:

(26) Mary believes that both towels are clean but that the red one is cleaner than the blue one.

[9]But this application of the economy principle appears to be different from Kennedy's intended application, which focuses on the fact that absolute adjectives have closed scales.

When the towels share the same comparison class, we obtain the first, contradictory belief, reading. But since the towels need not share the same comparison class, each of them can be at least as clean as its own counterparts without having the same degree of cleanness as the other towel. This means that this belief sentence need not be an attribution of a contradictory belief to Mary.

However, it appears that this proposal doesn't provide us with the resources to distinguish between the second reading and the third reading. Since the towels' comparison classes are supposed to be contextually given based on a certain public standard, this proposal only gives us the second, public-standard, reading. So invoking comparison classes alone isn't sufficient for generating the third, private-standard, reading. We need some way of allowing the belief subject's own belief states to determine the relevant comparison classes, but it hasn't been provided yet.

Let me end with a worry about their motivating idea. Since they analyze the truth conditions of maximal standard absolute adjectives in their positive form in terms of comparisons between an (actual) individual and its counterparts — that is, they assume that comparisons between (actual) individuals are irrelevant to their truth-conditions — their proposal has some counter-intuitive consequences. For example, suppose this is how we are going to fix our understanding of the positive form of 'clean' in our conversation: We decide that an object's degree of cleanness depends solely on the quantities of germs it has per square centimeter. Let's say Porky the pig is less clean than Tom the towel by this standard. This means that Tom ought to count as clean whenever Porky counts as clean — this is what Kennedy's account would predict without the economy principle.[10] But, according to Toledo and Sassoon's account, this is not true: If Porky has been so well taken care of that he is as clean as he could possibly be, and Tom hasn't yet been boiled and disinfected and could have been cleaner, Porky can be clean without Tom being clean. The following diagrams illustrate this counter-intuitive result:

[10]This is because of his hypothesis about the relation (6) between the basic meaning of a gradable adjective and its positive form meaning, which I repeat here:

(i) $pos(g) = \lambda x[g(x) \geq s(g)]$, where g is an adjective's degree function, $pos(g)$ is the adjective's meaning in its positive form, and $s(g)$ is the contextual standard for the adjective.

(27) (p t]

(28) ($\bigcirc p$ $t \bigcirc$]

The first circle and the second circle in (28) represent the contextually salient comparison classes for Porky and Tom respectively. Since Porky is at least as clean as every member in his comparison class, he is clean. But since Tom is less clean than some member in his comparison class, he is not clean, even though he is cleaner than Porky. The counter-intuitive feel of this result gets stronger if we replace Porky in our example by an object that is typically quite dirty. (You can think of your favorite objects that are in their nature to be dirty.)

Here is what I suspect to be the source of the problem. 'Clean' does seem to have a reading that is close to Toledo and Sassoon's intuitions. On that reading, 'clean' means something like *clean for its kind*. So there is the intuition that we need to compare an individual with its counterparts, or prototypes of its kind, in order to determine whether it is clean. However, this reading is only a plausible resolution of the indeterminacy of the meaning of 'clean'. As our example shows, it is not clear it ought to be taken to be the central meaning of the positive form of 'clean', because the resolution on which whether an object counts as clean depends solely on its location on the clean scale and the contextual standard is certainly possible and, to my mind, perfectly natural.[11]

So I conclude that neither Lasersohn's pragmatic halos nor Toledo and Sassoon's granularity shift and local absolutism help the absolutist answer our objection.

1.4 The Proposal

I now present a relativist-friendly modification to Kennedy's account of the positive-form meanings of maximal standard absolute adjectives. Before I present my implementation of my proposal, let me start by explaining its motivating ideas and how they relate to Lasersohn's

[11] In fact, it is not necessary to modify Kennedy's hypothesis about the meaning of the positive form in order to derive the *clean for its kind* reading, because that reading can be derived from the degree function of 'clean' by restricting the domain to possible and actual objects of a certain kind (Kennedy 2007, example (26), p.16). So to say that something is clean for its kind is just to say that its degree of cleanness stands out against the contextually salient possible or actual objects of its kind. Of course, while the absolutists would say that to stand out is to be assigned the maximal degree by the domain-restricted function, the relativists would say that having a high enough degree is sufficient.

and Toledo and Sassoon's approaches, and forestalling a few initial worries (§1.4.2).

1.4.1 Motivating ideas

We have seen in the last section that the major problem with both Lasersohn's and Toledo and Sassoon's accounts is that they fail to deliver the private standard reading of maximal standard absolute adjectives in belief contexts. So our central task is to account for that reading, along with the public standard reading, and the contradictory belief reading as a special case of the public standard reading.[12]

I side with Toledo and Sassoon in taking the view that the imprecision of maximal standard absolute adjectives is semantic rather than pragmatic. But since invoking degree functions of different granularities may raise complications such as the one we discussed in the previous section, I will adopt a different approach. I adopt wholesale Kennedy's basic account but exploit the scale structure of maximal standard absolute adjectives to represent imprecision. When the contextual standard for a maximal standard absolute adjective coincides with the scale's endpoint, its positive form's denotation, which is the set of objects that has the property denoted by the adjective, is the smallest and maximally precise. The more the contextual standard deviates from the endpoint, the larger and the more imprecise the positive form's denotation. This simple idea bears some resemblance to Lasersohn's idea of pragmatic halos, but it is based entirely on the semantics of maximal standard absolute adjectives, and it requires no additional machinery. My claim here is that the contextual standards for maximal standard absolute adjectives are determined by the discourse participants' standard of precision, which tends to be sufficiently high given the principles of rational communication, rather than by the economy principle.

This much is trivial. We now motivate the novel part of our proposal. According to Kennedy's basic account, when a gradable adjective occurs in its positive form, it is preceded immediately by a phonologically null morpheme *pos* (for 'positive form'), which maps the basic meaning of the gradable adjective, which we should recall is a degree function, into its

[12] The contradictory belief reading can also be treated as a special case of the private standard reading. For simplicity, I assume that it is a special case of the public standard reading only.

positive form meaning, which denotes a set. The following is the meaning of *pos*:

(29) $[\![pos]\!] = \lambda g \lambda x[g(x) \geq \boldsymbol{s}(g)]$ (Kennedy 2007, p.17)

It is a function that maps the basic meaning of a gradable adjective to its positive form meaning, which denotes a set of objects whose degrees measured by the adjective is at least as high as the contextual standard for the adjective $\boldsymbol{s}(g)$. The function $\boldsymbol{s}$ is a contextually given function that maps each gradable adjective to its contextual standard. This example shows how we derive the positive form meaning of 'clean':

(30) a. $[\![clean]\!] = \lambda x[x\text{'s degree of cleanness}]$

b. $[\![pos]\!] = \lambda g \lambda x[g(x) \geq \boldsymbol{s}(g)]$

c. $[\![pos]\!]([\![clean]\!]) = \lambda x[x\text{'s degree of cleanness} > \boldsymbol{s}([\![clean]\!])]$

[(a), (b), and Function Application]

It is by positing *pos* and its meaning that Kennedy achieves the conversion from the basic meaning of a gradable adjective to its positive form meaning. Another option, which Kennedy approves, is to introduce a typeshifting rule *POS* that maps the basic meaning of a gradable adjective to its positive form meaning. It can be stated in the following way:

(31) $POS(g) = \lambda x[g(x) \geq \boldsymbol{s}(g)]$, where g is an adjective's degree function, $POS(g)$ is the adjective's meaning in its positive form, s is a contextually given function that maps each gradable adjective to its contextual standard, and $\boldsymbol{s}(g)$ is the contextual standard for the adjective.

This is why we discuss these conversion strategies. Notice that, whichever conversion strategy we adopt, the meaning of *pos* or the typeshifting rule *POS* introduces a contextually given function $\boldsymbol{s}$ that maps each adjective g to its contextual standard $\boldsymbol{s}(g)$. But as we discussed in §1.3, if the conversion can only be achieved by invoking a specific contextual standard, we are unable to account for the private standard reading, because no such standard seems appropriate. This worry applies equally to the contextually given function $\boldsymbol{c}$ Toledo and Sassoon posits to map a gradable adjective and an individual to a comparison class.

I propose a novel way to think about the contextually given function $\boldsymbol{s}$. Instead of thinking that it is literally given by the context, its possible values are being ruled out by discourse participants. The conversion from the basic meaning of a gradable adjective g to the positive form meaning never requires there be a particular contextual standard $\boldsymbol{s}(g)$. The result of the conversion can be thought of as a set of possible denotations which correspond to all possible contextual standards, instead of a single denotation.[13] As discourse participants, we are generally uncertain about the salient value(s) of the function $\boldsymbol{s}$. But insofar as we manage to contain our uncertainty within a certain narrow range, we can extract the intended information about the world from a speaker's utterance to a sufficient degree.

This is how this new way of thinking about the function $\boldsymbol{s}$ allows us to account for the private standard reading. We can imagine that the subject to whom we attribute beliefs are in a situation similar to ours, entertaining multiple possible values for the function $\boldsymbol{s}$ at where she is. So to say that Mary believes that the towels are clean by her own standard is to say that, for every possible value of the function $\boldsymbol{s}$ she entertains, the towels' degrees of cleanness she believes them to have are at least as high as the standard $\boldsymbol{s}(clean')$. This is the motivating thought of how we obtain the private standard reading.

We will implement these ideas within a Kaplanian double-indexing framework. We extend each Kaplanian formal context and circumstance of evaluation with a possible value for the function $\boldsymbol{s}$. Just as typical Kaplanian formal contexts determine all possible values for 'I' and 'now', the new extended formal contexts determine all possible values for $\boldsymbol{s}$, which in turn determine all possible contextual standards for every gradable adjective. The major benefit of adopting the Kaplanian framework is that we can help ourselves to the diagonal proposition of 'both towels are clean, but the red one is cleaner than the blue one'. This special proposition provides the belief content we need for generating the private standard reading. I will elaborate on the details more fully shortly.

Before we proceed further, let me forestall three concerns about extending the Kaplanian framework with a $\boldsymbol{s}$-function.

[13] There are views in this neighborhood. See Barker (2002, 2013) and Lassiter (2009).

1.4.2 Objections and Replies

The threshold function variable isn't an indexical

One may object that a possible value for the function s isn't exactly analogous to the possible values of indexicals such as 'I' and 'now', so it is not entitled sit alongside with them in a Kaplanian formal context.

In response, even if there are disanalogies between them, the fact that our proposed extension will yield the private standard reading, along with the public standard reading and the contradictory belief reading, will provide sufficient justification for our extension, unless there are other alternatives on offer. But let me assuage this worry the best I can by stressing a couple of similarities between them.

Let me first use an example to clarify what I mean by 'context of utterance' and Kaplanian 'formal context'. Suppose you visit a toy shop and see a stuffed animal pig with a pink tag that reads 'hug me now'. As it turns out, the sentence is generated by a computer program, written by a group of dull programmers, who work for a lonely boss who wants to be hugged by every potential consumer. The context of utterance for 'hug me now' is the physical circumstance you are in, with the stuffed animal pig, its pink tag, and your surroundings. A formal context is a n-tuple storing the possible values of various indexicals. While it is often assumed that there is an one-to-one correspondence between contexts of utterance and formal contexts, I don't make that assumption,[14] because we want to avoid speculating which formal context best characterizes the context of utterance for 'hug me now'.

There is overwhelming evidence that the context of utterance doesn't give us the referents of indexicals independently of speaker intention. There are examples such as the much-discussed answering machine example.[15] There are examples of deferred uses of indexicals, such as Nunberg (1993)'s 'I am parked out back'. There are also examples of free indirect discourse, in which the speaker adopts the perspective of another person, who isn't located at the speaker's

[14] Lasersohn (2005, 2016) also doesn't assume there is an one-to-one correspondence between contexts of utterances and formal contexts, because he holds that there is no fact of the matter as to which judge is the real judge of a context of utterance.

[15] There is a huge literature on this. See, for example, Predelli (1998, 2011).

context of utterance.[16] There are languages in which 'John thinks I am a hero' has a reading on which 'I' denotes John rather than the speaker, and the sentence is given a *de se* reading.[17] So I assume that the referents of indexicals are resolved in a context of utterance by discourse participants, rather than given directly by the context of utterance.

Since the referents of indexicals are not given directly by the context, we should take seriously the idea that we are often uncertain about their referents. It is entirely possible that a speaker is uncertain about the referent of 'I'. For example, suppose Mary suffers from amnesia. When she utters 'I'm pleased to meet you', she may not be able to tell to whom 'I' refers. In this case, her uncertainty about who is speaking can be represented by a set of formal contexts with different speaker (or self) coordinates.[18] This elegant idea has been applied to account for the *de se* reading of 'John says that I am a hero' in languages such as Amharic and Zazaki.[19] What is it for the speaker to speak truly as she utters the sentence and intends the *de se* reading? The natural answer is that it is true when, for every doxastic possibility compatible with what John says, the self in that possibility is a hero in that same possibility.

I want to pursue two similarities between indexicals and the function variable s. First, just as we can be uncertain about the referents of indexicals, we can be uncertain about the contextual standards for maximal standard absolute adjectives (and gradable adjectives more generally), as we often are. Our uncertainty about the contextual standards can be represented by a set of formal contexts with different s functions. This gives us a natural way to account for the private standard reading: We can say that the private standard reading of 'Mary believes that the towels are clean' is true when, for every doxastic possibility compatible with Mary's beliefs, the towels' degrees of cleanness in that possibility are at least as high as the contextual standard for 'clean' in that same possibility. These truth conditions are just like those for the *de se* reading of 'John thinks that I am a hero' in languages such as Amharic and Zazaki.

Second, when our belief sentence is given the public standard, or the wide scope, reading, the value of the threshold function variable s is to be resolved by the discourse participants

[16] See Predelli (1998) and Recanati (2004), among others.

[17] See Schlenker (2003), Anand and Nevins (2004), and Anand (2006). See Ninan (2010) for a more accessible discussion.

[18] This way of representing an individual's uncertainty or ignorance about the referents of indexicals and demonstrates is due to Lewis (1979a).

[19] See, in particular, Anand and Nevins (2004), and Anand (2006).

rather than determined by Mary's belief states. This reading is analogous to the non *de se*, speaker-referring, reading of 'I' when it is embedded under belief contexts. This similarity gives us the truth conditions for the public standard reading, and the contradictory belief reading as a special case.

Certain ambiguities

One may worry that, since the public standard and the private standard readings are just like the scopal ambiguities of the following sentences, there is no need to motivate any new tools to account for them:

(32) I thought your yacht was larger than it was.

(33) (This draft is 10 pages) The paper is required to be exactly 5 pages longer than that (Heim 2000).

The first sentence, due to Bertrand Russell (1905), is ambiguous between the contradictory reading, which says the speaker thought the yacht was larger than itself, and the overestimation reading, which says that the speaker thought the yacht was larger than it actually was. The second sentence also has two readings. On the first reading, it says that the paper has to be exactly 15 pages long. On the second reading, it says that a 15-page long paper is certainly acceptable, but it leaves open whether it can be longer than that.

There are lots of discussions about these ambiguities.[20] One may worry that the ambiguity between the public standard reading and private standard readings is merely an instance of these ambiguities, and that the tools that account for these ambiguities can account for the public standard and the private standard readings as well. To assuage this worry, we assess the likelihood that an account of these ambiguities can readily explain the public standard and the private standard readings by looking at Heim (2000)'s analysis of these ambiguities.

We should notice an initial disanalogy between the belief sentence we have been focusing on and the sentences (32) and (33): The gradable adjectives in them are in their comparative form,

[20] See, for example, Postal (1974), Horn (1981), von Stechow (1984), Heim (2000).

rather than in the positive form. According to Heim, Russell's ambiguity, as in (32), can be accounted for if we allow each predicate to contain an explicit world-argument, which can be either free or bound. Following her suggestion, and staying as closely as we can to Kennedy's account of gradable adjectives, we can take the meaning of 'large' to be a function that maps a possible world and an individual to a degree:[21]

(34) $[\![large]\!] = \lambda w[\lambda x[x\text{'s size in world } w]]$

We can then introduce a rigidification operation that maps this new meaning of 'large' to a function that maps a possible world and an individual to that individual's size in the actual world:

(35) $r([\![large]\!]) = \lambda w[\lambda x[x\text{'s size in world } w_0]]$,

where r is an operation that maps a meaning g of type $\langle i, \langle e, d \rangle \rangle$ to the meaning $\lambda w[\lambda x[g(w_0)(x)]]$

With these modifications, along with appropriate modifications throughout the semantics, we can generate the two readings along the following lines (I ignore the tense here and take no stand on the exact analyses for '-er' and the than-clause):

(36) a. For every world w' compatible with what I think in the actual world w_0, the size of your yacht in w' is bigger than the size of your yacht in w'.

b. For every world w' compatible with what I think in the actual world w_0, the size of your yacht in w' is bigger than the size of your yacht in the actual world w_0.

The first and the second representations correspond to the contradictory reading and the overestimation readings respectively. So this strategy nicely accounts for Russell's ambiguity. However, it doesn't seem to provide a ready explanation for the ambiguity between the public standard and the private standard readings. Even if we allow explicit quantification over worlds

[21] In her discussion, the meaning of a gradable adjective is a two-place relation between a degree and an individual. Here I aim to present the core idea of her proposal and be neutral between her assumptions about the meanings of gradable adjectives and Kennedy (2007)'s.

in the syntax, there is still the issue of converting the basic meaning of a gradable adjective into its positive-form meaning. If the conversion is still achieved by introducing a contextually salient function $\boldsymbol{s}$, this strategy doesn't make it any easier for us to generate the private standard reading.[22] To see this, consider the new positive form meaning for 'large':

(37) $pos([\![large]\!]) = \lambda w[\lambda x[x\text{'s size in world } w \geq \boldsymbol{s}([\![large]\!], w)]]$

Quite naturally, the contextually salient function $\boldsymbol{s}$ is now a two-place function that maps a gradable adjective and a world to the adjective's contextual standard in that world.[23] Given the context for the private standard reading, the speaker is in no position to tell what that function ought to be. Further, the additional world-argument in the function $\boldsymbol{s}$ seems to make it harder for the speaker to know what it ought to be, not easier, because the function is supposed to determine the relevant contextual standard for each of Mary's belief worlds. So explicit quantification over worlds isn't what we need to explain the reading.

We now turn to the second sentence (33). According to the Heim's analysis, the constituent responsible for its scopal ambiguities is the degree quantifier 'exactly 5 pages -er than that', rather than the adjective itself. When it takes narrow scope under the modal context 'it is required that', we get the first, exactly 15 pages, reading. When it takes wide scope over the modal context, we get the second, at least 15 pages, reading. It will take us too far afield to discuss Heim's analysis in detail. But since our belief sentence doesn't contain any degree quantifiers, it seems safe to assume that the ambiguity between the public standard and the private standard readings isn't quite the same as the ambiguity of (33).

[22]The problem we focus on primarily concerns the degree-based account of the positive form of gradable adjectives, such as Kennedy's. Since the non-degree-based approach takes the basic meaning of a gradable adjective to be set denoting, it may use explicit quantification over worlds to account for the apparent wide scope and narrow scope readings. For example, the narrow scope and the wide scope readings of 'Mary believes that Porky is clean' can be analyzed as follows:

(i) a. For every world w' compatible with what Mary believes in the actual world w_0, the denotation of 'Porky' in w' is in the set of clean objects in w' (narrow scope).
b. For every world w' compatible with what Mary believes in the actual world w_0, the denotation of 'Porky' in w' is in the set of clean objects in the actual world w_0 (wide scope).

See, Klein (1980), van Benthem (1982), van Rooij (2011), and Burnett (2014) for examples of the non-degree-based approach to gradable adjectives.

[23]One can keep the function a single-place function just as before. My point doesn't depend on whether the function is one-place or two-place.

Propositional radicals

The third objection is of a more skeptical nature.[24] Its main thrust is that we have been trying to solve a pseudo-problem, because the proper semantic content of our belief sentence is a propositional radical with the function s unresolved, and the two readings are both due to the pragmatic context.

However, this line of objection doesn't have much force unless it can explain why the following sentence isn't ambiguous in a way analogous to our belief sentence:

(38) Mary believes that she loves it/ that.

Like the adjective 'clean' in our belief sentence, the pronoun 'it' or the demonstrative 'that' also asks for a certain value from the context. So, in principle, it is possible for the pragmatic context to generate a reading on which the value is determined by Mary's belief states alone. However, the only available reading is that Mary believes that she loves a certain contextually salient object. Regardless of what the pragmatic context is, there isn't a reading analogous to the private standard reading where she believes that she loves the object determined by her belief states. It seems our semantic explanation is more satisfactory than the pragmatic one because it makes definite predictions about which ambiguities exist and which don't.

We are finished addressing a few initial concerns about our account. Let's now proceed to fleshing it out in more detail.

1.4.3 Double-indexing with s-functions

Here is our plan. We will proceed by first laying out our double-indexing framework. After that, we enrich it with s-functions as promised, and explain how we embed Kennedy's basic account within it. Then we will discuss how we obtain the public standard reading and the private standing reading (§1.4.4).

While its specifics may vary from author to author depending on their aims, the hallmark of a Kaplanian double-indexing framework is that the formal contexts are kept distinct from the

[24]Someone sympathetic to Bach (1994, 2005)'s view on the semantics-pragmatics distinction might adopt this line of thought. Of course, I don't suggest that Bach would raise the following objection himself.

circumstances of evaluation. Following Anand (2006), I assume that both formal contexts and circumstances of evaluation are n-tuples storing the possible values of the standard indexicals. Let's assume, for simplicity, that they are 4-tuples of the following form:

(39) $\langle SELF, LOCATION, TIME, WORLD \rangle$, where SELF is an individual serving as a possible value of 'I', LOCATION a location serving as a possible value of 'here', TIME a time interval serving as a possible value of 'now', and WORLD a possible or actual state of the world.

Let's get a flavor of how our framework works by walking through the semantic composition of 'I am here now'. The lexical meanings of 'I', 'am', 'here', and 'now' are as follows:

(40) a. $[\![I]\!] = \lambda c \lambda i[SELF(c)]$

b. $[\![am]\!] = \lambda c \lambda i \lambda y \lambda x[x \text{ is located at } y \text{ at } TIME(i) \text{ in } WORLD(i)]$

c. $[\![here]\!] = \lambda c \lambda i[LOCATION(c)]$

d. $[\![now]\!] = \lambda c \lambda i \lambda p_{\langle i,t \rangle}[p(i[TIME(i)/TIME(c)])]$

Let's unpack these meanings. The lexical meaning of 'I' is a function that takes a formal context c as input, and outputs a function that takes a circumstance of evaluation i as input and outputs the self coordinate $SELF(c)$ of the formal context c. The lexical meanings of 'am' and 'here' can be read off from the formulas above in a similar way. Before we unpack the meaning of 'now', we'll introduce some helpful vocabularies. We will follow Kaplan in calling functions such as the lexical meaning of 'I' *characters*, the result of applying a character to a formal context a *content* or an *intension*, and the result of applying a content to a circumstance of evaluation an *extension*. 'Now' is essentially a sentential operator. Its extension relative to a formal context c and a circumstance of evaluation i is a function that takes the content of a sentence at c as input,[25] and maps it to TRUE if the content is true at a circumstance of evaluation that is just like i, except that its TIME coordinate is replaced by that of c.

To show how the lexical meanings above combine to give the truth conditions of 'I am here

[25]The rule content F.A. (41-b) ensures that the content of the sentence is fixed by the formal context argument of the character of 'now'. See the full derivations for details.

now', I will introduce the following type-sensitive function application rules:

(41) *Type-sensitive Function Application (extension, content, and character F.A.):*

a. If α is of type $\langle c, \langle i, \langle \alpha, \beta \rangle\rangle\rangle$ and β is of type $\langle c, \langle i, \alpha \rangle\rangle$, then $\alpha(\beta) = \lambda c[\lambda i[\alpha(c)(i)(\beta(c)(i))]]$

b. If α is of type $\langle c, \langle i, \langle i\alpha, \beta \rangle\rangle\rangle$ and β is of type $\langle c, \langle i, \alpha \rangle\rangle$, then $\alpha(\beta) = \lambda c[\lambda i[\alpha(c)(i)(\beta(c))]]$

c. If α is of type $\langle c, \langle i, \langle ci\alpha, \beta \rangle\rangle\rangle$ and β is of type $\langle c, \langle i, \alpha \rangle\rangle$, then $\alpha(\beta) = \lambda c[\lambda i[\alpha(c)(i)(\beta)]]$

Here I adopt the approach of generalizing to the worst case: Since the official meaning of each lexical item is its character, I assume that semantic composition operates directly on characters. Let c and i be an arbitrary formal context and an arbitrary circumstance of evaluation respectively. The first rule essentially combines the extension of α at $\langle c, i \rangle$ and the extension of β at $\langle c, i \rangle$. So it's called 'extension function application'. The second rule essentially combines the extension of α at $\langle c, i \rangle$ with the intension of β at c. So it is called 'content function application' or 'intension functional application'. The third rule essentially combines the extension of α at $\langle c, i \rangle$ with the character of β. It is called 'character function application'.

Here is the derivation of the meaning (character) of 'I am here now':

(42) I am here now:

a. $[\![I]\!] = \lambda c \lambda i[SELF(c)]$

$[\![am]\!] = \lambda c \lambda i \lambda y \lambda x[x \text{ is located at } y \text{ at } TIME(i) \text{ in } WORLD(i)]$

$[\![here]\!] = \lambda c \lambda i[LOCATION(c)]$

$[\![now]\!] = \lambda c \lambda i \lambda p_{\langle i,t \rangle}[p(i[TIME(i)/TIME(c)])]$

b. $[\![am\ here]\!] = [\![am]\!]([\![here]\!])$

$= \lambda c \lambda i[[\![am]\!](c)(i)([\![here]\!])(c)(i)]$ [extension F.A.]

$= \lambda c \lambda i \lambda x[x \text{ is located at } LOCATION(c) \text{ at } TIME(i) \text{ in } WORLD(i)]$

c. $[\![I\ am\ here]\!] = [\![am\ here]\!]([\![I]\!])$

$= \lambda c \lambda i[SELF(c) \text{ is located at } LOCATION(c) \text{ at } TIME(i) \text{ in } WORLD(i)]$

[extension F.A.]

d. $[\![I\ am\ here\ now]\!] = [\![now]\!]([\![I\ am\ here]\!])$

$= \lambda c \lambda i[[\![now]\!](c)(i)([\![I\ am\ here]\!](c))]$ [content F.A.]

$= \lambda c \lambda i[\lambda p_{\langle i,t \rangle}[p(i[TIME(i)/TIME(c)])]([\![I\ am\ here]\!](c))]$

$= \lambda c \lambda i[[\![I\ am\ here]\!](c)(i[TIME(i)/TIME(c)])]$

$= \lambda c \lambda i[SELF(c)$ is located at $LOCATION(c)$ at $TIME(i[TIME(i)/TIME(c)])$ in $WORLD(i[TIME(i)/TIME(c)])]$

$= \lambda c \lambda i[SELF(c)$ is located at $LOCATION(c)$ at $TIME(c)$ in $WORLD(i)]$

Let's check whether the resulting meaning of the sentence is correct. As we do so, the reader should pay attention to how our interpretation of it differs from the standard interpretation due to Kaplan, on which the sentence is true at every formal context. Suppose Mary utters 'I am here now' at her local animal sanctuary at 6:00 pm on 1/1/2017 in our world, and we know these particulars of her utterance maximally precisely. For simplicity, we can assume that the set of 4-tuples representing our uncertainty about the speaker, the time and the location, and the state of the world relevant to our interpretation of Mary's utterance is the singleton containing the 4-tuple ⟨Mary, Mary's local animal sanctuary, 6:00 pm on 1/1/2017, the actual world⟩. This means that the content or the proposition we extract from the character of 'I am here now', based on our uncertainty about the context of utterance, is as follows:

(43) $\lambda i[$Mary is located at Mary's local animal sanctuary at 6:00 pm on 1/1/2017 in $WORLD(i)]$

This proposition (intension) is true at any circumstance of evaluation whose world coordinate is the actual world, but it can be false at circumstances of evaluation with a different world. Notice that, unlike Kaplan's account, our own framework does not predict that, for every formal context, the sentence is true at the world coordinate and the time coordinate of that formal context, because we don't assume that there is an one-to-one correspondence between contexts of utterance and formal contexts, nor do we stipulate that, for every formal context $\langle s, l, t, w \rangle$, the self s is located at location l at time t and world w — in short, we allow what

Kaplan calls 'improper contexts'.[26]

It's also important to note that there need not be a unique best content or proposition that can be identified with what a sentence literally says. For example, the proposition(s) we extract from the character of the sentence 'I am here now' depends on our uncertainty about the referents of the indexicals in the sentence. Suppose we allow an one-minute margin of error for our belief about the time of Mary's utterance. Instead of extracting a single proposition from the sentence's character, we extract from it a set of propositions, each of which is based on a precise time between 6:00 pm and 6:01 pm.

Let's walk through one more example before we embed Kennedy's basic account into our framework. The example is the result of embedding 'I am here now' inside the belief context 'Esther believes that'. It is a good warm-up before we discuss the public standard and the private standard readings of our belief sentence. The lexical meaning of 'believe' is as follows:

(44) $[\![believe]\!] = \lambda c \lambda i \lambda p_{\langle i,t \rangle} \lambda x[\forall i'(i'R(x)(i) \rightarrow p(i') = 1)]$, where '$i'R(x)(i)$' reads: circumstance of evaluation i' is compatible with the beliefs of individual x at circumstance of evaluation i.

Let's unpack its meaning. Its extension at an arbitrary formal context c and an arbitrary circumstance of evaluation i takes both the content (intension) of the sentence embedded under it at c and an individual x (i.e. the belief subject) as inputs, and maps them to TRUE if the content is true at (the world and the time coordinates of) every circumstance of evaluation that is compatible with the beliefs of x at i.

Here is the semantic composition for 'Esther believes that I am here now':

(45) a. $[\![Esther]\!] = \lambda c \lambda i[Esther]$

b. $[\![believe]\!] = \lambda c \lambda i \lambda p_{\langle i,t \rangle} \lambda x[\forall i'(i'R(x)(i) \rightarrow p(i') = 1)]$

c. $[\![I\ am\ here\ now]\!]$

[26] Kaplan's assumptions that there is an one-to-one correspondence between contexts of utterance and formal contexts, and that every formal context is proper lead to some undesirable consequences. For example 'I am here now' is predicted to be true whenever and wherever it is uttered. However, if someone leaves 'I am not here now' as a recorded message on her phone, her utterance can naturally be interpreted to be true by her callers. See Predelli (1998, 2011) for further discussions.

$= \lambda c \lambda i[SELF(c)$ is located at $LOCATION(c)$ at $TIME(c)$ in $WORLD(i)]$

d. $[\![$*believe I am here now*$]\!] = [\![$*believe*$]\!]([\![$*I am here now*$]\!])$

$= \lambda c \lambda i \lambda x[\forall i'(i' R(x)(i) \rightarrow SELF(c)$ is located at $LOCATION(c)$ at $TIME(c)$ in $WORLD(i'))]$

[(b), (c), content F.A.]

e. $[\![$*Mary believe I am here now*$]\!] = [\![$*Mary*$]\!]([\![$*believe I am here now*$]\!])$

$= \lambda c \lambda i[\forall i'(i' R(Esther)(i) \rightarrow SELF(c)$ is located at $LOCATION(c)$ at $TIME(c)$ in $WORLD(i'))]$

[(a), (d), extension F.A.]

Let's check whether the resulting meaning is correct. Suppose the context of utterance is the same as above: Mary utters the sentence at her local animal sanctuary at 6:00 pm on 1/1/2017 in our world. And our uncertainty about the context of utterance is again the singleton containing the 4-tuple $\langle$Mary, Mary's local animal sanctuary, 6:00 pm on 1/1/2017, the actual world$\rangle$. So the proposition we extract from the character of 'Esther believes that I am here now' is as follows:

(46) $\lambda i[\forall i'(i' R(Esther)(i) \rightarrow$ Mary is located at Mary's local animal sanctuary at 6:00 pm on 1/1/2017 in $WORLD(i'))]$

This proposition is intuitively correct, because it is true at a circumstance of evaluation i just in case, for every circumstance of evaluation i' compatible with Esther's beliefs at i, Mary is located at Mary's local animal sanctuary at 6:00 pm on 1/1/2017 in the world of i'. Notice that the interpretations of 'I', 'here', 'now' aren't determined by the belief states of Esther, but are 'given by the context' — or, more cautiously, resolved based on our uncertainty about the context of utterance. This is the result we want.

We are now finished laying out our own double-indexing framework. Let me explain why it is a natural choice for our purpose. In order to account for the public standard reading, we want the $\boldsymbol{s}$-function to be 'given by the context' rather than determined by the belief states of the belief subject, so that the subject bears the belief relation to a content that is determined by the discourse participants' public standard of precision. This can be achieved by extending the formal contexts with possible values of the $\boldsymbol{s}$-function. For a reason that will become clear in the next subsection (§1.4.4), we will extend the circumstances of evaluation with $\boldsymbol{s}$-functions as well. So both of them are now 5-tuples like the following:

(47) $\langle SELF, LOCATION, TIME, WORLD, \boldsymbol{s}\rangle$, where $SELF$ is an individual, $LOCATION$ a location, $TIME$ a time interval, $WORLD$ a possible or actual state of the world, and $\boldsymbol{s}$ a function that maps each maximal standard absolute adjective (and, in fact, each gradable adjective) to a degree on its scale.

Our next step is to convert the basic meaning of each maximal standard absolute adjective into an appropriate character. Here is a natural way to proceed. The basic meaning for a maximal standard absolute adjective like 'clean' is now a constant function from formal contexts to functions from circumstances of evaluation to degree functions (i.e. of type $\langle c, \langle i, \langle e, d\rangle\rangle\rangle$). Here is my justification for this modification: The degree function of 'clean' can vary across circumstances of evaluation having different world coordinates (and time coordinates) because an object can have different degrees of cleanness in different worlds (and at different times). But that the intension of 'clean' is constant across formal contexts is nothing but a simplifying assumption.

Let's turn now to how we convert the basic meaning of a maximal standard absolute adjective into its positive form meaning. As I said in §1.4.1, what is novel about our proposal is that the conversion doesn't depend on there being a specific contextual standard. This can be achieved once we work with the character of the positive form, whose extensions can be thought of as a collection of sets corresponding to all possible contextual standards, instead of its extension, which is only a single set. I propose the following typeshifting rule that maps the basic meaning of a maximal standard absolute adjective to its positive form meaning:

(48) *New account of the positive form:*

$pos(g_{\langle ci,ed\rangle}) = \lambda c\lambda i\lambda x[g(c)(i)(x) \geq \boldsymbol{s_c}(g,i)]$, where g is the character of a maximal standard absolute adjective, and $\boldsymbol{s_c}$ is the $\boldsymbol{s}$-function of formal context c, which maps a gradable adjective g and a circumstance of evaluation i to a point on the adjective's scale at i.

Let's assume that the adjective g is 'clean'. This is what the definition says. The positive-form meaning of 'clean' is a character such that, given a formal context c, an object x falls into the set of clean objects at the world and the time coordinates of a circumstance of evaluation i just in case the object's degree of cleanness at those coordinates is at least as high as the contextual standard for 'clean' at i, i.e. $\boldsymbol{s_c}(\llbracket clean \rrbracket, i)$. For simplicity, we will assume that the function $\boldsymbol{s_c}$ is *rigid* in the sense that it maps 'clean' to the same contextual standard at every circumstance of evaluation, that is:

(49) $\forall c \forall i \forall i' (\boldsymbol{s_c}(\llbracket clean \rrbracket, i) = \boldsymbol{s_c}(\llbracket clean \rrbracket, i'))$

While this new account of the positive form meaning looks unduly complicated, the underlying idea is very simple. I'll use numerical values to illustrate the idea — but it is important to note that my proposal doesn't identify degrees with numerical values. Let's suppose the content of 'clean' at a certain circumstance of evaluation i encodes the scale (0,1]. Let's say the discourse participants' standard of precision determines the contextual standard for 'clean' to be 0.8. Only those objects whose degrees of cleanness at (the world and the time coordinates of) circumstance of evaluation i are at least as high as 0.8 will fall into the set of clean objects at (the world and the time coordinates of) i. When the public standard of precision becomes higher and sets the contextual standard to 0.9, the set of clean objects at (the world and the time coordinates of) i will shrink, as fewer objects meet the more stringent standard of cleanness. The idea here is similar to Lasersohn's pragmatic halos, but it is implemented within the semantics, with no additional machinery.

It may also be helpful to illustrate the idea graphically. The character of a sentence like 'Porky is clean' can be thought of as a table with infinitely many rows and columns. The following is a finite snapshot of it:

(50)

	i_1 [1]	i_2 [0.9]	i_3 [0.8]
c_1 [1]	$g(i_1)(p) = 0.8$	$g(i_2)(p) = 0.83$	$g(i_3)(p) = 0.85$
c_2 [0.9]	$g(i_1)(p) = 0.8$	$g(i_2)(p) = 0.83$	$g(i_3)(p) = 0.85$
c_3 [0.8]	$g(i_1)(p) = 0.8$	$g(i_2)(p) = 0.83$	$g(i_3)(p) = 0.85$

The numbers in square brackets are the contextual standards for 'clean' in their respective formal contexts and circumstances of evaluation.[27] '$g(i_1)(p) = 0.8$' reads the degree of cleanness of Porky at (the world and the time coordinates of) circumstance of evaluation i_1 is 0.8. The first row captures the truth conditions Kennedy's enhanced account would assign to the sentence: Since the contextual standard at c_1 is 1, to be clean at (the world and the time coordinates of) a circumstance of evaluation is to be maximally clean in that circumstance's world and time coordinates. That's why Porky isn't clean in (the world and the time coordinates of) every circumstance of evaluation on the first row. But, unlike Kennedy's account, we allow for the possibility that the positive-form has imprecise, non-absolutist, semantic contents, such as those represented by the remaining rows. The third row says that Porky eventually counts as clean at (the world and the time coordinates of) every circumstance of evaluation listed here when the contextual standard is 0.8.

The alert reader should notice I mentioned repeatedly the world and the time coordinates of circumstances of evaluation in parentheses. Here is the reason why I did so. A proposition is true or false at a circumstance of evaluation because of the world and the time coordinates it has. On my view, a proposition can't have different truth values at two distinct circumstances of evaluation unless they have different world or time coordinates. It is important to stress this point because I don't assume that a proposition can have different truth values at two distinct circumstances of evaluation in virtue of their having different coordinates which are neither worlds nor times (e.g. judges).[28] However, since the locution 'being true at the world and the time coordinates of a circumstance of evaluation' is a bit cumbersome, I shorten it as 'being true at a circumstance of evaluation', and trust that the reader understands it in my intended way.

[27]In this example, we assume that the circumstances of evaluation have different world coordinates or different time coordinates.

[28]See Lasersohn (2005, 2016) and Stephenson (2007) for their justifications of this assumption.

For a similar reason, instead of using the locution ‘Porky’s degree of cleanness at the world and the time coordinates of a circumstance of evaluation’, I use the shortened form ‘Porky’s degree of cleanness at a circumstance of evaluation’ instead.

We are now ready to account for the public standard reading. It is important to keep in mind that the underlying intuition of the public standard reading is very simple: It is no different from that of the only reading of ‘Esther believes that I am here now’ in English, because the contextual standard for ‘clean’ is ‘given by the context’ just like the English indexicals ‘I’, ‘here’, and ‘now’.

To simplify the semantic composition, instead of using our original belief sentence (51-a), we will call the red towel ‘Ruby’ and the blue towel ‘Tom’, and simplify the sentence as (51-b):

(51) a. Mary believes that both towels are clean but that the red one is cleaner than the blue one.

b. Mary believes that Ruby is clean and Tom is clean but that Ruby is cleaner than Tom.

Here is the semantic composition in full:

(52) Lexical meanings:

a. $[\![Mary]\!] = \lambda c \lambda i[Mary]$ (likewise for ‘Ruby’ and ‘Tom’)

b. $[\![believe]\!] = \lambda c \lambda i \lambda p_{\langle i,t \rangle} \lambda x[\forall i'(i' R(x)(i) \rightarrow p(i') = 1)]$

c. $[\![and]\!] = [\![but]\!] = \lambda c \lambda i \lambda q_t \lambda p_t [p \wedge q]$

d. $[\![clean]\!] = \lambda c \lambda i \lambda x[x\text{'s degree of cleanness at } i]$

e. $[\![-er]\!] = \lambda c \lambda i \lambda g_{\langle ci,ed \rangle} \lambda y \lambda x[g(c)(i)(x) > g(c)(i)(y)]$

f. $[\![is]\!] = [\![than]\!] = \lambda c \lambda i \lambda \alpha[\alpha(c)(i)]$

(53) Ruby is cleaner than Tom:

a. $[\![cleaner]\!] = [\![-er]\!]([\![clean]\!])$

$= \lambda c \lambda i \lambda y \lambda x[[\![clean]\!](c)(i)(x) > [\![clean]\!](c)(i)(y)]$

[(52-d), (52-e), and character F.A.]

b. $[\![\textit{than Tom}]\!] = [\![\textit{than}]\!]([\![\textit{Tom}]\!]) = \lambda c\lambda i[[\![\textit{Tom}]\!](c)(i)] = [\![\textit{Tom}]\!]$

[(52-a), (52-f), and character F.A.]

c. $[\![\textit{cleaner than Tom}]\!] = [\![\textit{cleaner}]\!]([\![\textit{than Tom}]\!])$

$= \lambda c\lambda i\lambda x[[\![\textit{clean}]\!](c)(i)(x) > [\![\textit{clean}]\!](c)(i)(Tom)]$

[(53-a), (53-b), and extension F.A.]

d. $[\![\textit{Ruby is cleaner than Tom}]\!] = [\![\textit{cleaner than Tom}]\!]([\![\textit{Ruby}]\!])$

$= \lambda c\lambda i[[\![\textit{clean}]\!](c)(i)(Ruby) > [\![\textit{clean}]\!](c)(i)(Tom)]]$

[(52-a), (53-c), and extension F.A.]

(54) Ruby is clean (likewise for 'Tom is clean'):

a. $[\![\textit{clean}]\!]^{pos} = \lambda c\lambda i\lambda x[[\![\textit{clean}]\!](c)(i)(x) \geq s_c([\![\textit{clean}]\!], i)]$

[by the new account of the positive form (48)]

b. $[\![\textit{is clean}]\!] = [\![\textit{is}]\!]([\![\textit{clean}]\!]^{pos})$

$= \lambda c\lambda i[[\![\textit{clean}]\!]^{pos}(c)(i)] = [\![\textit{clean}]\!]^{pos}$

[(52-f), (54-a), and character F.A.]

c. $[\![\textit{Ruby is clean}]\!] = [\![\textit{is clean}]\!]([\![\textit{Ruby}]\!])$

$\lambda c\lambda i[[\![\textit{clean}]\!](c)(i)(Ruby) \geq s_c([\![\textit{clean}]\!], i)]$

[(54-a), (52-a), and extension F.A.]

(55) Ruby is clean and Tom is clean:

a. $[\![\textit{and Tom is clean}]\!] = [\![\textit{and}]\!]([\![\textit{Tom is clean}]\!])$

$= \lambda c\lambda i\lambda p_t[p \wedge [\![\textit{clean}]\!](c)(i)(Tom) \geq s_c([\![\textit{clean}]\!], i)]$

[(52-c), (54), and extension F.A.]

b. $[\![\textit{Ruby is clean and Tom is clean}]\!]$

$= [\![\textit{and Tom is clean}]\!]([\![\textit{Ruby is clean}]\!])$

$= \lambda c\lambda i[[\![\textit{clean}]\!](c)(i)(Ruby) \geq s_c([\![\textit{clean}]\!], i) \wedge [\![\textit{clean}]\!](c)(i)(Tom) \geq s_c([\![\textit{clean}]\!], i)]$

[(54-c), (55-a), and extension F.A.]

(56) Ruby is clean and Tom is clean but that Ruby is cleaner than Tom:

a. $[\![\textit{Ruby is clean and Tom is clean but Ruby is cleaner than Tom}]\!]$

$= \lambda c \lambda i[[\![clean]\!](c)(i)(Ruby) \geq s_c([\![clean]\!], i) \wedge [\![clean]\!](c)(i)(Tom) \geq s_c([\![clean]\!], i) \wedge$

$[\![clean]\!](c)(i)(Ruby) > [\![clean]\!](c)(i)(Tom)]$

[(52-c), (53), (55), and extension F.A.]

(57) Mary believes Ruby is clean and Tom is clean but that Ruby is cleaner than Tom:

a. $[\![$*believes Ruby is clean and Tom is clean but Ruby is cleaner thanTom*$]\!]$

$= [\![believe]\!]([\![$*Ruby is clean and Tom is clean but Ruby is cleaner than Tom*$]\!])$

$= \lambda c \lambda i[[\![believe]\!](c)(i)([\![$*Ruby is clean and Tom is clean but Ruby is cleaner than Tom*$]\!](c))]$

[(52-b), (56-a), and content F.A.]

$= \lambda c \lambda i \lambda x[\forall i'(i'R(x)(i') \rightarrow [\![clean]\!](c)(i')(Ruby) \geq s_c([\![clean]\!], i') \wedge [\![clean]\!](c)(i')(Tom) \geq$

$s_c([\![clean]\!], i') \wedge [\![clean]\!](c)(i')(Ruby) > [\![clean]\!](c)(i')(Tom))]$

b. $[\![$*Mary believes Ruby is clean and Tom is clean but Ruby is cleaner thanTom*$]\!]$

$= [\![$*believes Ruby is clean and Tom is clean but Ruby is cleaner thanTom*$]\!]([\![Mary]\!])$

$= \lambda c \lambda i[\forall i'(i'R(Mary)(i) \rightarrow [\![clean]\!](c)(i')(Ruby) \geq s_c([\![clean]\!], i') \wedge [\![clean]\!](c)(i')(Tom) \geq$

$s_c([\![clean]\!], i') \wedge [\![clean]\!](c)(i')(Ruby) > [\![clean]\!](c)(i')(Tom))]$

[(52-a), (57-a), and extension F.A.]

The following is our belief sentence (57)'s extension relative to an arbitrary formal context c and an arbitrary circumstance of evaluation i:

(58) $\forall i'(i'R(Mary)(i) \rightarrow [\![clean]\!](c)(i')(Ruby) \geq s_c([\![clean]\!], i') \wedge [\![clean]\!](c)(i')(Tom) \geq s_c([\![clean]\!], i') \wedge$

$[\![clean]\!](c)(i')(Ruby) > [\![clean]\!](c)(i')(Tom))$

It says that, for every circumstance of evaluation i' compatible with Mary's beliefs at i, both towels' degrees of cleanness at i' are at least as high as the contextual standard determined by the function s_c, and that the red towel's (Ruby's) degree of cleanness at i' is higher than the blue towel's (Tom's) degree of cleanness at i'.

For simplicity, let's suppose we have precise knowledge of the contextual standard determined by the function $\boldsymbol{s_c}$. If the contextual standard is 0.8 on the clean scale, then our belief sentence is true at a circumstance of evaluation i if and only if, in every circumstance of evaluation compatible with Mary beliefs at *i*, the towels are clean up to degree 0.8 and that the red towel is cleaner than the blue one. Clearly, this means that Mary can believe coherently that the red towel is cleaner than the blue one.

If the contextual standard is maximal, then our belief sentence is true at *i* if and only if, for every circumstance of evaluation compatible with her beliefs at *i*, the towels are clean up to the maximal degree and that the red towel is cleaner than the blue one. And we know that this means that it is true only when Mary has a contradictory belief.

So we have obtained the desired truth conditions for the public standard reading, and for the contradictory belief reading as a special case.

1.4.4 Diagonalization and the private standard reading

Let's now discuss how we obtain the private standard reading. As we go through the details, it is important to bear in mind that the motivating idea is very simple. As we discuss in §1.4.1, to say that Mary believes that the towels are clean by her own standard is to say that, for every contextual standard of 'clean' compatible with her beliefs, the degrees of cleanness she believes the towels to have are at least as high as that standard. These truth conditions are analogous to the truth conditions of the *de se* reading of 'John says that I am a hero' in languages such as Amharic and Zazaki: If John says that he himself is a hero, then for every possible value of 'I' compatible with what he says, that value is a hero. Anand (2006) delivers these truth conditions by diagonalizing the character of 'I am a hero'.[29] Given the similarity between the private standard reading and the *de se* reading of 'I' in say-context, we are going to use a similar method to account for the private standard reading.

As a warm-up, let's pretend that English is Amharic, and that 'John says that I am a hero' does have a *de se* reading. If the extension of 'I am a hero' is always insensitive to the modal context 'John says that', then we clearly can't obtain its *de se* truth conditions. To allow the extension of 'I am a hero' to properly vary with the circumstances of evaluation compatible with what John says, Anand and Nevins (2004) introduce an operation that maps the character of 'I am a hero' to a proposition that is true at a circumstance of evaluation *i* just in case the self

[29]See Stalnaker (1978) for other applications of the diagonalization operation.

coordinate of i is a hero at the world and the time coordinates of i. Once we can help ourselves to this proposition (i.e. the diagonal proposition), we can say that the *de se* reading of 'John says that I am a hero' is true just in case that proposition is true at every circumstance of evaluation compatible with what John says.

Let's discuss Anand and Nevin's proposal in more detail. The following table illustrates the diagonal proposition of the character of 'I am a hero':

(59)

	i_1	i_2	i_3
c_1	**PH**	PH	PH
c_2	EH	**EH**	EH
c_3	MH	MH	**MH**

This table is a finite snapshot of the character of 'I am a hero'. The self coordinates of the formal contexts c_1, c_2, and c_3 are Porky, Esther, and Mary respectively. So if our uncertainty about the context of utterance is best represented by the singleton containing the formal context c_1, then the content we extract from the character is the proposition that Porky is a hero (abbreviated as 'PH' on the table). The same reasoning applies to the formal contexts c_2 and c_3. The diagonal going from the top left corner to the bottom right corner is a proposition that is true at i_1 if Porky is a hero at the world and the time coordinates of i_1, at i_2 if Esther is a hero at the world and the time coordinates of i_2, and at i_3 if Mary is a hero at the world and the time coordinates of i_3. This proposition is what we need to obtain the *de se* reading of John says that I am a hero. Here is why. Suppose i_1 up to i_3 are the only circumstances of evaluation compatible with what John says. Then Porky at i_1, Esther at i_2, and Mary at i_3 are the persons whom John takes himself to be at those circumstances. The *de se* reading of the sentence is therefore true just in case Porky is a hero at i_1, Esther is a hero at i_2, and Mary is a hero at i_3.

We are going to apply Anand and Nevins's elegant proposal to obtain the private standard reading. The truth conditions we are going to aim for is this. The private standard reading of our belief sentence is true at a circumstance of evaluation i if and only if:

(60) For every circumstance of evaluation i' compatible with Mary's beliefs at i, the towels'

degrees of cleanness at i' are at least as high as the contextual standard for 'clean' determined by the $\boldsymbol{s}$-function of i', namely $\boldsymbol{s}_{i'}(\llbracket clean \rrbracket, i')$, and the red towel's degree of cleanness at i' is higher than the blue towel's degree of cleanness at i'.

The following table may help consolidate our intuition about these truth conditions:

(61)

i_1 [0.9]	i_2 [0.8]	i_3 [0.7]
$g(i_1)(r) = 0.95$ $g(i_1)(b) = 0.9$	$g(i_2)(r) = 0.85$ $g(i_2)(b) = 0.8$	$g(i_3)(r) = 0.75$ $g(i_3)(b) = 0.7$

Let's suppose i_1 up to i_3 are the only circumstances of evaluation compatible with Mary's beliefs at some circumstance of evaluation i. The contextual standards for 'clean' determined by the $\boldsymbol{s}$-functions of i_1 up to i_3 are 0.9, 0.8, and 0.7. We can observe that the towels' degrees of cleanness at each circumstance of evaluation are at least as high as the contextual standard at that same circumstance of evaluation. This appears to be an adequate representation of Mary's belief (at i) that the towels are clean by her standard, because those contextual standards are determined by her own belief states, rather than by the context or by the discourse participants. We should notice as well that, in each circumstance of evaluation, the red towel's degree of cleanness is higher than the blue towel's degree of cleanness. So Mary believes that the red towel is cleaner than the blue towel as well. So this table as a whole represents Mary's belief (at i) that the towels are clean by her own standard and that the red one is cleaner.

We will arrive at these truth conditions by diagonalization. Just as we derive the content of what John says from the relevant diagonal of the character of 'I am a hero', we derive the content of Mary's belief from the relevant diagonal of the character of 'both towels are clean, but the red one is cleaner than the blue one'. The relevant diagonal we need is true at a circumstance of evaluation i if and only if the towels' degrees of cleanness at i are at least as high as the contextual standard for 'clean' determined by the $\boldsymbol{s}$-function at i, i.e. s_i. The following table, which shows a finite snapshot of the character of 'both towels are clean, but the red one is cleaner than the blue one', may help consolidate our intuition about what this diagonal is (I trust that the reader can easily fill in the truth values in the cells):

(62)

	i_1 [0.9]	i_2 [0.8]	i_3 [0.7]
c_1 [0.9]	$\boldsymbol{g(i_1)(r) = 0.95}$ $\boldsymbol{g(i_1)(b) = 0.9}$	$g(i_2)(r) = 0.85$ $g(i_2)(b) = 0.8$	$g(i_3)(r) = 0.75$ $g(i_3)(b) = 0.7$
c_2 [0.8]	$g(i_1)(r) = 0.95$ $g(i_1)(b) = 0.9$	$\boldsymbol{g(i_2)(r) = 0.85}$ $\boldsymbol{g(i_2)(b) = 0.8}$	$g(i_3)(r) = 0.75$ $g(i_3)(b) = 0.7$
c_3 [0.7]	$g(i_1)(r) = 0.95$ $g(i_1)(b) = 0.9$	$g(i_2)(r) = 0.85$ $g(i_2)(b) = 0.8$	$\boldsymbol{g(i_3)(r) = 0.75}$ $\boldsymbol{g(i_3)(b) = 0.7}$

The followings are the details of the combinatorics. The typeshifting rule below maps the character of a sentence to a constant character that maps every formal context to the diagonal we need:

(63) *Diagonalization along* $\boldsymbol{s}$*-function:*

If E is a sentence of type $< S, \langle c, \langle i, t \rangle \rangle >$, with meaning χ, then E^D is an expression of the same type, with meaning $\chi^D = \lambda c[\lambda i[\chi(c[\boldsymbol{s}(i)])(i)]]$, where $c[\boldsymbol{s}(i)]$ is just like c except that its $\boldsymbol{s}$-function is replaced by that of i.[30]

I am now going to show the semantic composition for the private standard reading of our belief sentence. To simplify our derivation, instead of using our original belief sentence, we will use the simplified sentence we used in the last subsection, which I repeat here:

(64) Mary believes that Ruby is clean and Tom is clean but that Ruby is cleaner than Tom.

Much of the groundwork has already been done in the last subsection. The reader may focus on the effect of the diagonalization operation on the character of the embedded sentence:

(65) a. $[\![Mary]\!] = \lambda c \lambda i[Mary]$ (likewise for 'Ruby' and 'Tom')

b. $[\![believe]\!] = \lambda c \lambda i \lambda p_{\langle i,t \rangle} \lambda x[\forall i'(i'R(x)(i) \rightarrow p(i') = 1)]$

c. $[\![Ruby\ is\ clean\ and\ Tom\ is\ clean\ but\ Ruby\ is\ cleaner\ than\ Tom]\!]$

[30]Notice that it is a partial diagonalization targeting the standard of precision coordinate only. If this operation targets the SELF coordinate as well, we will arrive at the undesirable result that the extension of 'I' can vary with the doxastic possibilities of a belief subject, which clearly isn't we want if English is our object of study. It is also worth emphasizing that this operation isn't obligatory but optional, or we can't deliver the reading we discussed in the last subsection.

$= \lambda c \lambda i[\llbracket clean \rrbracket(c)(i)(Ruby) \geq s_c(\llbracket clean \rrbracket, i) \wedge \llbracket clean \rrbracket(c)(i)(Tom) \geq s_c(\llbracket clean \rrbracket, i) \wedge$

$\llbracket clean \rrbracket(c)(i)(Ruby) > \llbracket clean \rrbracket(c)(i)(Tom)]$

d. $\llbracket$*Ruby is clean and Tom is clean but Ruby is cleaner than Tom*$\rrbracket^D$

$= \lambda c \lambda i[\llbracket clean \rrbracket(c[\boldsymbol{s}(i)])(i)(Ruby) \geq s_{c[\boldsymbol{s}(i)]}(\llbracket clean \rrbracket, i) \wedge \llbracket clean \rrbracket(c[\boldsymbol{s}(i)])(i)(Tom) \geq$

$s_{c[\boldsymbol{s}(i)]}(\llbracket clean \rrbracket, i) \wedge \llbracket clean \rrbracket(c[\boldsymbol{s}(i)])(i)(Ruby) > \llbracket clean \rrbracket(c[\boldsymbol{s}(i)])(i)(Tom)]$

[Diagonalization (63)]

$= \lambda c \lambda i[\llbracket clean \rrbracket(c)(i)(Ruby) \geq \boldsymbol{s_i}(\llbracket clean \rrbracket, i) \wedge \llbracket clean \rrbracket(c)(i)(Tom) \geq \boldsymbol{s_i}(\llbracket clean \rrbracket, i) \wedge$

$\llbracket clean \rrbracket(c)(i)(Ruby) > \llbracket clean \rrbracket(c)(i)(Tom)]$

[$\llbracket clean \rrbracket$ is a constant character, and $\boldsymbol{s}_{c[\boldsymbol{s}(i)]} = \boldsymbol{s_i}$]

e. $\llbracket$*believe Ruby is clean and Tom is clean but Ruby is cleaner than Tom*$\rrbracket$

$= \lambda c \lambda i \lambda x[\forall i'(i' R(x)(i) \rightarrow \llbracket clean \rrbracket(c)(i')(Ruby) \geq \boldsymbol{s_{i'}}(\llbracket clean \rrbracket, i')$

$\wedge \llbracket clean \rrbracket(c)(i')(Tom) \geq \boldsymbol{s_{i'}}(\llbracket clean \rrbracket, i') \wedge \llbracket clean \rrbracket(c)(i')(Ruby) > \llbracket clean \rrbracket(c)(i')(Tom))]$

f. $\llbracket$*Mary believe Ruby is clean and Tom is clean but Ruby is cleaner than Tom*$\rrbracket$

$= \lambda c \lambda i[\forall i'(i' R(Mary)(i) \rightarrow \llbracket clean \rrbracket(c)(i')(Ruby) \geq \boldsymbol{s_{i'}}(\llbracket clean \rrbracket, i')$

$\wedge \llbracket clean \rrbracket(c)(i')(Tom) \geq \boldsymbol{s_{i'}}(\llbracket clean \rrbracket, i') \wedge \llbracket clean \rrbracket(c)(i')(Ruby) > \llbracket clean \rrbracket(c)(i')(Tom))]$

So here is the extension of our belief sentence at an arbitrary formal context c and an arbitrary circumstance of evaluation i:

(66) $\forall i'(i' R(Mary)(i) \rightarrow \llbracket clean \rrbracket(c)(i')(Ruby) \geq \boldsymbol{s_{i'}}(\llbracket clean \rrbracket, i')$

$\wedge \llbracket clean \rrbracket(c)(i')(Tom) \geq \boldsymbol{s_{i'}}(\llbracket clean \rrbracket, i') \wedge \llbracket clean \rrbracket(c)(i')(Ruby) > \llbracket clean \rrbracket(c)(i')(Tom))$

It is true at a circumstance of evaluation i if and only if, for every circumstance of evaluation i' compatible with Mary's beliefs at i, the towels' degrees of cleanness at i' are at least as high as the contextual standard for 'clean' determined by the $\boldsymbol{s}$-function of i', and the red towel's degree of cleanness at i' is higher than the blue towel's degree of cleanness at i'. These are precisely the truth conditions (60) we desire.

1.5 Possible Extensions

We now review the intended scope of our proposal before we consider its possible applications to other domains.

While our approach to the imprecision of maximal standard absolute adjectives is semantic rather than pragmatic, we haven't introduced any new machinery to model their imprecision. We essentially adopt an *economy-principle-free* approach to the imprecision of maximal standard absolute adjectives (§1.4.1&1.4.3): Once we drop the principle and count on discourse participants' rationality to resolve the contextual standards for maximal standard absolute adjectives to a sufficiently high degree, the imprecise contents of maximal standard absolute adjectives count as their proper semantic contents. Since our aim isn't to develop a novel semantic account of imprecision in general, we are not going to discuss whether or how our approach to the imprecision of maximal standard absolute adjectives can be extended to other cases, such as the imprecision of number terms, and metalinguistic comparison (Morzycki 2011). That said, we will discuss shortly a potential contribution our account can make to the discussions on the imprecision of number terms.

The novel part of our proposal is motivated by the need to account for the private standard reading of our belief sentence, along with its public standard reading, and the contradictory belief reading as a special case of the public standard reading. We pursued an analogy between indexicals in languages such as Amharic and Zazaki and the function variable s introduced by Kennedy's phonologically null morpheme *pos*. We argued that, instead of being given by the context, its possible values are stored in Kaplanian formal contexts and are ruled out or resolved by discourse participants in contexts.

As we discussed in §1.3, a nice feature of our proposal is that our treatment of the function s can readily be applied to accounting for the public-private ambiguity of gradable adjectives in general. We can force the public standard reading by using contexts such as 'not merely believe but know' and 'mistakenly believe'. And contexts such as 'say' and 'believe' tend to be compatible with both the public standard reading and private standard reading:

(67) a. Mary not merely believes but knows that she is tall/ wet/ clean.

b. Mary says/ believes that she is tall/ wet/ clean.

Our proposal offers the degree-based theorist — who holds that each gradable adjective's degree function is more basic than its set-denoting meaning — a way to account for this type of ambiguity. But one potential worry, which must be addressed in future work, is that we are yet to be able to explain why only maximal standard absolute adjectives are imprecise, less susceptible to sorites reasoning, and modifiable by slack regulators such as 'perfectly' and 'absolutely'.

Let's now turn to a potential contribution our proposal can make to the discussions on the imprecision of number terms. According to Sauerland and Stateva (2011), every numerical expression, such as '5m', is interpreted relative to a granularity contextual parameter, which maps its point denotation to an interval containing it.[31] For example, the followings are the extensions of '5m' under different granularity assignment functions:

(68) a. $gran_{fine}(5m) = [4.95m, ..., 5.00m, ..., 5.05m]$

b. $gran_{mid}(5m) = [4.75m, ..., 5.00m, ..., 5.25m]$

c. $gran_{coarse}(5m) = [4.50m, ..., 5.00m, ..., 5.50m]$

They implement their proposal by taking the granularity functions to be interval-assignment functions analogous to variable assignment functions. Slack regulators and approximators such as 'exactly' and 'approximately' can abstract over the function granularity variable of an expression such as '5m' and select for it an appropriate granularity function. For example, 'exactly' forces '5m' to accept the granularity function $gran_{fine}$.

We won't get into the difficult topic of the semantics of slack regulators and approximators. Instead, we consider the extent to which it is fruitful to absorb Sauerland and Stateva's proposal about the granularity contextual parameter into our framework. The following sentence is assertable because the content of Mary's belief seems to be determined by the discourse participants' own standard of precision:

[31] See also Krifka (2007).

(69) Mary mistakenly believes the dinosaur is one million year old. It is 1 million year and three days to be exact.

We will set aside the pragmatic approach on which the sentence 'Mary believes that the dinosaur is one million year old' is almost always an attribution of a false belief to Mary. We follow Sauerland and Stateva in adopting the semantic approach. But instead of positing a granularity function assignment, we posit a granularity function variable $\boldsymbol{g}$ analogous to Kennedy's function variable $\boldsymbol{s}$. A 'contextually given' value for the function variable will map the point denotation of 'one million year', which we can assume is a point on the scale of the adjective 'old', to a set of points surrounding the point denotation on that scale. We will store the possible values of the $\boldsymbol{g}$ function in formal contexts, just as we do for the possible values of the variable $\boldsymbol{s}$. The resulting meaning of 'The dinosaur is one million year old' will be a character, such that each of its contents is based on a possible value for the granularity function $\boldsymbol{g}$.

Let's me spell out this idea in more detail. We will introduce a typeshifting rule $GRAN$ that maps the point denotation of a numerical expression to a function that maps the basic meaning of a gradable adjective to an appropriate positive form meaning, which is true of an object if that's object degree, as measured by the adjective, is a member of the relevant set of degrees surrounding the point denotation:

(70) $GRAN(N) = \lambda c \lambda i \lambda g_{\langle ci,ed \rangle} \lambda x[g(c)(i)(x) \in \boldsymbol{g_c}(N)]$, where N is the point denotation of a numerical expression, and g is the basic meaning of a gradable adjective, and $\boldsymbol{g_c}$ is the granularity function of the formal context c, which maps each numerical expression's point denotation to a set of points containing that point denotation.

This is the semantic composition for 'Mary believes that Rocky the dinosaur is one million year old':

(71) Lexical items

a. $[\![Mary]\!] = \lambda c \lambda i[Mary]$

(likewise for 'Rocky the dinosaur', which we simplify as 'Rocky')

b. $[\![one\ million\ year]\!] = \lambda c \lambda i[d_{million}]$

c. $[\![old]\!] = \lambda c \lambda i \lambda x[x$'s degree of oldness at the world and the time coordinates of $i]$

d. $[\![believe]\!] = \lambda c \lambda i \lambda p_{\langle i,t \rangle} \lambda x[\forall i'(i'R(x)(i) \rightarrow p(i') = 1)]$

(72) a. $[\![one\ million\ year]\!]^{GRAN} = \lambda c \lambda i \lambda g_{\langle ci,ed \rangle} \lambda x[g(c)(i)(x) \in \boldsymbol{g_c}(d_{million})]$

[Typeshifting rule *gran* (70)]

b. $[\![one\ million\ year\ old]\!] = [\![one\ million\ year]\!]^{GRAN}([\![old]\!])$

$= \lambda c \lambda i \lambda x[[\![old]\!](c)(i)(x) \in \boldsymbol{g_c}(d_{million})]$ [extension F.A.]

c. $[\![Rocky\ is\ one\ million\ year\ old]\!] = [\![one\ million\ year\ old]\!]([\![Rocky]\!])$

$= \lambda c \lambda i[[\![old]\!](c)(i)(Rocky) \in \boldsymbol{g_c}(d_{million})]$ [extension F.A.]

d. $[\![believe\ Rocky\ is\ one\ million\ year\ old]\!] = [\![believe]\!]([\![Rocky\ is\ one\ million\ year\ old]\!])$

$\lambda c \lambda i \lambda x[\forall i'(i'R(x)(i) \rightarrow [\![old]\!](c)(i')(Rocky) \in \boldsymbol{g_c}(d_{million}))]$ [content F.A.]

e. $[\![Mary\ believe\ Rocky\ is\ one\ million\ year\ old]\!]$

$= [\![believe\ Rocky\ is\ one\ million\ year\ old]\!]([\![Mary]\!])$

$= \lambda c \lambda i[\forall i'(i'R(Mary)(i) \rightarrow [\![old]\!](c)(i')(Rocky) \in \boldsymbol{g_c}(d_{million}))]$

Let's suppose the salient margin of error for 'one million year' is a day from one million year. Our uncertainty about the function $\boldsymbol{g_c}$ will be characterized by a set of formal contexts each of which contains a granularity function that maps 'one million year' to the interval (1.m.y. - 0.5d, 1.m.y. + 0.5d). The proposition that we obtain based on our uncertainty about the function $\boldsymbol{g_c}$ is true at a circumstance of evaluation i if and only if, for every circumstance of evaluation i' compatible with Mary's beliefs at i, the dinosaur's degree of oldness at i' is a member of the set of points determined by $\boldsymbol{g_c}$. So we seem to have an adequate representation of the meaning of the sentence.

While the public standard reading is certainly possible for numerical expressions in belief contexts, I don't have a strong intuition that a private standard reading analogous to that of gradable adjectives exists. It is not clear whether the following sentence has a private standard reading, or, if it does, how we can make that reading salient:

(73) Mary the paleontologist believes that the dinosaur is one million years old.

Is there a reading on which the discourse participants defer to Mary's own standards of precision, given that she has better knowledge about dinosaurs than they do? If there is, we can obtain that reading by diagonalization. But since my intuitions on this aren't very clear, I don't claim that extending our approach to numerical expressions is maximally fruitful.

1.6 Conclusion

I have defended a version of the relativist view on the positive-form meanings of maximal standard absolute adjectives. I have argued that, instead of taking the imprecise contents of maximal standard absolute adjectives in their positive form to be the contents we pragmatically convey, we should treat them as their semantic contents, and that we can represent those contents by enriching the double-indexing framework with a s-function coordinate. I motivate my view with the puzzle that the sentence 'Mary believes that both towels are clean but that the red one is cleaner than the blue one' can be used to attribute coherent beliefs to Mary, and that the sentence has three readings: the contradictory belief reading, the public standard reading, and the private standard reading. The main consideration in favor of my view is that it provides a plausible solution to the puzzle. The extent to which it is fruitful to extend my approach to other cases of imprecision, such as the imprecision of number terms and metalinguistic comparison, must be assessed in future work.

Chapter 2

Binding without Binding Assumption:

From Overgeneration to Simplicity

2.1 Introduction

According to Stanley, that there are bound readings of a certain expression provides good evidence for the presence of an individual variable in the syntactic structure of that expression. Let call this *the binding assumption*.[1]

This assumption plays an important role in Stanley and Szabo (2000a)'s argument that there is an individual variable in generalized quantifiers, such as 'every bottle', and in definite descriptions, such as 'the corner'. For example, based on the bound readings of 'every bottle' and 'the corner' in the following sentence, they argue that there exists an individual variable and a function variable in both 'every bottle' and 'the corner':

(1) a. In every room in John's house, every bottle is in the corner (Stanley and Szabo 2000a, example #25).[2]

b. In every room x in John's house, [every bottle $f(x)$] is in [the corner $g(x)$].

[1]Here are some of Stanley's formulations of the binding assumption: 'If α and β are within the same clause, and α semantically binds β — that is, 'the interpretation of β systematically depends on the values introduced by α' — then α either is, or introduces, a variable-binding operator which is co-indexed with, and stands in a certain specified structural relation to, a variable which is either identical to, or is a constituent of β' (Stanley 2000, p.412). '[I]f there is a genuine bound reading of a certain construction, that supports the hypothesis that the quantifier in question binds a variable in the syntactic structure of the sentence' (Stanley 2007b, p.213). To avoid lengthy exegesis of Stanley's formulations, we adopt a more informal formulation of the binding assumption. Nothing in our discussion turns on how finely the binding assumption is formulated. The crucial commitment of the assumption is that the bound readings of an expression supports an existence claim about individual variable(s).

[2]Examples like this were given by Heim (1991) and von Fintel (1994).

On the bound reading of the sentence above, the bottles and the corner referred to covary with the rooms in John's house, so 'every bottle' and 'the corner' can be paraphrased as 'every bottle in that room' and 'the corner of that room' respectively. Stanley and Szabo argue that we obtain that reading because the individual variables in 'every bottle' and 'the corner' are bound by 'every room in John's house', and the function variables in them (i.e. f and g) are both saturated by a function that maps a room to the things in it. The following gloss captures the truth conditions they assign to (1-a):

(2) For every room x in John's house, every bottle in x is in the corner of x.

Let's call Stanley and Szabo's approach *the grammatical approach*, because it says that the truth conditions of sentences containing generalized quantifiers and definite descriptions result directly from their syntax and semantics. It is opposed to *the pragmatic approach*, which says that 'every bottle' and 'the corner' literally mean every bottle in the world and the only corner in the world respectively, and that there are no individual or function variables in 'every bottle' and 'the corner' that ask for qualifications on their meanings. On this view, the reason why we can typically understand what a speaker means by uttering them is that we are able to read the relevant qualifications into their semantic contents. The following gloss shows how one might obtain the bound reading of (1-a) by reading qualifications into the semantic contents of 'every bottle' and 'the corner':

(3) In every room in John's house, every bottle (in that room) is in the corner (of that room).

The major benefit of the pragmatic approach is that, by doing away with variables, it keeps the syntax and the semantics maximally simple.[3] It is therefore no surprise that its proponents revel in data that could be used to show that the grammatical approach overgenerates variables and thereby commits itself to a syntax and a semantics that are unduly complicated. In this paper, we will focus on two types of such data. The first type of data are based on the unbound instances of generalized quantifiers and definite descriptions where the individual variables in

[3] See Stanley and Szabo (2000a, p.240), and Bach (2000, p.269) who agrees with Stanley and Szabo's assessment.

them appear unexploited:

(4) a. Every man is mortal.

b. The prime numbers are numbers which are only divisible by 1 and themselves.

This is why these sentences appear to show that the grammatical approach overgenerates variables. Since there are sentences where 'every man' and 'the prime numbers' are bound (e.g. 'Every government serves every man'; 'Every examinee doesn't enjoy seeing the prime numbers'), it follows from Stanley's binding assumption that there are individual variables (and function variables) in them. However, 'every man' and 'the prime numbers' are unbound in these sentences, and no qualifications on their meanings seems necessary because the sentences seem to be general claims about all men and about all prime numbers. So the variables in 'every man' and 'the prime numbers' seem idle in explaining what the sentences convey.

The second type of data are based on the bound instances of generalized quantifiers and definite descriptions where they appear to be bound by more than one generalized quantifiers:

(5) a. Every professor believes that every student misses some assignment (she assigns her).

b. Every professor asks every student to enjoy the assignment (she assigns her).

Under suitable contexts, 'some assignment' and 'the assignment' are each bound by both 'every professor' and 'every student'. Apparently, this isn't possible if there is only one individual variable in them. So, to be consistent with their explanation for the bound reading of (1-a), it seems Stanley and Szabo would have to posit in the common noun 'assignment' a two-place function variable, which is to be saturated by a function that maps a pair of individuals to the assignment the first individual assigns to the second individual, and two individual variables, such that the first individual variable is bound by 'every professor' and the second individual variable is bound by 'every student'. The following glosses show how the two individual variables in 'assignment' contribute to the truth conditions of the sentences:

(6) a. Every professor x believes that every student y misses some assignment x assigns y.

b. Every professor x asks every student y to enjoy the assignment x assigns y.

However, since there are now at least two individual variables in 'some assignment' and 'the assignment', at least one individual variable becomes unexploited when they occur in sentences where they only have a single binder, and when they occur unbound in sentences like (4). So the grammatical approach appears to once again overgenerate variables.

Throughout his work, Stanley insists that these overgeneration worries are ill-founded. He dismisses the worry based on the first type of data by arguing that the individual and the function variables are always exploited,[4] and the worry based on the second type of data by arguing that the generalized quantifiers and the definite descriptions are only bound by a single generalized quantifier, despite appearance to the contrary.[5] However, I believe that these worries are well-founded. But this is no good news to the pragmatic approach, because I am going to argue that the grammatical approach has enough resources to respond to the worries, and that a more adequate response to them strengthens its case against the pragmatic approach.

Here is how we going to proceed. After arguing that the overgeneration worries are well-founded, I will argue that there are data about gradable adjectives that are parallel to those driving the worries, and discuss a proper response to them, which will become a model for addressing the worries. Here are the relevant data:

(7) Every Pentium processor is slow (by today's standard).

(8) Every running team demands that no member is too slow (for the level at which the team competes and for the member's age and gender group).

This is why these data are parallel to those driving the overgeneration worries. Since there are sentences where 'slow' is bound (e.g. 'Every team has members who are slow'), Stanley's binding assumption suggests that there is an individual variable in 'slow', which we can think

[4]See Stanley and Szabo (2000a, 2000b)

[5]See Stanley's response to Breheny's objection (2007b, pp.222-225) and his response to Jacobson's objection (2007a, pp.251, fn.2).

of as a controller of the relevant comparison class for the adjective. However, even without much context, 'slow' in (7) isn't bound, because it conveys that every Pentium processor falls below the same standard set by the context, such as a certain salient clock speed. This means that whether or not each Pentium processor has its own comparison class plays no essential role in explaining what the sentence conveys. So the individual variable in 'slow' seems unexploited. The second example parallels the second type of data because, under a context which I will discuss in some length, 'too slow' is bound by both 'every running team' and 'no member'.

The reason for this detour to gradable adjectives is twofold. First, Kennedy (2007)'s account of sentences like (7) provides us with an important idea which we can adapt to respond to the first overgeneration worry. Here is the upshot. Instead of positing a comparison class variable in 'slow' to account for its bound readings, he introduces a comparison class argument in it by typeshifting it (i.e. changing its syntactic and semantic types in a rule-governed manner). This means that, instead of following Stanley's binding assumption, we can use a typeshifting rule to introduce the materials required for binding only when they are needed. This idea is precisely what we need to address the first overgeneration worry.

Second, Kennedy's important idea, as we will see, is in a natural alliance with Jacobson (1999)'s variable-free account of binding; the latter will play a key role in our response to the second overgeneration worry. Here is the significance of the variable-free approach to the debate between the grammatical approach and the pragmatic approach. First, instead of building the structures required for binding into the lexical meanings of bindable expressions in anticipation of the most complicated binding scenario, it accounts for their bound readings by using an incremental approach, which ensures that the expressions bound into always have just enough structures to be bound by the binders preceding it. In short, it keeps the syntactic structures and the meanings of the expressions bound into maximally simple. So this approach to binding is precisely what Stanley and Szabo need to block the second overgeneration worry. And the good news for them is that they are fully entitled to use it, because, since their main contention is that a sentence's truth conditions result directly from its syntax and semantics, they can be neutral between their variable-ful approach and the variable-free approach.[6] Second,

[6]Stanley appears to be neutral between the variable-ful approach and the variable-free approach. See his (2007a

the variable-free approach makes the striking prediction that if gradable adjectives, definite descriptions, and generalized quantifiers can be bound, they can be bound by two generalized quantifiers, because the doubly bound readings of those expressions can be accounted for using the same principles that account for the singly bound cases. The ability to make these precise predictions can't be matched by the pragmatic approach as it stands, because not only does it fail to predict which expressions can be bound, it also fails to explain the connection between the singly-bound cases and the doubly-bound cases, or so I argue.

The rest of our discussion is structured as follows. In the next section (§2.2), we will review Stanley and Szabo's account of domain restriction, and argue, *pace* Stanley, that the overgeneration worries are well-founded. After that, I will discuss data about gradable adjectives that are parallel to those driving the overgeneration worries, and propose an account of them that is based on Kennedy's work on gradable adjectives and Jacobson's on binding (§2.3). I then discuss how we can adapt that account to respond to the overgeneration worries and how our response strengthens the case against the pragmatic approach (§2.4). We conclude our discussion in the last section (§2.5).

2.2 Two Well-founded Overgeneration Worries

2.2.1 Stanley and Szabo on domain restriction

Based on von Fintel (1994)'s work on quantifier domain restriction, Stanley and Szabo (2000a) argue that the bound reading of (1-a), repeated here as (9-a), is best accounted for by positing in each common noun a function variable, which is assigned an one-place function from individuals to sets relative to a context, and an individual variable, which is assigned an individual relative to a context:

(9) a. In every room in John's house, every bottle is in the corner.

b. In every room x in John's house, [every [bottle $f(x)$]] is in [the [corner $g(x)$]].

c. $[\![bottle\, f(x)]\!]^{\sigma} = \lambda z[bottle'(z) \wedge z \in \sigma(f)(\sigma(x))]$, where σ is a context-dependent

p.215; pp.250-251) for his discussions on the variable-free approach.

assignment function that maps an individual variable to an individual, and a function variable to an one-place function from individuals to sets.

d. $[\![corner\, g(x)]\!]^{\sigma} = \lambda z[corner'(z) \wedge z \in \sigma(g)(\sigma(x))]$, where σ is a context-dependent assignment function that maps an individual variable to an individual, and a function variable to an one-place function from individuals to sets.

This is how their account explains the bound reading of (9-a). (9-c) and (9-d) show how the variables located at the common nouns 'bottle' and 'corner' constrain their interpretations. Since they are parallel to each other, we will only walk through (9-c) in detail: It says that an object falls into the extension of 'bottle $f(x)$' relative to a context just in case it falls into the intersection of the extension of 'bottle' and the set which is the result of applying the value of f in that context to the value of x in that context.

This is how their analysis of 'bottle' and 'corner' interacts with the clause 'in every room in John's house'. As we can see in (9-b), the generalized quantifier 'every room in John's house' binds the individual variable x in the common nouns 'bottle' and 'corner', while the function variables f and g are left free. So when the assignment function σ (whose value is determined by the context) maps both function variables to the function that maps each room to the objects in that room, we obtain the reading that says that, for every room in John's house, every bottle in that room is in the corner of that room. These seem to be the truth conditions we want for the bound reading of (9-a).

We should note that Stanley and Szabo proposed their account as a competitor to von Fintel (1994)'s account. According to von Fintel, the function variable and the individual variable are located at the determiners (e.g. 'every', 'the') rather than at the common noun, and it can account for the bound reading of (9-a) equally well.[7] Although I favor von Fintel's account,[8] I

[7] According to von Fintel, the function variable has an arbitrary arity, and the number of individual variables located at the quantifier is identical to the arity of the function variable (1994, p.31). So he does not suggest, as Stanley and Szabo do, that the function variable is always one-place. However, he has given no examples where the function variable has to be at least two-place.

[8] Consider:

(i) a. [Every $f(x)$] fake philosopher is from Disneyland (von Fintel).
b. Every fake [philosopher $f(x)$] is from Disneyland (Stanley and Szabo).

If we use von Fintel's account to obtain the reading that every contextually salient fake philosopher is from Disneyland, we can have 'fake philosophers' pick out the set of fake philosophers, and have that set intersect with the set of am going to remain neutral between his account and Stanley and Szabo's in this discussion, because both are vulnerable to the overgeneration worries we are about to discuss.

2.2.2 The first overgeneration worry: the challenge from unbound instances

Kent Bach (2000), who is the perhaps the most vocal advocate of the pragmatic approach,[9] raises an interesting overgeneration worry against Stanley and Szabo's grammatical approach.[10] Based on the following examples, he argues that the individual variables and the function variables in the common nouns often play no essential role in explaining what sentences containing them convey:

(10) a. All [men $f(i)$] are mortal.

b. There are more [ants $f(i)$] than [mosquitoes $f(i)$].

His worry seems to be that, since both sentences convey universal claims about men, ants, and mosquitoes in general, no qualifications on the meanings of 'men', 'ants', and 'mosquitoes' are necessary, and the variables in the common nouns seem idle in explaining what the sentences convey.

While I am sympathetic with his general worry, since he doesn't provide the contexts for his examples, Stanley and Szabo can easily construct contexts in which qualifications on the meanings of the common nouns are essential for explaining what the sentences convey: Suppose (10-a) is uttered by a reporter in the world of Spiderman, who is informed of the immortality of various superheroes. In order to interpret the reporter's utterance as a true claim about her world, it is arguable that we need to exclude the immortal superheroes from the domain of the generalized quantifier 'all men' (since they are fairly good examples of immortal

contextually salient fake philosophers, which is picked out by $f(x)$ in the determiner 'every'. But if we adopt Stanley and Szabo's approach, we can't have $f(x)$ pick out a subset of the set of (real) philosophers, because the meaning of 'fake' can't map that subset of (real) philosophers to the contextually salient set of fake philosophers. It won't do to have $f(x)$ pick out a contextually salient set of fake philosophers either, because the result of intersecting that set with the set of (real) philosophers is the empty set, which, of course, can't be mapped to the contextually salient set of fake philosophers by the meaning of 'fake'. Similar worries have been raised by Breheny (2003).

A further problem with Stanley and Szabo's approach is that common nouns that aren't immediately preceded by determiners are not bindable. Consider, for example, 'Every man loves philosophers $f(x)$'. If there is an individual variable in 'philosopher', we should expect that it can be bound by 'every man'. But the sentence doesn't seem to have a bound reading.

[9] See also his (1994).

[10] A similar objection has also been raised by Breheny (2003).

men). With his second example (10-b), even without much context, what we take it to convey naturally depends on the relevant location the claim is about, because it may be a false claim about the entire universe and about a certain country, but a true claim about the earth. So it is arguable that qualifications on the meanings of 'ants' and 'mosquitoes' are essential for explaining what the sentence conveys.

To block these possible responses, Bach might want to use the following examples instead:

(11) a. Every [prime number $f(i)$] is divisible by 1 and itself.

b. The [prime numbers $f(i)$] are numbers which are divisible by 1 and themselves.

If a teacher uses these sentences in a mathematics class to define what the prime numbers are, I can't think of any reason to restrict the domains of the generalized quantifier 'every prime number' and the definite description 'the prime numbers' — the domain-restricted interpretations are bad for her pedagogical purpose because they imply that the statements aren't true of some number outside of the domain. So I agree with Bach that there are unbound uses of generalized quantifiers and definite descriptions where the variables in them are unexploited.

2.2.3 The second overgeneration worry: the challenge from double or multiple binders

Generalized quantifiers

Breheny, who is no proponent of the pragmatic approach, is the first author who raised an overgeneration worry against Stanley and Szabo's account based on the bound readings of generalized quantifiers.[11] One of his examples is the following:

(12) Every student was feeling particularly lucky and thought no examiner would notice **every mistake** (2003, p.63, example #23a).

[11] He also raised some important worries about placing the variables at the common noun. See also the postscript of Stanley (2007a) for his most recent view on this issue.

He describes the context very briefly as: '[S]tudents write a number of papers which are each marked by three examiners' (2003, p.63). Since I don't think that the exact number of examiners for each student matters in his example, we will make the context as simple as possible by assuming that each student writes two papers and submits each of them to a different examiner, and that there are only two examiners for all students (let's say there are 20 of them). Let's also make the example more vivid by imagining a student Socrates,[12] who answers ten questions in each of his two papers, and makes nineteen mistakes in total. According to Breheny's intended reading of the sentence, since Socrates was feeling rather lucky, he thought that none of his examiners would notice every mistake in the paper he submits to her. So 'every mistake' doesn't refer to all the nineteen mistakes he makes (since none of his examiners can notice the mistakes he makes in the paper she doesn't mark). Nor does it refer to all the mistakes each examiner will notice as she marks the papers she receives from the students (there must be plenty of them if Socrates is already above average). It refers specifically to the nine or ten mistakes in one of his papers.

Breheny's concern is that, if his example does have his intended reading, then Stanley and Szabo ought to posit two individual variables and a two-place function variable in the common noun 'mistake', so that the first and the second individual variables are bound respectively by 'every student' and 'no examiner', and the two-place function variable can be assigned the two-place function that maps a student and an examiner to the mistakes the student makes and are examined by the examiner. Of course, his concern is meant to suggest that even a two-place function variable isn't sufficient, because there may be examples which demand function variables of multiple places.[13] The significance of his concern to the debate between the grammatical approach and the pragmatic approach is that, if the function variable is at least two-place, then at least one individual variable is unexploited when the common nouns are singly bound or unbound.

[12] 'Unexamined life is not worth living' (Plato's Socrates).

[13] Here is a potential example:

(i) Every CEO asks every vice president to order every regional director to fire some managers.

In response, Stanley contends that we can obtain Breheny's intended reading once we observe that the noun 'examiner' in his example isn't a common noun, but a relational noun having the form of 'examiner-of x'. He suggests that the following gloss captures Breheny's intended reading, and he makes explicit for us the requisite value for the one-place function variable in the common noun 'mistake':

(13) a. Every student x thought no examiner of x would notice every mistake made on a paper x turned in and the examiner of x examines (2007b, p.223, example #35, modified for readability).

b. 'We may assume 'f' is assigned a function from students to their exam questions. So we can straightforwardly predict a reading of (12) according to which every student thought no examiner of that student would notice every mistake on that student's exam' (2007b, p.223).

For the moment, we will set aside the issue whether it is helpful to assume that 'examiner' is a relational noun — it isn't as I'll explain shortly — and focus on the main problem with Stanley's response: It appears that he misses the context Breheny provides and hence the intended reading of his example.[14] The reading Stanley assigns to the generalized quantifier 'every mistake' can be paraphrased as 'every mistake in *his* paper that is examined by *his* examiner', which suggests that we can obtain that reading by having the one-place function variable in 'mistake' be saturated by the function that maps each student to the exam questions (or to the mistakes, more specifically) in his paper that are examined by his examiner, or more simply, by the function that maps each student to his exam questions (or his mistakes, more specifically), just as Stanley suggests. However, according to the context Breheny provides, each student writes at least two papers, and has at least two examiners. If the requisite one-place function is the 'function from students to their exam questions' as Stanley suggests, then it maps Socrates to the twenty questions on his two papers, and 'every mistake' refers to the intersection of the set of mistakes (which, of course, contains Socrates nineteen mistakes) and

[14]This is understandable because Breheny's description of the context is quite brief.

the set containing his twenty questions. If we follow Stanley in assuming that an incorrectly answered question is a mistake,[15] that intersection contains his nineteen mistakes, rather than the nine or ten mistakes in one of his papers. So the reading Stanley provides isn't identical to Breheny's intended reading.

Since part of the confusion surrounding Breheny's example results from its artificiality, let's confirm that his worry is well-founded by using a more natural example:

(14) No semantics teacher finds it surprising that every student turns in some assignment late.

Let's suppose that every semantics teacher has a rather lenient late work policy, but they don't know about other teachers' late work policies. The sentence we constructed naturally conveys what each semantics teacher would expect at the beginning of a semester: Every student of hers would turn in some assignment(s) she assigns him/her late because of her lenient late work policy. It is important to add that, since each student may receive different assignments from the same teacher based on their interests and abilities (e.g. there are both undergraduate and graduate students in each class), we do not assume that the students in a given semantics class receives the same set of assignments from their teacher.

Can we construct an one-place function for 'some assignment $g(i)$' that can deliver the intended reading? If 'some assignment' is bound by 'no semantics teacher' only, we clearly can't have the relevant set of assignments vary with the students. So this option is out. So let's consider the option of having 'some assignment' be bound by 'every student' only. Let's assume for the sake of argument that 'student' is a relational noun. So it now has the form 'student-of x'. If the variable x is left free and saturated by the context, then we get an odd reading on which the relevant students referred to by 'every student' need not be taught by the semantics teachers. So, instead of being left free, the variable should be bound by 'no semantics teacher', so that 'every student' can now be paraphrased as 'every student of that semantics teacher'. So far so good. But if the individual variable i in 'some assignment $g(i)$' is only bound

[15] I think Stanley could have described the function more clearly as 'the function that maps each student to his mistakes in his paper'.

by 'every student', then we run into the following problem. Suppose Aristotle loves semantics so much that he is enrolled in two semantics classes, which are taught by two different teachers. According to the intended reading of our example, each of Aristotle's teachers expects that he turn in some assignment she assigns him late — rather than that he turn in some assignment for any of his classes late, because she doesn't know about other teachers' late work policies. But if 'some assignment $g(i)$' is only bound by 'every student', we can't have the relevant domain of 'some assignment' vary with his teachers. So an one-place function variable and a single individual variable in 'assignment' are not enough. In order to allow the relevant domain of 'some assignment' to vary with his teachers, we need to posit in the common noun 'assignment' a two-place function variable (which is to be saturated by the function that maps a pair of individuals to the assignments the first individual assigns to the second individual) and two individual variables, so that the generalized quantifier 'some assignment' can be bound by both 'no semantics teacher' and 'every student'. So we conclude that the assumption that 'student' is a relational noun doesn't help Stanley respond to Breheny's worry.

For the sake of completeness, let's drop the assumption that 'student' is a relational noun. 'Student' is a common noun with the form 'student $f(y)$'. If we allow the individual variable y to be free and saturated by the context, we will run into a problem similar to the one we just discussed: The students need not be taught by the semantics teachers. So the variable y should be bound by 'no semantics teacher', and f should be saturated by a function that maps a teacher to her students, so that we obtain the desired reading for 'every student', which can be paraphrased as 'every student taught by that semantics teacher'. But if the individual variable i in 'some assignment $g(i)$' is still meant to be bound by 'every student' only, then we run into the Aristotle problem we discussed above again: We can't have the relevant domain of 'some assignment' vary with his semantics teachers. So Breheny's overgeneration worry seems well-founded, whether or not 'student' is a common noun or a relational noun.

Definite descriptions

To follow up on Breheny's worry, Jacobson, who is certainly not a proponent of the pragmatic approach, raises the following counterexample to Stanley and Szabo's assumption that a single individual variable and an one-place function variable in common nouns suffice for accounting for all bound readings of definite descriptions:

(15) Consider a multidisciplinary conference of linguists, philosophers, and psychologists, with an equal number of participants from each field. To maximize interaction among the participants, each linguist is assigned to pour a drink for one and only one philosopher at the party, and also one and only one psychologist. And each philosopher will pour a drink to their assigned linguist and to their assigned psychologist, and similarly for the psychologists. It turns out to be a particularly bad year for jobs for recent PhDs in philosophy. And so the philosophers at the conference — all being philosophers of language — hatch a plot: they decide [to] kill off the linguists, thereby creating some new openings. And so each philosopher serves a poisoned drink to her/his appointed linguist (a foolproof plot, for they can always blame the psychologists for the mass die-off of linguists). But the plot failed: the [linguists] figured it out ... No philosopher managed to get any linguist to actually accept **the drink** (Jacobson 2014, pp.380-381).[16]

For ease of exposition, let's assume that we are in a strange world where all philosophers are male, and all linguists are female. Given the elaborate context Jacobson provides, 'the drink' is naturally understood as 'the drink *he* pours for *her*'. This seems to suggest that Stanley and Szabo need to posit a two-place function variable in the common noun 'drink', so that it can be assigned the function that maps a pair of individuals to the drink poured by the first individual for the second individual.

In response, Stanley rightly points out that, since each philosopher pours a drink for the linguist assigned to him, we can obtain the desired truth conditions for Jacobson's example by having the individual variable x in 'the drink $f(x)$' be bound by the generalized quantifier 'no philosopher', and by having the function variable f be saturated by the function that maps a philosopher to the drink he pours for the linguist assigned to him. A more informal way of explaining Stanley's suggestion is that we can obtain Jacobson's intended reading by paraphrasing 'the drink' into 'the drink *he* pours for the linguist assigned to *him*'. This paraphrase clearly suggests that an one-place function (i.e. the function that maps an individual to the drink he pours for the linguist assigned to him) suffices for delivering Jacobson's intended reading.[17]

Stanley could also obtain Jacobson's intended reading by having the individual variable x in

[16]Stanley's response is based on an example given in Jacobson (2006), to which I do not have access. I believe that the example quoted here is the same as the one Stanley responded to.

[17]See Stanley (2007a, p.251).

'the drink $f(x)$' be bound by the generalized quantifier 'any linguist', and by having the function variable f be saturated by the function that maps a linguist to the drink she receives from the philosopher who pours a drink for her (whom we will refer to as her 'philosopher drink pourer'). The intuition behind this alternative strategy is that 'the drink' can be understood as 'the drink *she* receives from *her* philosopher drink pourer'.

However, it appears that Stanley's strategy only works if every philosopher pours a drink for a single linguist, and if every linguist receives a drink from a single philosopher. Suppose we modify Jacobson's example slightly by requiring that every philosopher pours a poisoned drink for two linguists (and pours a normal drink for a single psychologist), and that every linguist receives a drink from two philosophers (and a normal drink from a single psychologist). It seems we can still get the reading on which 'the drink' can be paraphrased as 'the drink *he* pours for *her*' or 'the drink *she* receives from *him*'. But since every philosopher pours a drink for two linguists, and every linguist has two philosopher drink pourers, the paraphrases 'the drink he pours for the linguist assigned to him' and 'the drink she receives from her philosopher drink pourer' are no longer apt. This suggests that Stanley and Szabo can't avoid positing a two-place function variable and two individual variables in 'the drink' in order to account for its intended reading.[18] This appears to be good news for the pragmatic approach, because at least one individual variable is unexploited when 'the drink' is only bound by a single generalized quantifier.

[18] I believe that we can construct examples that are more natural than Jacobson's. For example, consider Schlenker's example below:

(i) Every Dean asked every part-time instructor not to give an A to a majority of **the students** (Schlenker 2005, p.62, example #101).

Each Dean is clearly not paired with a single part-time instructor. Given our background knowledge about the administrative structure of a college, even in the absence of an elaborate context, we can still get a reading on which the relevant students referred to by 'the students' are crucially a function of both the Deans and the part-time instructors (e.g the students who belonged to a given Dean's college and were taught by a given part-time instructor). Since each Dean is not paired with a single part-time instructor, we can't use Stanley's strategy to construct an one-place function that delivers the intended reading.

Here is another possible example, which is similar to the example I constructed for generalized quantifiers (§2.3.1):

(ii) Every semantics teacher asks every student to enjoy the assignments.

Even without much context, 'the assignments' can be understood as 'the assignments s/he assigns him/her'. To ensure that a two-place function and two individual variables are essential for delivering the intended reading, we can assume that each student need not receive the same set of assignments from the same teacher, and that a student can be enrolled in multiple semantics classes, and that the same semantics teacher can offer multiple semantics classes.

But we will see in the next two sections that it is not.

2.3 Some Parallels between Domain Restriction and Adjectival Domain Restriction

I just argued that the overgeneration worries against Stanley and Szabo's grammatical approach to domain restriction are well-founded. In this section, I argue that there are data about gradable adjectives that are parallel to those driving those worries. Not only are these data interesting in their own right, our account of them, which will be based on both Kennedy's work on gradable adjectives and Jacobson's on binding, will also provide us with a model for responding to the overgeneration worries. We will proceed by first discussing Kennedy's account of gradable adjectives and how he would explain the truth conditions of 'Every Pentium processor is slow'. I will then provide an example where 'slow' is plausibly bound by two generalized quantifiers (§2.3.2), and explain why this phenomenon is predicted by the result of combining a typeshifting principle posited by Kennedy, which we'll call 'the k-rule', and Jacobson's variable-free approach to binding (§2.3.3).

2.3.1 Kennedy on adjectival domain restriction

According to Kennedy (2007), the meaning of a gradable adjective is a measure function that maps the objects in its domain onto a set of degrees. The set of degrees is naturally thought of as a scale, which is essentially a totally ordered set of points along a certain dimension, such as height and speed. For example, the meaning of 'slow' is a function that maps Mary and Paul to their speeds (or degrees of slowness) on the slow scale:

(16) $[\![slow]\!] = \lambda x[slow'(x)] = \lambda x[x\text{'s degree of slowness}]$

This meaning gives us a very straightforward truth condition for the comparative 'Mary is slower than Paul'; the sentence is true just in case Mary's degree of slowness is higher than Paul's degree of slowness. In order to account for the truth conditions of the positive form such as 'Mary is slow', Kennedy proposes that the positive form of a gradable adjective results from the combination of the meaning of a phonologically null morpheme, *pos* (for 'positive form'), and the basic meaning of the adjective. The intuition behind his proposal is that the *pos* morpheme introduces a contextual threshold, such that if an object's degree (e.g. its degree of slowness) that is measured by a certain gradable adjective (e.g. 'slow') is at least as high as that threshold, that object counts as satisfying the positive form of the adjective. The following example illustrates how Kennedy would covert the basic meaning of 'slow' into its positive-form meaning:

(17) a. $[\![slow]\!] = \lambda x[slow'(x)]$

b. $[\![pos]\!] = \lambda g[\lambda x[g(x) \geq \boldsymbol{s}(g)]]$, where $\boldsymbol{s}$ is a contextually given function that maps a gradable adjective g to its contextual threshold $\boldsymbol{s}(g)$.

c. $[\![pos]\!]([\![slow]\!]) = \lambda x[slow'(x) \geq \boldsymbol{s}([\![slow]\!])]$

[(a)&(b), Function Application]

Here is why the semantics of gradable adjectives and von Fintel's and Stanley and Szabo's work on domain restriction are interconnected. Kennedy observes that the domain of an adjective's measure function can be contextually restricted in a way parallel to the domains of generalized quantifiers and definite descriptions. For example, the domain of the measure function of 'small' can be restricted either implicitly or explicitly by a for-phrase such as 'for an elephant'.

(18) a. Jumbo is small.

b. Jumbo is small for an elephant.

c. #This fly is small for an elephant.

The sentence 'Jumbo is small' seldom conveys that Jumbo's is small relative to everything in the world; it is often understood as conveying that he is small relative to a comparison class, such as the set of elephants. This phenomenon suggests that the domain of a gradable adjective can be implicitly restricted in a way parallel to the domain of a generalized quantifier or a definite description. As the second and the third examples show, the domain can also be explicitly restricted by a for-phrase. When the comparison class is explicitly restricted, predicating the adjective of an object that doesn't fall into the restricted domain results in infelicity.

A more striking piece of evidence supporting Kennedy's observation is that the domain of a gradable adjective can be bound by a generalized quantifier preceding it, in a way parallel to the domain of a generalized quantifier or a definite description. For example, the following sentence has a reading on which each family member is evaluated for his/ her relative tallness based on his/her own comparison class (e.g. people of his/ her age and gender):

(19) Every member of my family is tall.

Kennedy accounts for these instances of adjectival domain restriction by introducing a type-shifting rule that maps the measure function of a gradable adjective to a function that maps a contextually given set to a new measure function that is just like the adjective's original measure function except that its domain is now identical to the contextually given set. Let's call it 'the k-rule':

(20) *The k-rule* (Kennedy 2007, example #26):

For any gradable adjective with meaning $[\![A]\!]$, it can be typeshifted into the following meaning:

$$[\![A']\!] = \lambda f_{\langle e,t\rangle}[\lambda x[f(x).[\![A]\!](x)]]$$

The intuition behind this rule is that, while the basic meaning of a gradable adjective doesn't ask for a comparison class, its typeshifted meaning does, and the comparison class it asks for just is its new domain — the notation '$f(x).[\![A]\!](x)$' says that the meaning of a gradable adjective A maps x to a degree only when x falls into the comparison class f, and that $[\![A]\!](x)$ is undefined if x doesn't fall into the comparison class f. This rule accounts for the implicit and the explicit domain restrictions we saw in (18), because the value of the function argument f in the typeshifted meaning can either be given by the context when the domain restriction is implicit, or explicitly given by a for-phrase, which plausibly denotes a set.

To account for the bound reading of (19), there are two ways to proceed based on the k-rule. The first option is to combine it with Jacobson's variable-free account of binding, which is the approach we will eventually adopt. The second option, which we will focus on for the moment, is to generalize the k-rule into the following rule schema:

(21) *The generalized k-rule*:

For any gradable adjective with meaning $[\![A]\!]$, and for any natural number $n \geq 0$, it can be typeshifted into the following meaning:

$$[\![A']\!] = \lambda f_{\langle e...e,\langle e,t\rangle\rangle} \lambda y_1 ... \lambda y_n \lambda x[f(y_1)...(y_n)(x).[\![A]\!](x)]$$

Assumption: When $n = 0$, this rule delivers the k-rule.

To see the connection between this rule and the bounding reading of (19), we can consider the result of applying this rule to 'tall' when $n = 1$:

(22) $[\![tall]\!]^{k_1} = \lambda f_{\langle e,et\rangle}[\lambda y[\lambda x[f(y)(x).tall'(x)]]]$

This meaning should remind us of the meanings Stanley and Szabo assign to common nouns. According to them, the common noun 'bottle $f(y)$' has a function variable f, which asks for a function from individuals (which are of type e) to sets (which are of type $\langle e, t\rangle$, with t being the type of truth values), and an individual variable y that can be bound. Here the typeshifted meaning of 'tall' has a function argument (represented by 'λf' in the formula) that asks for a function (of type $\langle e, et\rangle$) from individuals to sets, and an individual argument (represented by 'λy' in the formula) that can be bound. These materials allow us to account for the bound readings of 'tall' in a way analogous to how Stanley and Szabo account for the bound readings of 'every bottle' or 'the bottle'. For example, suppose the function argument is saturated by a function F that maps a person to the people of her age, and the individual argument is bound by the generalized quantifier 'every member of my family'. Let's say the five-year-old Mary and the ten-year-old Susan fall into domain of 'every member of my family'. The result of applying F to Mary is the set of five-year-olds, and the result of applying F to Susan is the set of ten-year-olds. So we can see that the additional materials in the typeshifted 'tall' are precisely introduced to allow its comparison class to covary with the individuals ranged over by a generalized quantifier preceding it.

For the sake of completeness, here is how Kennedy accounts for the truth conditions of (19) — but we will not preoccupy ourselves with the details of the semantic composition, because our main goal is to adapt Kennedy's idea to respond to the first overgeneration worry raised by Bach. Since the domain of 'tall' varies with the family members, the contextually given function $\boldsymbol{s}$ (demanded by the phonologically null morpheme *pos*) maps the resulting domain-restricted tall-functions to potentially different contextual thresholds, so that each family member is evaluated for her relative tallness based on a potentially different contextual threshold. The following glosses show the truth conditions Kennedy assigns to (19):

(23) a. Let F be a function that maps a family member to the set of people of her age: Every member of my family x is such that s/he falls into the set of people of his/her age, $F(x)$, and his/her height is at least as high as the contextual threshold for the measure function of 'tall' whose domain is restricted to $F(x)$.

b. $\forall x(member'(x) \rightarrow F(x)(x).tall'(x) \geq \boldsymbol{s}(\lambda y[F(x)(y).tall'(y)]))$, where '$F(x)(x)$' reads '$x$ falls into the comparison class $F(x)$', and '$F(x)(y)$' reads 'y falls into the comparison class $F(x)$'.

For our purposes, the most important upshot is that Kennedy doesn't adopt Stanley's binding assumption and posit an individual variable in 'tall' in response to its bound readings. Instead, he introduces a typeshifting rule (i.e. the k-rule or its generalized version) which adds additional structures to 'tall' so that it becomes bindable. So the reader can easily anticipate how Kennedy would account for the truth conditions of 'Every Pentium processor is slow': Since there is no binding, we need not introduce a comparison class argument in 'slow' by typeshifting it. The sentence is true in a context if and only if the degree of slowness of every Pentium processor is at least as high as the contextual threshold $\boldsymbol{s}(slow')$. I will show how his idea can be easily adapted to respond to Bach's overgeneration worry in §2.4.

As I mentioned, we will eventually combine Kennedy's k-rule with Jacobson's variable-free account of binding. The reason for this is twofold. First, one might find our assumption that the generalized k-rule, which I repeat here, delivers Kennedy's k-rule when n is identical to 0 unsatisfactory.

(24) *The generalized k-rule*:

For any gradable adjective with meaning $[\![A]\!]$, and for any natural number $n \geq 0$, it can be typeshifted into the following meaning:

$$[\![A']\!] = \lambda f_{\langle e...e,\langle e,t\rangle\rangle} \lambda y_1 ... \lambda y_n \lambda x [f(y_1)...(y_n)(x).[\![A]\!](x)]$$

Assumption: When $n = 0$, this rule delivers the k-rule.

Since the individual arguments (represented by the y_i's in the formula) have already been stipulated to exist, it is somewhat odd that they disappear when n is identical 0. So one may justifiably object that the generalized k-rule should be split into the k-rule and a modified generalized k-rule where n is greater than or equal to 1. But this amendment will make the resulting semantic fragment less elegant, because the connection between the k-rule and the modified generalized k-rule is lost. One nice benefit of combining Kennedy's idea with Jacobson's variable-free account of binding is that there is an independently motivated typeshifting principle in her framework (i.e. the geach-rule), which, together with Kennedy's k-rule, can deliver every instance of the generalized k-rule. This is why I said that Kennedy's account of adjectival domain restriction and Jacobson's variable-free account of binding form a natural alliance.[19]

The second part of our reason is this. I will argue shortly that gradable adjectives can not only be bound, but be bound by two generalized quantifiers, so the parallel between domain restriction and adjectival domain restriction is all the more striking. What is most relevant to the debate between the pragmatic approach and the grammatical approach is that Jacobson's

[19] In discussing how we can rule his k-rule (or example #26) to deliver the bound readings of gradable adjectives, Kennedy (2007) refers to Jacobson's variable-free semantics, but he doesn't discuss the generalized version of the k-rule or how we can use the g-rule in variable-free semantics to deliver the bindable meaning of 'tall' that is needed to account for the bound reading of 'every member of my family is tall'.

variable-free account of binding has predicted that gradable adjectives, generalized quantifiers, and definite descriptions — among other expressions — can potentially be bound by more than one generalized quantifiers, and that they can all be accounted for by using the geach-rule and a small number of independently motivated rules in her framework. So by drawing from her approach to binding, not only can the proponents of the grammatical approach address the second overgeneration worry, they can also strengthen their case against the pragmatic approach, which I will argue isn't in a position to make those precise predictions.

2.3.2 Adjectival binding with two binders

I argue that the comparison class of 'slow' in the following example is naturally understood to be bound by both 'every team' and 'no current member':

(25) Because competitive running teams care about winning races a lot, they are always very strict about membership. So every team demands that no current member is **too slow**. 24-year-old Paul is definitely too slow to stay in the regional team after his injury, but he still isn't too slow to stay in the local team. On the other hand, 42-year-old Mary, who certainly runs slower than Paul, isn't too slow to stay in both the local team and the regional team, because she is more likely to win races in her own race group(s) than Paul in his race group(s).

Can we obtain the intended reading of 'too slow' by having the comparison class of 'slow' vary with either the teams or the members but not both? Let's first rule out the possibility that the comparison class only varies with the running teams. We can imagine that the runners in the comparison class of the regional team are on average faster than those in the comparison class of the local team. This is why it is possible that Paul isn't too slow to stay in the local team, despite his being too slow to stay in the regional team. However, since Paul is faster than Mary, if Paul stands out against the regional team's comparison class in terms of his slowness, so should Mary. But we know that Mary isn't too slow to stay in the regional team, because she may well be extremely competitive in her own age and gender group.

Let's also rule out the possibility that the comparison class only varies with the members. If each member has a unique comparison class, then it is inexplicable why Paul is too slow to stay in the regional team without being too slow to stay in the regional team. As we recall from our discussion of Kennedy's analysis of 'every member of my family is tall', the reason why every family member has a potentially different contextual threshold is that they have potentially different comparison classes. If Paul only has a single comparison class, he can only be evaluated for his relative slowness based on a single contextual threshold, which means that it is not possible that he is too slow for the regional team without being too slow for the local team.

We can account for the intended reading of 'too slow' only if we allow the comparison class to vary with both the teams and the members. Here is one possible way in which the relevant comparison class varies with both the teams and the members. Since the regional team must be more selective than the local team, we can imagine that, while the regional team demands that its members are among the top 5% of their age and gender groups, the local team only demands that its members are among the top 10% of their age and gender groups. Since the comparison class is a function of both the teams and the members, we can now understand 'too slow' in a way that fits our scenario. Paul is not too slow for the local team, because he doesn't stand out against the top 10% of his age and gender group in terms of his slowness. But he is too slow for the regional team because they set a higher bar of excellence for his age and gender group. His misfortune, of course, has nothing to do with whether Mary's slowness stands out against the top 5% of her age and gender group.

To finish our argument, we make explicit our assumption about the meaning of 'too'. 'Too', like Kennedy's *pos* morpheme, is a degree morpheme that maps the basic meaning of a gradable adjective to a set-denoting meaning. Here we assume that it maps the basic meaning of a gradable adjective and an individual to TRUE in a context just in case that individual's degree of slowness exceeds the contextual standard for the gradable adjective by an unacceptably large margin:

(26) $[\![too]\!] = \lambda g_{\langle e,d\rangle}[\lambda x[g(x) \gg \mathbf{s}(g)]]$, where '$g(x) \gg \mathbf{s}(g)$' reads x's degree as measured

by the measure function g exceeds the contextual threshold of g, i.e. $\boldsymbol{s}(g)$, by an unacceptably large margin.

Since the job of 'too' is merely to lift the contextual threshold already assigned to an adjectival scale (by the contextually given threshold function $\boldsymbol{s}$), and it is not bindable itself, the doubly-bound reading must be due to the fact that the adjective 'slow' is bound by both generalized quantifiers preceding it.[20]

2.3.3 The Jacobson-Kennedy approach to adjectival domain restriction

We now move on to showing how we can account for the doubly-bound reading of our example above by combining Kennedy's k-rule with Jacobson's variable-free approach to binding. For ease of exposition, let's now simplify our example (25) into:

(27) Every team demands that no member is too slow.

If we were to account for its doubly bound reading by using the generalized k-rule (or its modified version), we would let n be 2, and typeshift the meaning of 'slow' into the following:

(28) $[\![slow]\!]^{k_2} = \lambda f_{\langle ee,et\rangle}[\lambda y_1[\lambda y_2[\lambda x[f(y_1)(y_2)(x).tall'(x)]]]]$

This meaning should remind us of the meanings Breheny and Jacobson (§2.2.3) argue Stanley and Szabo would have to assign to common nouns in response to the doubly bound readings of generalized quantifiers and definite descriptions. Since the individual arguments (represented by 'λy_1' and 'λy_2' in the formula) can be bound by the generalized quantifiers 'every team' and 'no member', we can allow the comparison class of 'slow' to covary with both the teams and

[20]One may worry that, since the bound reading may be due to the bound reading of 'too' rather than that of 'slow', we haven't yet constructed an example where 'slow' is bound by two generalized quantifiers. To forestall this worry, let's consider the result of dropping the degree morpheme 'too' in our example. The intended bound reading of 'slow' is admittedly slightly harder to get without the morpheme, but it is still possible. Suppose I say that *every team specializing in 800m races takes every step to ensure that no new recruit is slow*. How should 'slow' be understood? Given our background knowledge about how 800m races are typically organized, we can imagine that when each team recruits new talents, instead of setting the same standard of slowness for each of their potential recruits, they calibrate the standard based on the potential recruit's age and gender. This age-and-gender-graded understanding of 'slowness' is the most relevant as far as winning races is concerned. So the standard of slowness can plausibly vary with each potential recruit. Since a runner competing in regional races is expected to be faster than a runner competing in local races, we can imagine that the relevant standards set for the potential recruits increase with the levels at which the teams compete. So the standard of slowness can vary not only with the potential recruits but also with the teams as well.

the members in the domains of the quantifiers.

But as I said in §2.3.1, instead of adopting the generalized k-rule, we will keep Kennedy's k-rule and combine it with Jacobson's variable-free approach of binding, because the fact that the comparison class argument slot f can take two (or more) arguments instead of only one is a direct consequence of that approach to binding, and this fact has been pointed out explicitly by Jacobson under the heading of *paycheck generalization*.[21] We will see shortly that we can derive all instances of the generalized k-rule from the k-rule and a typeshifting rule called 'the geach-rule', or 'the g-rule' in short.

Jacobson's variable-free approach is so-called because it makes no essential use of variables in the syntax and the semantics (which shouldn't be confused with the claim that the theorist can't use variables in the meta-language to describe the syntax and the semantics). It is motivated by the hypothesis, due to Richard Montague, that the surface structures of natural language sentences, rather than their hidden logical forms, are the inputs for semantic composition, and by a host of empirical motivations, most of which are beyond the scope of this discussion.[22] I take no stand on whether the variable-free approach or the variable-ful approach to syntax and semantics is the most viable — this issue goes way beyond the scope of our discussion, and my own expertise. As I said at the outset, our main interest in Jacobson's approach to binding is that it not only complements Kennedy's account of adjectival domain restriction, but it also provides the resources the grammatical approach needs to respond to the second overgeneration worry the pragmatists could have raised.

Our goal in this section is modest. We will introduce the variable-free approach to binding, and show how we can derive every instance of the generalized k-rule from the k-rule and the g-rule, and how we can account for the doubly-bound reading of 'slow' by using the instance (28) of the generalized k-rule when $n = 2$ and the binding rule in Jacobson's framework, namely the z-rule. Let's first introduce the g-rule by walking through the following examples:

(29) Porky escapes.

a. $[\![Porky]\!]$ = Porky

[21] I refer the reader to her (2014, Chapter 17.4) for her statement of the generalization.

[22] I refer the interested reader to Jacobson (1999, 2000, 2014).

b. $[\![escapes]\!] = \lambda x_e[escapes'(x)]$

c. $[\![Porky\, escapes]\!] = [\![escapes]\!]([\![Porky]\!]) = escapes'(Porky)$

[(a)&(b), Function Application]

(30) It escapes.

a. $[\![It]\!] = \lambda x_e[x]$

b. $[\![escapes]\!] = \lambda x_e[escapes'(x)]$

The semantic composition for 'Porky escapes' is straightforward, but the semantic composition for 'It escapes' is less so: Since the meaning of the pronoun 'it' in variable-free semantics is an identity function from individuals (of type *e*) to individuals, we can't use function application to combine it with the meaning of 'escapes', which is a function from individuals to truth values (which are of type *t*). One may want to solve this apparent problem by introducing a function composition rule into the semantics, so that we can function-compose the meaning of 'it' with the meaning of 'escapes', and obtain the function that maps an individual to TRUE if it escapes (which we should notice just is the meaning of 'escapes'). This is certainly a workable strategy, but the variable-free approach substitutes the g-rule for the function composition rule, because the g-rule can be used to account for a host of phenomena, some of which we are about to see. Let me state the g-rule as follows (We will ignore the syntax in following, but the reader can observe in the sample derivation in the appendix that the syntax parallels the semantics in an elegant way):

(31) *The g-rule* (Jacobson 1999):

If f is a function of type $\langle a, b\rangle$, then $\boldsymbol{g}(f)$ is a function of type $\langle\langle c, a\rangle, \langle c, b\rangle\rangle$, where $\boldsymbol{g}(f) = \lambda V_{\langle c,a\rangle}[\lambda C_c[f(V(C))]]$

Let's call the result of applying the g-rule to the meaning of 'escapes' *g-escapes*. It takes as inputs a function f from individuals to individuals and an individual x, and returns TRUE if the image of that individual under the function (i.e. $f(x)$) escapes. For example, let's call the function that maps an individual to his/her best friend *the best friend function*. The result of applying *g-escapes* to the best friend function is the function that maps an individual to TRUE if his/ her best friend escapes. So we can easily see that the result of applying *g-escapes* to the identity function is the function that maps an individual to TRUE if the individual itself (which just is its image under the identity function) escapes. This function just is the meaning of 'escapes', which we should recall is the result of function-composing the meaning of 'it' with the meaning of 'escapes'. This means that we just used the g-rule to replicate the result of function-composing the meaning of 'it' with the meaning of 'escapes'. The following shows the semantic composition of 'It escaped' (We will ignore tense and intensions throughout, since they play no essential roles in our examples):

(32) It escaped.

a. $[\![It]\!] = \lambda x_e[x]$

b. $[\![escaped]\!] = \lambda x_e[escaped'(x)]$

c. $[\![escaped]\!]^g = \lambda h_{\langle e,e\rangle}[\lambda y[escaped'(h(y))]]$

[(b), the g-rule]

d. $[\![It\, escaped]\!] = [\![escaped]\!]([\![It]\!]) = \lambda y_e[escaped'(y)] = [\![escaped]\!]$

[(a)&(c), Function Application]

We now introduce the binding rule in variable-free semantics (i.e. the z-rule) by walking through the following example (The reader who is primarily interested in how we can derive every instance of the generalized k-rule from the k-rule and the g-rule can skip our discussion of the z-rule):

(33) Every man loves his mother.

a. $[\![every]\!] = \lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[\forall x(Px \rightarrow Qx)]]$

b. $[\![man]\!] = \lambda x_e[man'(x)]$

c. $[\![every\, man]\!] = [\![every]\!]([\![man]\!]) = \lambda Q_{\langle e,t\rangle}[\forall x(man'(x) \rightarrow Qx)]$

[(a)&(b), Function Application]

d. $[\![loves]\!] = \lambda y[\lambda x[loves'(x,y)]]$

e. $[\![his\ mother]\!] = \lambda x_e[x\text{'s mother}]$

We will set aside the domain restriction of the generalized quantifier 'every man', a topic we will return to in §2.4. The bound reading of our example can be glossed as: 'Every man x loves x's mother', but we are about to see that the variable-free approach can account for this reading without making use of variables in the syntax and the semantics. The meanings of 'every man', 'loves', and 'his mother' given above are uncontroversial: 'Every man' denotes a generalized quantifier that maps a set Q to TRUE if every man falls into that set; 'loves' maps individuals y and x to TRUE if x loves y; ignoring gender, 'his mother' denotes a function that maps an individual x to x's mother, which we will call *the mother function*. The variable-free approach achieves binding by typeshifting the meaning of 'love', so that it takes both the mother function and an individual (i.e. the lover) as inputs, and map those inputs to TRUE if the lover loves its image under the mother function (i.e. his mother). Let's call the typeshifted meaning of 'love' *z-love*. That someone z-loves the mother function just in case s/he loves his/her mother. So the bound reading of 'every man loves his mother' can be glossed as 'every man z-loves the mother function'.

The rule that typeshifts the meaning of 'love' into *z-loves* is the following:

(34) *The z-rule* (Jacobson 1999):

Given a function f of type $\langle a, \langle e, b\rangle\rangle$, $\boldsymbol{z}(f)$ is a function of type $\langle\langle e, a\rangle, \langle e, b\rangle\rangle$, where $\boldsymbol{z}(f) = \lambda G_{\langle e,a\rangle}[\lambda x[f(G(x))(x)]]$

Let's unpack this rule by instantiating f to the meaning of 'love', which is a two-place relation and is hence of type $\langle e, \langle e, t\rangle\rangle$. So a and b in the rule should be instantiated to e and t respectively. The typeshifted meaning of 'love', i.e. $z(love')$, is ready to take as inputs a function G of type $\langle e, e\rangle$, which is precisely the semantic type of the mother function, and an individual x, and map them to TRUE just in case x loves its image under G.

For the sake of completeness, the following is the semantic composition of 'every man loves his mother':

(35) Every man loves his mother

a. $[\![every\ man]\!] = [\![every]\!]([\![man]\!]) = \lambda Q_{\langle e,t\rangle}[\forall x(man'(x) \rightarrow Qx)]$

[see above]

b. $[\![loves]\!] = \lambda y[\lambda x[loves'(x,y)]]$

c. $[\![loves]\!]^z = \lambda G_{\langle e,e\rangle}[\lambda z[loves'(G(z))(z)]]$

[(b), the z-rule]

d. $[\![his\ mother]\!] = \lambda x_e[x\text{'s mother}]$

e. $[\![loves\ his\ mother]\!] = [\![loves]\!]^z([\![his\ mother]\!]) = \lambda z[loves'(z, z\text{'s mother})]$

[(c)&(d), Function Application]

f. $[\![Every\ man\ love\ his\ mother]\!] = [\![every\ man]\!]([\![loves\ his\ mother]\!])$

$= \forall x(man'(x) \rightarrow loves'(x, x\text{'s mother}))$

[(a)&(e), Function Application]

We should be careful to note that the z-rule need not apply, because 'his mother' in our example can refer to the mother of a contextually salient individual instead of being bound. On Jacobson's view, when 'his mother' isn't bound, our example expresses a propositional function that maps a contextually salient individual to TRUE if every man loves that individual's mother, and she would obtain that proposition function by passing the individual argument of 'his mother' to the sentence level by using the g-rule. The following demonstrates how it can be done (The crucial step is the transition from (d) to (e)):[23]

(36) Every man loves his mother (the unbound/ referential reading)

a. $[\![every\ man]\!] = \lambda Q_{\langle e,t\rangle}[\forall x(man'(x) \rightarrow Qx)]$

b. $[\![loves]\!] = \lambda y[\lambda x[loves'(x,y)]]$

c. $[\![his\ mother]\!] = \lambda x_e[x\text{'s mother}]$

[23] Here I simplify the derivation by applying the g-rule to 'every man loves'. But since I believe that the g-rule is intended to be applied to lexical items, the g-rule has to be applied to each word in that expression individually, which would make the derivation several steps longer.

d. $[\![Every\ man\ loves]\!] = [\![every\ man]\!]^g([\![loves]\!]) = \lambda z[\forall x(man'(x) \rightarrow loves'(z)(x))]$

[We use the g-rule to replicate the resulting of function composition]

e. $[\![Every\ man\ loves]\!]^g = \lambda G_{\langle e,e\rangle}[\lambda u[[\![Every\ man\ loves]\!](G(u))]]$

$= \lambda G_{\langle e,e\rangle}[\lambda u[\forall x(man'(x) \rightarrow loves'(G(u))(x))]]$

[(d), the g-rule]

f. $[\![Every\ man\ loves\ his\ mother]\!] = [\![Every\ man\ loves]\!]^g([\![his\ mother]\!])$

$= \lambda u[\forall x(man'(x) \rightarrow loves'(x, u\text{'s mother}))]$

As we can see, to pass the context-dependence of 'his mother' to the sentence level is a laborious process which demands repeated applications of the g-rule. So, for the sake of convenience, we will adopt what I call the *local saturation convention*, which allows the context-dependence of any expression to be discharged locally. For example, this convention allows us to saturate the individual argument of 'his mother' with a contextually salient individual locally. The following shows how we would redo the semantic composition above if we help ourselves to the convention (The crucial step is the transition from (c) to (d)):

(37) Every man loves his mother (Context: Paul is contextually salient.)

a. $[\![every\ man]\!] = \lambda Q_{\langle e,t\rangle}[\forall x(man'(x) \rightarrow Qx)]$

b. $[\![loves]\!] = \lambda y[\lambda x[loves'(x, y)]]$

c. $[\![his\ mother]\!] = \lambda x_e[x\text{'s mother}]$

d. $[\![his\ mother]\!]^s = [\![his\ mother]\!](Paul) = \text{Paul's mother}$

[(c), local saturation convention]

e. $[\![loves\ his\ mother]\!] = [\![loves]\!]([\![his\ mother]\!]^s) = \lambda x[loves'(x, \text{Paul's mother})]$

[(b)&(d), Function Application]

f. $[\![every\ man\ loves\ his\ mother]\!] = [\![every\ man]\!]([\![loves\ his\ mother]\!])$

$= \forall x(man'(x) \rightarrow loves'(x, \text{Paul's mother}))$

[(a)&(e), Function Application]

Let me emphasize that this convention is only introduced to simplify our semantic compositions. I take no stand on whether a sentence can express a propositional function as Jacobson suggests, or whether semantic processing requires context-dependence to be resolved locally. We will adopt the convention throughout our discussion.

We are now ready to see one of the empirical motivations of the g-rule: That rule, together with the z-rule, accounts for the paycheck readings of pronouns. The most important point we should focus on is that the principles that account for the paycheck readings of pronouns can be used to derive all possible instances of the generalized k-rule, and to account for the singly and the doubly bound readings of gradable adjectives. Consider the following examples of paycheck pronouns (Karttunen 1969):[24]

(38) Mary deposited her paycheck in the bank, but every man lost **it** (*his paycheck*).

(39) The new ruling requires that the semester report which professors write about their students be nonconfidential. However, each professor we interviewed told us that not every student would want to see **it** (*his report about him*) (Cooper 1979, p.79).

The intended reading of (38) says that every man lost his own paycheck, so the pronoun 'it' doesn't refer to a contextually salient paycheck. Nor is it bound in a way similar to 'his' in 'Every man loves his mother'. To obtain the intended reading of (38), it seems that we need to apply to the meaning of the sentence the function that maps an individual to his paycheck — which we will call *the paycheck function* — and have every man z-lost that function (we should recall that to z-love the mother function is to love your mum). Similarly, to obtain the intended reading of (39), we need to apply to its meaning a two-place report function that maps a professor and a student to that professor's report about that student, and have 'each professor' and 'every student' bind the first and the second arguments of that function respectively.

It is not an accident that the interpretations of some sentences are incomplete without us being able to apply to their meanings some contextually salient functions. Engdahl (1986)

[24] Karttunen's example is the following:

(i) The man who gave his paycheck to his wife was wiser than the man who gave it to his mistress.

observed a parallel between examples like those above and the following questions, which seem to ask for a function that maps an individual — or, more generally, a n-tuple of individual(s) — to an individual:

(40) Q: Who does every man love the most?

A: His mother.

(41) Q: What does *every father* tell **his children** never to forget?

A: How *he* has sacrificed himself for **them** (Engdahl 1986, p.177).

The variable-free approach can account for the intended readings of these functional questions in a way similar to how it accounts for paycheck pronouns, but we will only focus on its account of the latter, and discuss its connection with the singly and doubly bound readings of gradable adjectives.

Let's consider the second conjunct of (38):

(42) Every man lost it.

a. $[\![every\ man]\!] = \lambda Q_{\langle e,t \rangle}[\forall x(man'(x) \rightarrow Qx)]$

b. $[\![lost]\!] = \lambda y[\lambda x[lost'(x, y)]]$

c. $[\![it]\!] = \lambda x[x]$

As I said, its intended reading can be paraphrased as 'every man z-lost the paycheck function'. But as we can surmise from the meanings of 'every man', 'lost', and 'it' above, we can only obtain the referential reading of the pronoun (on which it refers to a contextually salient object) by combining them. The variable-free account solves this puzzle by applying the g-rule to the meaning of 'it', so that it becomes an function with an argument of type $\langle e, e \rangle$ (which is the semantic type of the paycheck function). While Jacobson would pass that argument to the sentence level, so that the meaning of the sentence asks for the paycheck function, we are going to saturate the argument locally using the local saturation convention:

(43) It (Context: the paycheck function is contextually salient)

a. $[\![it]\!] = \lambda x[x]$

b. $[\![it]\!]^g = \lambda G[\lambda z[[\![it]\!](G(z))]] = \lambda G[\lambda z[G(z)]]$

[(a), the g-rule]

c. $[\![it]\!]^{gs} = [\![it]\!]^g(\text{the paycheck function}) = \lambda z[z\text{'s paycheck}]$

[(b), the local saturation convention]

Once we have this key step, we can easily complete the rest of the semantic composition: We first use the z-rule to typeshift the meaning of 'lost' into *z-lost*. We then apply *z-lost* to the paycheck function, and apply the meaning of 'every man' to the meaning of 'z-lost the paycheck function'. For the sake of completeness, the rest of the semantic composition is as follows:

(44) Every man lost it.

a. $[\![lost]\!] = \lambda y[\lambda x[lost'(x, y)]]$

b. $[\![lost]\!]^z = \lambda G_{\langle e,e\rangle}[\lambda z[lost'(z, G(z))]]$

[(a), the z-rule]

c. $[\![lost\ it]\!] = [\![lost]\!]^z([\![it]\!]^{gs}) = \lambda z[lost'(z, z\text{'s paycheck})]$

[(b)&(43-c), Function Application]

d. $[\![Every\ man\ lost\ it]\!] = [\![Every\ man]\!]([\![lost\ it]\!])$

$= \forall x(man'(x) \rightarrow lost'(x, x\text{'s paycheck}))$

[(42-a)&(c), Function Application]

Let's now move onto the key step of the variable-free analysis of Cooper's example (39):

(45) [E]ach professor we interviewed told us that not every student would want to see **it** (*his report about him*)

This sentence seems to ask for a two-place report function that maps a pair of individuals to the report the first individual compiles for the second individual. And we can surmise from our discussion of the last example that we can't have the sentence ask for that function without manipulating the meaning of 'it'. A striking result achieved by the variable-free approach is that, if we can account for the intended reading of our last example, we can account for this example as well, with no additional machinery. To create an argument place that asks for a two-place report function, we apply the g-rule to the meaning of 'it' twice:

(46) It (Context: The two-place function that maps a pair of individuals to the report compiled by the first individual for the second individual is contextually salient)

a. $[\![it]\!]^g = \lambda G[\lambda z[[\![it]\!](G(z))]] = \lambda G_{\langle e,e\rangle}[\lambda z[G(z)]]$

[see above]

b. $[\![it]\!]^{gg} = \lambda H[\lambda y[[\![it]\!]^g(G(y))]] = \lambda H_{\langle e,ee\rangle}[\lambda y[\lambda z[H(y)(z)]]]$

[(a), the g-rule]

The next step is that we use the local saturation convention to saturate the newly created argument of type $\langle e, ee\rangle$ locally:

(47) $[\![it]\!]^{ggs} = [\![it]\!]^{gg}$(the 2-place report function)$= \lambda y[\lambda z[y$'s report for $z]]$

[(46-b), the local saturation convention]

To complete the semantic composition for (45), we will apply the z-rule to 'told' and to 'see', and have 'each professor' and 'not every student' bind the first and the second individual arguments of $[\![it]\!]^{ggs}$ respectively. The rather lengthy derivation is omitted here, but the reader can refer to the appendix for the derivation of a related example.

We are now ready to appreciate the connection between the variable-free approach of paycheck pronouns and the singly and the doubly bound readings of gradable adjectives. Just as we can obtain the two-place paycheck reading of 'it' in Cooper's example by applying the g-rule to the meaning of 'it' twice, we can obtain the instance (28) of the generalized k-rule when n is identical to 2 by applying the g-rule to the k-typeshifted meaning of 'slow' twice:

(48) a. $[\![slow]\!] = \lambda x[slow'(x)] = \lambda x[x$'s degree of slowness$]$

b. $[\![slow]\!]^k = \lambda f_{\langle e,t\rangle}[\lambda x[f(x).slow'(x)]]$

[(a), the k-rule]

c. $[\![slow]\!]^{kg} = \lambda h_{\langle e,et\rangle}[\lambda y[[\![slow]\!]^{k}(h(y))]] = \lambda h_{\langle e,et\rangle}[\lambda y[\lambda x[h(y)(x).slow'(x)]]]$

[(b), the g-rule]

d. $[\![slow]\!]^{kgg} = \lambda j_{\langle ee,et\rangle}[\lambda z[[\![slow]\!]^{kg}(j(z))]] = \lambda j_{\langle ee,et\rangle}[\lambda z[\lambda y[\lambda x[j(z)(y)(x).slow'(x)]]]$

[(c), the g-rule]

Notice that (48-c) is the instance of the generalized k-rule when n is identical to 1. As we discussed above, this is the meaning we need to account for the bound reading of 'every member of my family is tall'. (48-d) just is the instance (28) of the generalized k-rule when n is identical to 2. This is the meaning we need to account for the doubly bound reading of 'slow' in (27); the reader can refer to the appendix for a sample derivation, which I omit here. Here are the two most crucial steps in the derivation: First, we use the local saturation convention to saturate the argument of type $\langle ee, et\rangle$ in (48-d) (which is represented by 'λj' is the formula) with the two-place comparison class function that maps a team and a member to the comparison class whose membership depends on the standard the team sets for the member's age and gender group. Second, we apply the z-rule to 'demands' and to 'is', and have 'every team' and 'no member' bind respectively the first and the second arguments of the comparison class function.

We can easily verify that repeated applications of the g-rule on the k-typeshifted meaning of 'slow' delivers every instance of the generalized k-rule. So the result of combining Kennedy's k-rule with Jacobson's variable-free approach to binding is just as powerful as adopting the k-rule and the modified generalized k-rule, and it preserves the connections between the instances of the generalized k-rule, because all of them can be derived from the k-rule and the g-rule. We will call the resulting account of adjectival domain restriction the 'Jacobson-Kennedy approach'.[25] It will serve as as model for responding to the two overgeneration worries we discussed in §2.

[25] Here the authors are arranged in alphabetical order. The observation that gradable adjectives, which have a comparison class argument (which is either part of their lexical meanings or introduced by a typeshifting principle such as Kennedy's k-rule), has been made by Jacobson (2014). Here I apply that observation to Kennedy's account of gradable adjectives.

2.4 Response to Overgeneration Worries

2.4.1 The proposal: an analogue of the Jacobson-Kennedy approach

I now propose an analogue of the Jacobson-Kennedy approach which the grammatist can use to respond to the overgeneration worries. My own contribution is modest: I propose that the following typeshifting rule, which is analogous to Kennedy's k-rule, can be applied to common nouns:

(49) *The c-rule* ('c' for contextual domain and common noun):

For any common noun with meaning $[\![N]\!]$, it can be typeshifted into the following meaning:

$$[\![N']\!] = \lambda C_{\langle e,t \rangle}[\lambda x[[\![N]\!](x) \wedge C(x)]]$$

Here is the idea behind this rule. Just as gradable adjectives on Kennedy's account only ask for comparison classes when they are k-typeshifted, common nouns on my proposal only ask for contextual domains when they are c-typeshifted. When the c-typeshifted meaning of 'bottle' is saturated by a contextually salient set (e.g. the objects in a certain room), an object falls into the extension of the typeshifted 'bottle' just in case it is inside the intersection of the set of bottles and the contextually salient set.

As the reader can anticipate, we will combine the c-rule with Jacobson's variable-free framework, because its g-rule predicts that, if generalized quantifiers and definite descriptions have an argument asking for a contextual domain, that argument can become a function argument that asks for a function that maps n individuals to a contextual domain, in a way analogous to the comparison class argument of the k-typeshifted gradable adjectives. The following is the result of applying the g-rule to the c-typeshifted meaning of 'bottle' twice:

(50) a. $[\![bottle]\!] = \lambda x[bottle'(x)]$

[The lexical meaning of 'bottle']

b. $[\![bottle]\!]^c = \lambda C_{\langle e,t\rangle}[\lambda x[bottle'(x) \wedge C(x)]]$

[(a), the c-rule]

c. $[\![bottle]\!]^{cg} = \lambda f_{\langle e,et\rangle}[\lambda y[\lambda x[bottle'(x) \wedge f(y)(x)]]]$

[(b), the g-rule]

d. $[\![bottle]\!]^{cgg} = \lambda f'_{\langle ee,et\rangle}[\lambda z[\lambda y[\lambda x[bottle'(x) \wedge f'(z)(y)(x)]]]]$

[(c), the g-rule]

(50-c) is the meaning we need to account for the bound reading of 'In every room in John's house, every bottle is in the corner', and (50-d) is the meaning we need to account for the doubly bound readings of 'every bottle' and 'the bottle' (if they exist). It is easy to see that our approach is not vulnerable to the overgeneration worries we discussed in §2.2. Since we do not respond to the bindability of generalized quantifiers and definite descriptions by positing variables in them, we have a straightforward response to the first overgeneration worry, which we should recall is best brought out by sentences like the followings:

(51) a. Every prime number is divisible by 1 and itself.

b. The prime numbers are numbers which are divisible by 1 and themselves.

Since there is no binding, and no domain restriction is called for, the common noun 'prime number' need not be typeshifted, which means that there are no unexploited variables or arguments in it.

Our response to the second overgeneration worry is equally straightforward. While we need to create a two-place function argument and two individual arguments in the common nouns 'mistake' and 'drink' by applying the g-rule to them twice in order to account for the doubly bound readings of 'every mistake' and 'the drink', we only need to create an one-place function argument and a single individual argument in 'bottle' by applying the g-rule to it once in order to account for the singly bound readings of 'every bottle' and 'the bottle'. To account for ordinary domain restriction, such as the fact that 'every bottle is empty' seldom conveys that every bottle in the world is empty, we only need the c-typeshifted meaning of 'bottle', which asks for a contextual domain. When there is no binding and no domain restriction, we can stick to the lexical meanings of common nouns, which are kept as simple as possible by our approach.

I should note that we need not implement our general approach by introducing a typeshifting rule targeting common nouns. We could implement it by introducing a typeshifting rule that creates a contextual domain argument in the determiners:

(52) *The d-rule* ('d' for determiners):

For any determiner with meaning $[\![D]\!]$, it can be typeshifted into the following meaning:

$$[\![D']\!] = \lambda C_{\langle e,t\rangle}[\lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[[\![D]\!](P \cap C)(Q)]]]$$

Let me explain the idea behind this rule by walking through the example 'every bottle is empty'. 'Every', as we saw in (33-a), denotes a function, which asks for two sets P and Q, and returns TRUE if P is a subset of Q. The d-typeshifted meaning also asks for two sets P and Q, but it now has an additional argument that asks for a contextual domain C. It maps the contextual domain C and the two sets P and Q to the truth value that 'every' maps the intersection of P and C, i.e. $P \cap C$, and Q.

The same rule can be used to add a contextual domain argument to 'the', which I assume has the following meaning:

(53) $[\![the]\!] = \lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[\exists x(Px \wedge \forall y(Py \rightarrow x = y)).\exists x(Px \wedge Qx))]]$

What it says is essentially that the meaning of 'the' maps two sets P and Q to a truth value only if P is a singleton set, and that it maps them to TRUE just in cast the intersection of P and Q is non-empty.

For the sake of completeness, here is the result of applying the d-rule to 'the':

(54) $[\![the]\!]^d$

$= \lambda C_{\langle e,t\rangle}[\lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[[\![the]\!](P \cap C)(Q)]]]$

$= \lambda C_{\langle e,t\rangle}[\lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[\exists x(x \in P \cap C \wedge \forall y(y \in P \cap C \rightarrow y = x).\exists x(Px \wedge Cx \wedge Qx)))]]]$

So the d-typeshifted meaning of 'the' essentially adds to the truth conditions of 'the P is Q' the additional requirement that the object that has the property denoted by 'P' falls into the contextual domain, which is the result we want.

To account for the doubly bound readings of generalized quantifiers and definite descriptions, we can apply the g-rule to the d-typeshifted meanings of 'every' and 'the' twice. The following shows the result of applying the g-rule to the d-typeshifted meaning of 'every' twice, which is the meaning we need to account for the doubly bound reading of 'every mistake' in Breheny's example:

(55) a. $[\![every]\!] = \lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[\forall x(Px \rightarrow Qx)]]$

b. $[\![every]\!]^d = \lambda C_{\langle e,t\rangle}[\lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[\forall x(Px \wedge Cx \rightarrow Qx)]]]$

[(a), the d-rule]

c. $[\![every]\!]^{dg}$

$= \lambda f_{\langle e,et\rangle}[\lambda y_e[\lambda P_{\langle e,t\rangle}[\lambda Q_{\langle e,t\rangle}[\forall x(Px \wedge f(y)(x) \rightarrow Qx)]]]]$

[(b), the g-rule]

d. $[\![every]\!]^{dgg}$

$= \lambda f'_{\langle ee,et\rangle}[\lambda z_e[\lambda y[\lambda P[\lambda Q[\forall x(Px \wedge f'(z)(y)(x) \rightarrow Qx)]]]]]$

[(c), the g-rule]

The steps here are exactly parallel to those in (50), where we apply the g-rule to the c-typeshifted meaning of 'bottle' twice. We can derive the meaning of 'the' we need to account for the intended reading of 'the drink' in Jacobson's example by applying the g-rule to the d-typeshifted meaning of 'the' twice. The steps, which are exactly analogous to those above, are omitted here.

We have now completed our response to the overgeneration worries. We now discuss its significance to the debate between the grammatical approach and the pragmatic approach to domain restriction.

2.4.2 Its significance

Let's recall what's at stake in the debate. Since the pragmatist holds that the qualifications on the meanings of generalized quantifiers and definite descriptions are filled in by us, rather than being parts of their semantic contents, she holds that sentences containing generalized quantifiers and definite descriptions (e.g. 'every bottle is empty'; 'the corner is nice') are typically either false or without truth value. The strongest motivation for this view is that, by doing away with variables in generalized quantifiers and definite descriptions which are often unexploited, it keeps the syntax and the semantics maximally simple.

Here is the first significance of our new response to the overgeneration worries. It shows that the grammatical approach has the resources to avoid unexploited variables. So it weakens the major motivation for the pragmatic approach by showing that it isn't the only way to keep the syntax and the semantics maximally simple.

Our response also allows the grammatist to reorient her debate with the pragmatist in a way that is favorable to her — and constructive for both sides — by focusing on the predictive and the explanatory powers of her own approach and the pragmatist's. To see why she may want to reorient the debate, consider the following exchange between Szabo and Stanley and Bach:

> The obvious disadvantage [of the pragmatic approach] is that one has to abandon ordinary intuitions concerning the truth [or] falsity of most sentences containing quantifiers. This is worrisome because accounting for our ordinary judgements about the truth-conditions of various sentences is the central aim of semantics. Since these judgements are the data of semantic theorizing, we should be careful with proposals that suggest a radical revision of these judgements (Stanley and Szabo 2000a, p.240).
>
> Now I should have thought that the central aim of semantics is to account for semantic facts. 'Ordinary judgements' or 'intuitions' provide data for semantics, but it is an open question to what extent they reveal semantic facts, hence should be explained rather than explained away. Since they are often responsive to non-semantic information, they should not be given too much weight. Besides, they don't seem to play a role in ordinary communication. People don't have to be able to make accurate judgements about semantic facts to be sensitive to semantic information. In the course of speaking and listening to one another, people do not consciously reflect on the propositions semantically expressed by the sentences they hear, but are focused on what they are communicating and on what is being communicated to them (Bach 2000, pp.267-268).

Here we have what appears to be a stalemate over the evidential role of ordinary intuitions about sentences' truth conditions in semantic theorizing.[26] While Stanley and Szabo stress the fact their account respects ordinary intuitions about truth conditions as a point in favor of their approach, Bach seems to downplay the importance of such intuitions by suggesting that they may not be reliable guide to the propositions expressed by sentences, and contends that the central aim of semantics is to account for what he calls 'semantics facts', rather than to account for 'our ordinary judgements about the truth-conditions of various sentences'.

While I don't understand what Bach means by 'semantic facts', I agree with him that ordinary intuitions are not always reliable guide to the propositions semantically expressed by sentences.[27] For example, speakers often judge sentences like 'this table is flat' and 'Paul is 1.6m tall' to be true, even though nothing in the physical realm is perfectly flat and (it is very likely that) no one is exactly 1.6m tall. These ordinary intuitions can't decide for the theorist whether she should assign true propositions to those sentences (relative to their contexts of utterance), because that decision depends on whether the imprecise uses of maximal standard absolute adjectives and number terms contribute to their semantic contents proper or merely to what they pragmatically convey, but the theoretical status of imprecision is far from being settled.[28] Since one can reasonably doubt, as Bach does, that ordinary intuitions are reliable guide to the semantic contents of sentences containing generalized quantifiers and definite descriptions, the grammatist's respect for ordinary intuitions about truth conditions doesn't seem to be the most compelling reason in favor of her position.[29]

[26] This issue is a recurring theme in debates about the semantics-pragmatics distinction. See Cappelen and Lepore (2005) and the literature it inspires.

[27] Here are some examples that I am not going to discuss in length. The first example is scalar implicature. If Paul has three children, the sentence 'Paul has two children' should sound false to some ordinary speakers. But we can't infer from those speakers' judgments that the semantic content of the sentence is false, because if the semantic content of the sentence is that Paul has at least two children, the content that is judged false is the result of combining the semantic content of the sentence with its implicature that Paul doesn't have more than two children.

Another example is metalinguistic negotiation, which we can assume is a process where speakers negotiate the speaker meaning of a certain expression, rather than its semantic content. To use an example similar to Ludlow (2014)'s: If 'Secretariat' refers to a horse, the sentence 'Secretariat is an athlete' may sound false to some ordinary speakers without context or in a context where NFL players are salient. However, it is certainly possible that, after long discussions about the qualities of Secretariat that qualify her as an athlete, some ordinary speakers may come to accept that the sentence is true. So, once again, ordinary speaker's judgment about truth conditions or truth values need not track semantic contents.

[28] For discussions on imprecision, see Pinkal (1995), Lasersohn (1999), Kennedy and McNally (2005), Kennedy (2007), Sauerland and Stateva (2011). I argued in the first chapter that the imprecise uses of maximal standard absolute adjectives contribute to the adjectives' semantic contents.

[29] Here I am restricting to (i) intuitions about *truth conditions* or *truth values*, and to (ii) the claim that those intuitions are reliable guide to semantic contents. Not all intuitions that can be data for linguistics involve intuitions about sentences' truth conditions or truth values. Intuitions about entailment, covariation, and (in)felicity aren't intuitions about sentences' truth conditions or truth values: In judging whether a given entailment holds, ordinary speakers

I propose that the grammatist reorients the debate by focusing on the predictive and the explanatory powers of her own approach and the pragmatic approach, because the predictive and the explanatory powers it has, rather than its respect for ordinary intuitions about truth conditions, are the most compelling considerations in favor of her position, and it is difficult to see how the pragmatist can object to this reorientation, given that both approaches need to earn their keep by making non-trivial predictions, and by explaining them with independently motivated principles.

Our new response to the overgeneration worries, which draws from data that go beyond those on domain restriction, and which is based on an explicitly given semantic fragment, demonstrates the predictive and the explanatory power of the grammatical approach. Here I will make the point by using some results of Jacobson's approach to binding, but it can probably be made based on other frameworks and results. Her approach has made the striking prediction that gradable adjectives — rather than any expression we find in natural language — can potentially be bound by two generalized quantifiers in a way parallel to generalized quantifiers and definite descriptions. It is in a position to make that prediction based on the assumption that gradable adjectives have a comparison class argument (which is either part of their lexical meanings or introduced by a typeshifting principle such as Kennedy's k-rule), because repeated applications of the g-rule on a meaning having a comparison class argument result in meanings that have a functional comparison class argument (which asks for n individuals and returns a comparison class). That prediction is borne out by my example in §2.3.2, and, as we saw in §2.3.3, we can account for its doubly bound reading by using two independently motivated typeshifting principles, namely the g-rule and the z-rule, which play a key role in explaining the paycheck readings of pronouns, the singly and the doubly bound readings of generalized quantifiers and definite descriptions, among other phenomena.

Unless the pragmatist makes explicit the syntactic and the semantic fragments that are

need not know the actual truth values of the sentences, because she is only required to judge whether she ought to accept the conclusion on accepting the premise(s). In order to determine whether 'every man loves his mother' has a covariation reading, ordinary speakers need not know its truth value relative to a context. And, certainly, we can determine whether a given sentence is felicitous in a context without knowing its truth value in that context.

My assumption about these intuitions is that, if they yield a systematic and repeatable pattern (e.g. 'completely tall' is consistently judged by ordinary speakers to be unacceptable), they are data that call for an explanation. But I don't assume that the explanation must be either grammatical or pragmatic (and not both).

supposed to go with her pragmatic account, or/and substantially revise her account of how we 'read things into utterances', it is difficult to see how she can make such precise predictions.[30] As it stands, the pragmatic account can't predict why verbs (e.g. 'exist', 'is') don't display covariation behavior analogous to those of generalized quantifiers, definite descriptions, and gradable adjectives. Nor can it predict that the doubly bound readings of generalized quantifiers, definite descriptions, and gradable adjectives aren't surprising given that they can be singly bound. Its failure to make these predictions seems to put it at a serious disadvantage. So even if we suspend judgment on whether ordinary intuitions are reliable guide to the semantic contents of sentences containing generalized quantifiers and definite descriptions, there seem to be sufficient reasons to disfavor the pragmatic approach in its current form.[31]

2.5 Conclusion

I have argued that there are two well-founded overgeneration worries with Stanley and Szabo's account of domain restriction, and that they can respond to them by introducing a typeshifting rule analogous to Kennedy's k-rule (i.e. the c-rule or the d-rule), and by implementing binding in a variable-free framework. The interest of our discussion goes beyond the debate between

[30] It is also difficult to evaluate her claim that her approach keeps the syntax and the semantics maximally simple unless she makes her syntactic and semantic fragments explicit.

[31] Here is an additional worry with the pragmatic account. Its account of binding appears to lack generality. As we discussed in §1, it accounts for the bound reading of Stanley and Szabo's example (1-a), which I repeat here as (i), by suggesting that we read the relevant qualifications on the semantic contents of 'every bottle' and 'the corner':

(i) In every room in John's house, every bottle (in that room) is in the corner (of that room).

However, this approach to binding is unable to deliver the paycheck reading of pronouns, which not only involves binding, but involves the pronouns asking for a contextually salient function. In order to account for them by having us read additional contents into the semantic contents of the following sentences, the pragmatist would have to have us take out the pronouns in them:

(ii) Mary deposited her paycheck in the bank, but every man lost ~~it~~ (his paycheck).

(iii) The new ruling requires that the semester report which professors write about their students be nonconfidential. However, each professor we interviewed told us that not every student would want to see ~~it~~ (his report about him).

But, first of all, this approach seems rather *ad hoc*, because the pragmatist has never suggested that we can delete the semantic contents of certain words in the sentences we utter. Second, this deletion-based pragmatic account of paycheck pronouns runs counter to her explanation for why we have the audience read things into our utterances instead of making them explicit ourselves. On her view, the reason why we often allow our audience to read things into our utterances is because had we articulated the things that are supposed to be read into our utterances, we would 'slow things down' (2000, p.262) and be handicapped by the 'articulatory bottleneck of linguistic communication' (2000, p.262). But if this explanation is sound, then it's hard to see why we need to utter the pronoun 'it' in order to convey its paycheck reading, if its semantic content is supposed to be deleted by our audience.

the grammatical and the pragmatic approaches to domain restriction. We have seen that a gradable adjective can plausibly be bound by two generalized quantifiers, which strengthens the parallel between adjectival domain restriction and quantifier domain restriction already observed by Kennedy. We have also discussed the benefit of embedding Kennedy's k-rule inside a framework with the g-rule: The resulting framework effortlessly explains the connections between the k-typeshifted meaning of a gradable adjective, which asks for a comparison class, and the meanings that are needed to account for the bound readings of the adjective, which ask for functions from n individuals to a comparison class.

I hope that our discussion also contributes to reorienting various debates on the semantics-pragmatics distinction. It is often said that the aim of semantics is to account for ordinary intuitions about sentences' truth conditions (or truth values). This view about the aim of semantics encourages pessimism about the viability of truth-conditional semantics: The pessimist argues that, since ordinary judgements about the truth value of a sentence vary radically from context to context — even after factors such as indexicality, ambiguity, polysemy, and vagueness have been controlled for — there is no way we can assign truth conditions (meanings) to sentences in a principled way. However, our discussion suggests two reasons why we can do semantics without holding such view about its aim. First, we have seen that ordinary intuitions about sentences' truth conditions (or truth values) are not the only data available to the theorist: The infelicity of 'Jumbo the elephant is big for a fly', the bound readings of 'every member of my family is tall' and 'every government serves every citizen', and the doubly bound readings of Breheny's 'every mistake', Jacobson's 'the drink', and my own 'too slow' are also data for semantics, albeit of a more theoretical sort. Second, we have seen that we are in a position to adjudicate between different semantics-pragmatics packages, even if we take no stand on whether ordinary intuitions about truth conditions are reliable guide to semantic contents. Recall that the main reason we disfavored the pragmatical approach to domain restriction isn't that it doesn't respect ordinary intuitions about truth conditions, but that the pragmatist hasn't shown us her entire semantics-pragmatics package, which makes it very difficult for us to see how her approach can make predictions that are as non-trivial as the ones predicted by the grammatical approach.

Whether we can settle all debates on the semantics-pragmatics distinction by focusing on the predictive and the explanatory power of each theorist's semantics-pragmatics package must be addressed in future work.

2.6 Appendix

The following semantics framework, with the exception of the local saturation convention and Kennedy's typeshifting rule which I call the *k-rule*, is due to Jacobson (1999, 2014). Here I show the syntax and the semantics of our example in §3.2. The essential rules required for the derivation are provided here for the reader's convenience.

Functional Application:

If E_1 is an expression with $< B/A,\ \langle \alpha, \beta \rangle >$, and E_2 is an expression $< A, \alpha >$, then there is an expression E with $< B, \beta >$.

Typeshifting rules:

Let f be the input of these rules.

z: $< (B/C)/A,\ \langle \alpha, \langle \gamma, \beta \rangle \rangle >$

$\Rightarrow < (B/C)/A^C,\ \langle \langle \gamma, \alpha \rangle, \langle \gamma, \beta \rangle \rangle >$, where $z(f) = \lambda X_{\langle \gamma, \alpha \rangle}[\lambda c_\gamma[f_{\langle \alpha, \langle \gamma, \beta \rangle \rangle}(X(c))(c)]]$

g $:< B/A,\ \langle \alpha, \beta \rangle >$

$\Rightarrow < (B^C/A^C),\ \langle \langle \gamma, \alpha \rangle, \langle \gamma, \beta \rangle \rangle >$, where $g(f) = \lambda X_{\langle \gamma, \alpha \rangle}[\lambda c_\gamma[f_{\langle \alpha, \beta \rangle}(X(c))]]$

g′ $:< (C/B)/A,\ \langle \alpha, \langle \beta, \gamma \rangle \rangle >$

$\Rightarrow < (C^D/B^D)/A,\ \langle \alpha, \langle \langle \sigma, \beta \rangle, \langle \sigma, \gamma \rangle \rangle \rangle >$, where $g(f) = \lambda X_\alpha[\lambda Y_{\langle \sigma, \beta \rangle}[\lambda c_\sigma[f_{\langle \alpha, \langle \beta, \gamma \rangle \rangle}(X)(Y(c))]]]$

Kennedy's k-rule:

Let G (with 'G' standing of for 'gradable adjective') be the syntactic type of a gradable adjective and f be its meaning, which is of type $\langle e, d \rangle$:

k $:< G, \langle e, d \rangle >$

$\Rightarrow < G^N, \langle et, ed \rangle >$, where $k(f) = \lambda g_{\langle e,t \rangle}[\lambda x[g(x).f(x)]]$

Local saturation convention:

s $:< A^B,\ \alpha^\beta >$

$\Rightarrow < A, \alpha >$, where $s(f) = f(B)$, and B is a 'contextually given' value of type β

(56) Every team demands (that) no member is too slow.

Syntax:

$$
\begin{prooftree}
\AxiomC{\textit{Every}} \RightLabel{lex} \UnaryInfC{$(S/(S/NP))/N$}
\AxiomC{\textit{team}} \RightLabel{lex} \UnaryInfC{N}
\RightLabel{f.a.} \BinaryInfC{$S/(S/NP)$}
\AxiomC{\textit{demands}} \RightLabel{lex} \UnaryInfC{$(S/NP)/S$} \RightLabel{z} \UnaryInfC{$(S/NP)/S^{NP}$}
\AxiomC{\textit{no}} \RightLabel{lex} \UnaryInfC{$(S/(S/NP))/N$} \RightLabel{g$'$} \UnaryInfC{$(S^{NP}/(S/NP)^{NP})/N$}
\AxiomC{\textit{member}} \RightLabel{lex} \UnaryInfC{N}
\RightLabel{f.a.} \BinaryInfC{$S^{NP}/(S/NP)^{NP}$}
\AxiomC{\textit{is}} \RightLabel{lex} \UnaryInfC{$(S/NP)/(S/NP)$} \RightLabel{z} \UnaryInfC{$(S/NP)/(S/NP)^{NP}$} \RightLabel{g} \UnaryInfC{$(S/NP)^{NP}/(S/NP)^{NP^{NP}}$}
\AxiomC{\textit{too}} \RightLabel{Lex} \UnaryInfC{$(S/NP)/G$} \RightLabel{g} \UnaryInfC{$(S/NP)^{NP}/G^{NP}$} \RightLabel{g} \UnaryInfC{$(S/NP)^{NP^{NP}}/G^{NP^{NP}}$}
\AxiomC{\textit{slow}} \RightLabel{lex} \UnaryInfC{G} \RightLabel{k} \UnaryInfC{G^{N}} \RightLabel{g} \UnaryInfC{$G^{NP^{N^{NP}}}$} \RightLabel{g} \UnaryInfC{$G^{NP^{NP^{N^{NP^{NP}}}}}$} \RightLabel{sat} \UnaryInfC{$G^{NP^{NP}}$}
\RightLabel{f.a.} \BinaryInfC{$(S/NP)^{NP^{NP}}$}
\RightLabel{f.a.} \BinaryInfC{$(S/NP)^{NP}$}
\RightLabel{f.a.} \BinaryInfC{S^{NP}}
\RightLabel{f.a.} \BinaryInfC{S/NP}
\RightLabel{f.a.} \BinaryInfC{S}
\end{prooftree}
$$

Semantics (As the reader refers to the syntax, she is advised to read from the right to the left):

$\llbracket slow \rrbracket = \lambda x[slow'(x)] = \lambda x[x\text{'s degree of slowness}]$ [lex]

$\llbracket slow \rrbracket^{k} = \lambda f_{\langle e,t\rangle}[\lambda x[f(x).\llbracket slow \rrbracket(x)]] = \lambda f_{\langle e,t\rangle}[\lambda x[f(x).slow'(x)]]$ [Kennedy (2007), example #26]

$\llbracket slow \rrbracket^{kg} = \lambda h_{\langle e,et\rangle}[\lambda y[\llbracket slow \rrbracket^{k}(h(y))]] = \lambda h_{\langle e,et\rangle}[\lambda y[\lambda x[h(y)(x).\llbracket slow \rrbracket(x)]]]$ [Geach]

$\llbracket slow \rrbracket^{kgg} = \lambda j_{\langle ee,et\rangle}[\lambda z[\llbracket slow \rrbracket^{kg}(j(z))]] = \lambda j_{\langle ee,et\rangle}[\lambda z[\lambda y[\lambda x[j(z)(y)(x).\llbracket slow \rrbracket(x)]]]]$ [Geach]

$\llbracket slow \rrbracket^{kggs} = \lambda z[\lambda y[\lambda x[J(z)(y)(x).\llbracket slow \rrbracket(x)]]]$

[local saturation convention; J is a two-place contextually given function from a pair of individuals to a set (e.g. a two-place comparison class function that maps a pair of individuals a and b to the comparison class whose membership depends on the standard a sets for the b's age and gender group)]

$\llbracket too \rrbracket = \lambda m_{\langle e,d\rangle}[\lambda x[m(x) \gg \boldsymbol{s}(m)]]$ [lex]

['m' for measure function; '$\gg$' for 'exceeds by an unexpectedly large margin'; $\boldsymbol{s}(m)$ is the contextual threshold for m]

$\llbracket too \rrbracket^{g} = \lambda h_{\langle e,ed\rangle}[\lambda y[\llbracket too \rrbracket(h(y))]] = \lambda h_{\langle e,ed\rangle}[\lambda y[\lambda x[h(y)(x) \gg \boldsymbol{s}(h(y))]]]$ [Geach]

$\llbracket too \rrbracket^{gg} = \lambda j_{\langle ee,ed\rangle}[\lambda z[\llbracket too \rrbracket^{g}(j(z))]] = \lambda j_{\langle ee,ed\rangle}[\lambda z[\lambda y[\lambda x[j(z)(y)(x) \gg \boldsymbol{s}(j(z)(y))]]]]$ [Geach]

$\llbracket too\ slow \rrbracket = \llbracket too \rrbracket^{gg}(\llbracket slow \rrbracket^{kggs}) = \lambda z[\lambda y[\lambda x[J(z)(y)(x).\llbracket slow \rrbracket(x) \gg \boldsymbol{s}(\lambda x[J(z)(y)(x).\llbracket slow \rrbracket(x)])]]]$ [f.a.]

$\llbracket is \rrbracket = \lambda X_{\langle e,t\rangle}[X]$ [lex]

$\llbracket is \rrbracket^{z} = \lambda F_{\langle e,et\rangle}[\lambda a[\llbracket is \rrbracket(F(a))(a)]] = \lambda F_{\langle e,et\rangle}[\lambda a[F(a)(a)]]$ [z-rule]

$\llbracket is \rrbracket^{zg} = \lambda h_{\langle ee,et\rangle}[\lambda b[\llbracket is \rrbracket^{z}(h(b))]] = \lambda h_{\langle ee,et\rangle}[\lambda b[\lambda a[h(b)(a)(a)]]]$ [Geach]

$\llbracket is\ too\ slow \rrbracket = \llbracket is \rrbracket^{zg}(\llbracket too\ slow \rrbracket)$

$= \lambda b[\lambda a[J(b)(a)(a).\llbracket slow \rrbracket(a) \gg \boldsymbol{s}(\lambda x[J(b)(a)(x).\llbracket slow \rrbracket(x)])]]$ [f.a.]

$= \lambda b[\lambda a[J(b)(a)(a).\llbracket slow \rrbracket(a) \gg \boldsymbol{s}(\lambda y[J(b)(a)(y).\llbracket slow \rrbracket(y)])]]$ ['x' is bound; variable change]

$\llbracket member \rrbracket = \lambda x[member'(x)]$ [lex]

$\llbracket no \rrbracket = \lambda X_{\langle e,t\rangle}[\lambda Y_{\langle e,t\rangle}[\forall x(Xx \rightarrow \neg Yx)]]$ [lex]

$\llbracket no \rrbracket^{g'} = \lambda X_{\langle e,t\rangle}[\lambda T_{\langle e,et\rangle}[\lambda c[\forall x(Xx \rightarrow \neg T(c)(x))]]]$ [Geach$'$]

$\llbracket no\ member \rrbracket = \llbracket no \rrbracket^{g'}(\llbracket member \rrbracket) = \lambda T_{\langle e,et\rangle}[\lambda c[\forall x(member'(x) \rightarrow \neg T(c)(x))]]$ [f.a.]

$[\![no\ member\ is\ too\ slow]\!] = [\![no\ member]\!]([\![is\ too\ slow]\!])$

$\lambda c[\forall x(member'(x) \rightarrow J(c)(x)(x).[\![slow]\!](x) \not\succ \boldsymbol{s}(\lambda y[J(c)(x)(y).[\![slow]\!](y)]))]$ [f.a.]

$[\![demands]\!] = \lambda p_t[\lambda x[demand'(x,p)]]$ [lex]

$[\![demands]\!]^z = \lambda F_{\langle e,t\rangle}[\lambda e[[\![demands]\!](F(b))(b)]] = \lambda F_{\langle e,t\rangle}[\lambda e[demand'(e, F(e))]]$ [z-rule]

$[\![demand\ no\ member\ is\ too\ slow]\!] = [\![demands]\!]^z([\![no\ member\ is\ too\ slow]\!])$

$= \lambda e[demand'(e, \forall x(member'(x) \rightarrow J(e)(x)(x).[\![slow]\!](x) \not\succ \boldsymbol{s}(\lambda y[J(e)(x)(y).[\![slow]\!](y)])))]$ [f.a.]

$[\![team]\!] = \lambda x[team'(x)]$ [lex]

$[\![Every]\!] = \lambda X_{\langle e,t\rangle}[\lambda Y_{\langle e,t\rangle}[\forall x(Xx \rightarrow Yx)]]$ [lex]

$[\![Every\ team]\!] = [\![Every]\!]([\![team]\!]) = \lambda Y_{\langle e,t\rangle}[\forall x(team'(x) \rightarrow Yx)]$ [f.a.]

$= \lambda Y_{\langle e,t\rangle}[\forall z(team'(z) \rightarrow Yz)]$ ['x' is bound; variable change]

$[\![Every\ team\ demand\ no\ member\ is\ too\ slow]\!] = [\![Every\ team]\!]([\![demand\ no\ member\ is\ too\ slow]\!])$

$= \forall z(team'(z) \rightarrow demand'(z, \forall x(member'(x) \rightarrow J(z)(x)(x).[\![slow]\!](x) \not\succ \boldsymbol{s}(\lambda y[J(z)(x)(y).[\![slow]\!](y)]))))$ [f.a.]

[Notice that the contextual standard varies with each team z and each member x, as desired. To simplify the derivation, the quantifier domains of 'every' and 'no' are assumed to be unrestricted, but they can easily be restricted using the same tools that handle adjectival domain restriction, i.e. the c-rule (49) or the d-rule (52), and the g-rule.]

Chapter 3

Against Indexical Predicates

3.1 Introduction

Let's set aside tense, intensions, and the formal representations of context, and assume that truth-conditional semantics is concerned with constructing extensional fragments for natural languages.[1] So understood, truth-conditional semantics says that the truth value of each sentence is a function of the extensions of its parts and how those parts are put together syntactically, but nothing else. For example, the truth value of 'Porky escapes' is determined solely by the extension of 'Porky', which is Porky himself, the extension of 'escapes', which is the set of escapees, and the syntax of the sentence, which determines that the sentence is true just in case the extension of 'Porky' is a member of the extension of 'escapes'. That every sentence's truth value is compositionally derived from the extensions of its parts (in a way just like the truth value of 'Porky escapes') is a central tenet of truth-conditional semantics, because, without it, truth-conditional semantics is unable to explain our ability to understand sentences we haven't encountered before.

Some theorists, while they believe that the semantics ought to be compositional in some sense, are skeptical of there being a deterministic mapping from the extensions of a sentence's parts to that sentence's truth value. According to them, whether or not a sentence contains indexicals (e.g. 'I', 'now'), we can't speak of its truth value without situating it in the context in which it is uttered, and its truth value relative to a given context depends not only on the extensions of its parts and the way they are put together syntactically, but also on a myriad of

[1] Here I follow Rothschild and Segal's assumptions about truth-conditional semantics.

factors about the context, such as the intentions and the interests of the discourse participants, and what is perceptually or conversationally salient to them.[2]

Examples like the following, originally due to Charles Travis, are often said to lend considerable support to those theorists' view:

(1) The leaves are green.
Contexts: Pia's Japanese maple is full of russet leaves. Believing that green is the colour of leaves, she paints them. Returning, she reports, 'That's better. The leaves are green now.' She speaks truth. A botanist friend then phones, seeking green leaves for a study of green-leaf chemistry. 'The leaves (on my tree) are green,' Pia says. 'You can have those.' But now Pia speaks falsehood (Travis 1997).

Here is what is special about these cases. Each of these examples features a sentence in natural language (e.g. the leaves are green) and presents us with two contexts (which shouldn't be confused with their formal representations such as Kaplanian formal contexts).[3] It is usually set up in such a way that there is some plausibility to the claims that the featured sentence has different truth values in the two contexts, and that the change in truth value can't be explained in terms of indexicality, ambiguity, polysemy, vagueness, or non-literality. Once we accept these two claims, the skeptics of truth-conditional semantics can then argue that there is not always a deterministic mapping from the extensions of a sentence's parts to that sentence's truth value. Here I illustrate how this argument would go by using Travis's example:[4]

(2) a. The sentence 'The leaves are green' is true in the context of Pia's soliloquy, but false in the context of her conversation with her botanist friend.
[based on Travis's story]

b. The extension of 'are green' remains constant, because the two uses of the sentence don't involve indexicality, ambiguity, polysemy, or non-literality, and the vagueness of 'green' plays no essential role in determining what the sentence conveys in each use.

[2] See Bezuidenhout (2002) and the authors who defend the position that is often loosely referred to as 'truth-conditional pragmatics' or 'radical contextualism'.

[3] A Kaplanian formal context is a 4-tuple containing an agent, a location, a time, and a possible world. See Kaplan (1989).

[4] My statement of the first premise is based on Rothschild and Segal's formulation (2009, p.469), but they use 'Keble College is red' as their example.

c. The extension of 'the leaves' remains constant, because it refers to the same objects in both contexts.[5]

d. Therefore, there is not a deterministic mapping from the extensions of 'the leaves' and 'are green' to the truth value of 'the leaves are green'.
[from (a) – (c)]

I should note that these cases (which we will call 'Travis cases' from now on) need not be based on color adjectives (e.g. 'green', 'red'). The reason why they seem to pose a general problem to truth-conditional semantics is that, with enough ingenuity, they can be constructed based on common nouns (e.g. 'baseball'), mass nouns (e.g. 'beer', 'milk'), and verbs (e.g. 'weighs').[6] So they seem to call for a general response from those of us who think that truth-conditional semantics is a viable research program.

Let's consider some early responses by focusing on the argument I stated above. Cappelen and Lepore (2005) would deny the very first premise, because they think that ordinary truth value intuitions are about one of the many possible propositions 'the leaves are green' could pragmatically convey,[7] rather than about the sentence's semantic value, which they argue never vary from context to context.[8] Their view happens to be rather unpopular, because it is often said that it denies to truth-conditional semantics its very data by significantly downplaying the theoretical relevance of ordinary truth value intuitions to semantic values.[9] The most popular option is to deny the second premise by identifying ambiguity or indexicality in the sentence. For example, according to Szabo (2001), 'green' has a hidden indexical that asks for the relevant part at which a given object is supposed to be green, and another hidden indexical that asks for the relevant comparison class.[10]

In their recent response to the subclass of Travis cases based on color adjectives, Rothschild

[5]This premise blocks the trivial response that the first utterance of 'the leaves' refers to the leaves together with the paint, and the second refers to the leaves underneath the paint.

[6]See Travis (1985), Bezuidenhout (2002), Cappelen and Lepore (2005), among others, for more examples of this kind.

[7]Cappelen and Lepore call the propositions I refer to as 'what sentences pragmatically convey' *speech act contents*. I use my terminology instead of Cappelen and Lepore's because it is likely to be more familiar to the reader. I refer the reader to Cappelen and Lepore (2004,2005) for their own discussions of speech act contents.

[8]Cappelen and Lepore hold that there is an exhaustive list of context-sensitive expressions, and that proper names and gradable adjectives are not on that list. They propose three tests of context-sensitivity, and argue that only those expressions on their list pass those tests. But see Szabo (2006) for worries about those tests.

[9]See King and Stanley (2007, p.160) for an elaboration on this objection. See also Stanley and Szabo (2000a, p.240).

[10]See also Hansen (2011) for another example of this strategy.

and Segal (2009) also choose to deny the second premise, but, unlike Szabo, they make color adjectives full-fledged indexicals with minimal semantic constraints on their possible extensions. They defend their account primarily by arguing against Cappelen and Lepore's and Szabo's accounts. Their main objection to Cappelen and Lepore is methodological: By allowing the semantic value (i.e. truth value) of sentences like 'the leaves are green' to vary from context to context, they contend that their account is in a better position to preserve the methodological assumption that 'our intuitions about the truth conditions [truth values] of utterances provide reliable data for a semantic theory' (p.474). Their main objection to Szabo is empirical: They contend that the 'theoretical strength' of their account is at least as good as Szabo's, and that their proposed syntax for color adjectives is simpler.

The main goals of this paper are to argue that Rothschild and Segal's objections do not go through, and to propose, in light of the failure of their objections, an alternative response to Travis cases. Here is how we are going to proceed: I first argue that Rothschild and Segal's objection to Cappelen and Lepore fails, because a component about explanation or prediction is missing in their methodological assumption about the evidence of semantics, and, once we fill in that component, their assumption is in fact neutral between their account and Cappelen and Lepore's. I then argue that Szabo's account, with a small modification, can make predictions that can't be made by Rothschild and Segal's account. Finally, I will argue that we can effectively defend semantics against Travis cases by clarifying the relation between ordinary intuitions, linguistic phenomena, and explanation and prediction.

The rest of our discussion is structured as follows. In the next section, we discuss Rothschild and Segal's account of color adjectives. I then assess their objection to Cappelen and Lepore's account (§3.3), and discuss their objections to Szabo's account and how Szabo could respond to them (§3.4). After that, I propose an alternative response to Travis cases (§3.5). We conclude our discussion in the last section (§3.6).

3.2 Rothschild and Segal on Indexical Predicates

The following example is the target of Rothschild and Segal's analysis:

(3) a. **The Greengrocer**
The greengrocer stocks two types of watermelons. Both types are green on the outside, one has red flesh and the other has yellow flesh. A customer asks for a red watermelon. The greengrocer points to one and says, 'How about this one? It's red' (p.467).

b. **The Artist's Studio**
An artist is painting a still life. On his desk is a red-skinned apple and that same watermelon, still green on the outside and red-fleshed. The artist points to the apple and says 'It's red'. He then points to the watermelon and says, 'It's not red' (p.468).

Let's christen the watermelon 'Melon'. We can observe that this example has the features characteristic of a Travis case: There is some plausibility to the claim that the featured sentence 'Melon is red' is true in the Greengrocer context, but false in the Artist's Studio context. And it seems quite plausible that we can't explain the change in truth value in terms of indexicality, ambiguity, polysemy, vagueness, or non-literality. As I explained above, these claims jointly pose a *prima facie* challenge to truth-conditional semantics.

Rothschild and Segal deny the second claim by proposing that, contrary to appearance, 'red' is itself an indexical which changes its extensions across contexts based on the contexts' conversational standards (p.472).[11] The notion of context they have in mind is meant to be technical rather than intuitive: The tokens of 'red' that are used in the Greengrocer scenario are grouped together under a particular syntactic subtype (e.g. red$_{greengrocer}$). Likewise for the tokens of 'red' used in the Artist's Studio scenario. Tokens that belong to a certain syntactic subtype can only exist within a context in their technical sense.[12] Let's distinguish contexts in their sense from the intuitive notion of context we have been operating with by calling them *t-contexts* ('t' for technical) — and we will continue to use the word 'context' to refer roughly to the situation in which an utterance is made. If we grant Rothschild and Segal that two distinct t-contexts can be assigned to the Greengrocer and the Artist's Studio contexts, they are able to explain why 'red' can have different extensions in them.

Here is a word about their formal semantics. Let's suppose that 'red$_g$' is the syntactic

[11] I assume that Rothschild and Segal's aim isn't to argue that their account of color adjectives, which constitute a subclass of gradable adjectives, is the most empirically adequate account out there. Its major motivation, as I understand it, is to provide a response to a significant subclass of Travis cases, and to argue that it is more adequate than Cappelen and Lepore's and Szabo's accounts, which can plausibly be interpreted as the leading responses to Travis cases. So we are primarily concerned with the relative empirical adequacy of Rothschild and Segal's account against Cappelen and Lepore's and Szabo's. How their account fares against other accounts of gradable adjectives (Cresswell (1977), Klein (1980), Kennedy (1999, 2007), Heim (2000), Kennedy and McNally (2005, 2010), among others) is beyond the scope of our discussion.

[12] See pp.471-472, and fn. 10, where they say that 'context' is a 'slightly technical term'.

subtype under which the tokens of 'red' in the Greengrocer context are grouped. They construct a simple T-theory which proves theorems of the following form:

(4) When it is uttered in the t-context assigned to the Greengrocer scenario, 'Melon is red_g' is true if and only if Melon is red_g

Theorems like this present the truth conditions of 'Melon is red' when it is uttered in different t-contexts. They also represent the indexicality of 'red', because the relevant syntactic subtype of 'red' varies from t-context to t-context. However, since nothing in Rothschild and Segal's defense of their position rests on this particular formal implementation, we'll make our later comparison between their account and Szabo's easier by recasting their formal semantics using the Montagovian framework adopted by Szabo.[13] Since 'red' is an indexical, we can assume that its extension (i.e. the set of red objects), just like the value of a variable, is always determined by an assignment function given by a t-context.[14] So instead of using T-theorems to represent its indexicality, we use the following simple lexical item to do so:

(5) $[\![red]\!]^{\sigma} = \sigma$('red') = the extension of 'red' as determined by the assignment function σ, where the value of σ is determined by a t-context

One additional benefit that comes with this recasting is that Rothschild and Segal can now do away with t-contexts and avoid the theoretical costs involved in positing them. The most plausible reason why they introduce t-contexts and syntactic subtypes of 'red' is that they want to give the truth-conditions of 'Melon is red' recursively by proving the theorem I stated above. But their introduction comes at a cost. First, it is unclear from Rothschild and Segal's description of t-contexts how they can reliably map each context presented in a Travis case to a t-context that determines the same extension for 'red' (or an extension that is sufficiently similar to the one determined by that context). If there is no guarantee that two suitable t-contexts

[13] The Montagovian framework is represented by the following textbooks: Dowty (1979), Heim and Kratzer (1998), Chierchia and McConnell-Ginet (2000), Jacobson (2014). The framework Rothschild and Segal adopt, which assigns truth-conditions to sentences by proving for them theorems like the one I showed above, is represented by Larson and Segal's (1995).

[14] This assumption is harmless, because their examples don't involve indexicals in intensional contexts (e.g. Everything now beautiful will become more beautiful).

can be assigned to their Greengrocer and Artist Studio contexts, then it is difficult to see how t-contexts help explain the change in the truth value of 'Melon is red'. Second, the syntactic subtypes of 'red' are in their object language, so their account appears to have the implausible commitment that 'red' is massively ambiguous between its possible syntactic subtypes. In order to focus on the most central part of Rothschild and Segal's account, we will iron out these wrinkles by assuming that the value of the assignment function σ is determined by a context rather than a t-context.

Rothschild and Segal have mentioned a number of *prima facie* virtues of their view (p.474), such as its avoidance of hidden logical forms, and its potential generalizability to a few other predicates (e.g. 'tall', 'know', 'water'). But we will only focus on their claim that their account preserves 'the idea that our intuitions about the truth conditions [or truth values] of utterances provide reliable data for a semantic theory' (ibid.), and their claim that their account covers at least as much data as Szabo's account but does so with a simpler syntax for color adjectives. Those are the main reasons they have for objecting to Cappelen and Lepore's and Szabo's accounts. Assessing them is the task for the next two sections.

3.3 Rothschild and Segal's Methodological Objection to Semantic Minimalism

Let me flesh out Cappelen and Lepore's response to Travis cases before we discuss Rothschild and Segal's objection to it. According to Cappelen and Lepore, both proper names and color adjectives are context-insensitive in the sense that their extensions do not vary from context to context. Since they hold that the truth value of a sentence (i.e. its semantic value) is compositionally derived from the extensions of its parts, they hold that the truth value of 'Melon is red' does not vary from context to context. But they allow what that sentence pragmatically conveys to vary across contexts depending on a myriad of contextual factors, such as the intentions of the speaker, and what the speaker and the hearer know about their previous conversations. They respond to Travis cases by suggesting that ordinary intuitions about the featured sentences' truth values in different contexts may be about what those sentences pragmatically convey, rather than about their semantic values.

Here is Rothschild and Segal's main objection to Cappelen and Lepore's view:

> One problem with this view is that it rejects the idea that intuitions about the truth values of utterances of sentences provide good evidence about those sentences' truth conditions [semantic values]. This makes it very hard to understand what the data for a semantic theory are supposed to be. In other words, it makes it very hard to see how one is supposed to tell whether a semantic account of some fragment of natural language is correct (pp.469-470).

Before we assess this objection, a word about 'truth conditions' is in order. Since Rothschild and Segal assume as we do that semantics is extensional semantics, there are two plausible things they can mean by the term. It can either mean sentences' truth values (i.e. their semantic values), or specifications of how sentences' semantic values are compositionally derived from their parts' semantic values such as the followings:

(6) 'Porky escapes' is true if and only if the extension of 'Porky' is a member of the extension of 'escapes' (Truth conditions in Montagovian extensional semantics).

(7) 'Porky escapes' is true if and only if Porky escapes (Truth-conditions in T-theory).

Either way, it refers to certain theoretic constructs of a semantics fragment, rather than contents that are available to ordinary speakers through introspection. Here I adopt the first interpretation, but my concern with Rothschild and Segal's objection still stands if we adopt the second interpretation.

Rothschild and Segal's objection seems to be the following:

(8) a. Since ordinary intuitions about the truth value of 'Melon is red' vary with the contexts in which it is uttered, and since it is a good methodological practice to assume that those intuitions 'provide good evidence about' that sentence's semantic value, there is a *prima facie* reason to favor an account that allows the semantic value of 'Melon is red' to vary across contexts over an account that doesn't.

b. Their account allows the semantic value of 'Melon is red' to vary across contexts and Cappelen and Lepore's doesn't.

c. Therefore, there is a *prima facie* reason to favor their account.

This objection seems compelling at first blush, but whether the first premise holds depends on what it is for ordinary intuitions about the truth values of sentences uttered in different contexts to 'provide good evidence about' semantic values. Let's rule out what Rothschild and Segal can't mean by 'good evidence' by considering the following toy indexical account of 'red': Suppose a hypothetical semanticist constructs a certain number of Travis cases featuring the sentence 'Melon is red' — let's assume that there is no upper limit on how many Travis cases she can construct. She then sends out surveys to solicit her respondents' intuitions about that sentence's truth values in the contexts described by her Travis cases. If a high enough percentage of her respondents agree on that sentence's truth value at a certain context, she assigns that truth value to that sentence relative to that context, and makes sure that compositionality holds by adjusting the semantic values of 'red' accordingly. She takes no stand on the truth value of that sentence relative to the contexts that are either not covered by her survey or fail to meet a certain percentage threshold. (She could say, as many semanticists do, that those contexts determine the extensions of 'red' and hence the truth values of the sentence).

This account not only allows the semantic value of 'Melon is red' to vary across contexts, but also ensures a very good agreement between the sentence's semantic values in different contexts and ordinary truth value intuitions.[15] But we are reluctant to say that the respondents' truth value intuitions provide good evidence for the semantic values the semanticist assigns to the sentence, because her account merely reports, without explaining or predicting, the data she obtains from her survey. So what Rothschild and Segal mean by 'good evidence' must include a component about explanation or prediction. Here I my attempt to capture what they

[15]It is one thing for the truth values of 'Melon is red' to vary across contexts. It is another for them to agree with ordinary intuitions about them. Rothschild and Segal's indexical account can only guarantee the former. They can't guarantee the latter, because it is possible that the truth values of 'Melon is red' varies across contexts in a way that is severely at odds with ordinary truth value intuitions. In order to guarantee the latter, they need to make the controversial assumption that ordinary speakers have reliable intuitions about sentences' truth values (i.e. semantic values).

mean by ‘good evidence’:

(9) *No good evidence without good explanation or prediction:*

Ordinary truth value intuitions are good evidence for a semantics fragment only if they suggest some phenomenon that can be explained or predicted well by that fragment.

The idea behind this rendering is that no observations are good evidence for a given theory unless that theory provides a good explanation or prediction about some phenomenon suggested by those observations. For example, our observations of various combustion events suggest the phenomenon that combustion is more efficient in open space than in enclosed space. But those observations aren’t evidence for any theory until some theory provides an explanation or prediction for that phenomenon (or other phenomena suggested by those observations). So explanation or prediction of phenomena is necessary for evidencehood. Although the phenomenon that combustion is more efficient in open space can be explained by the phlogiston theory,[16] we don’t say that the observations that suggest it are good evidence for that theory, because the oxygen theory offers a better explanation for it. So good explanations or predictions make good evidence. Like our observations of various combustion events, ordinary truth value intuitions aren’t good evidence for a given semantic fragment unless it suggests some phenomena about which that fragment can provide good explanations or predictions.[17] In this discussion, I take no stand on the precise scope of the phenomena that are plausibly suggested by ordinary intuitions. But I assume that they at least include the unacceptability, ambiguity, entailment phenomena we saw presented in semantics textbooks and journal articles.

If this is how we should understand ‘good evidence’, we need to reassess whether Cappelen and Lepore have failed to respect some good evidence for semantics. Which phenomena are plausibly suggested by ordinary intuitions about the truth values of ‘Melon is red’ in different contexts? Let’s consider the following two possible candidates and their consequences for

[16] According to the phlogiston theory, combustion is due to the release of phlogistons from the fuel to the air, which carries away the phlogistons released. Since the air has a limited capacity to carry away phlogistons, combustion becomes less efficient when the air is saturated with phlogistons.

[17] See also Ludlow (2011, Ch.5), from whose discussion I learn the distinction between ordinary intuitions and phenomena. On his view, ordinary intuitions are evidence for phenomena, which in turn are evidence for a given semantic theory in virtue of being explained or predicted by that theory. My own contribution is to give a plausible characterization of when intuitions are good evidence for a semantic theory based on Ludlow’s distinction between ordinary intuitions and phenomena.

Rothschild and Segal's objection:

(10) a. A particular distribution of truth values over different contexts (e.g. the Greengrocer, the Artist's Studio)

b. The phenomenon that ordinary intuitions about the truth value of 'Melon is red' tend to vary across contexts

Let's start with the first candidate.[18] Recall from our discussion in the last section that Rothschild and Segal's indexical account says that the extension of 'red' depends on the conversational standard of a context. But they don't offer an account of how it does so. For example, they don't explain what a conversational standard is, how a context fixes its conversational standard, or how conversational standards fix the extensions of 'red'. Since they don't have an account of how contexts determine the extensions of 'red', they can't explain or predict the specific truth values of 'Melon of red' in different contexts. So the first candidate can't be what Rothschild and Segal has in mind as they level their objection against Cappelen and Lepore's account.

The second candidate looks more promising as the basis for Rothschild and Segal's objection, because they could explain and predict it in terms of indexicality resolution and the general reliability of ordinary intuitions about the truth values of sentences: Since 'red' is an indexical, its extension varies across contexts, and, due to compositionality, so do the truth value of 'Melon is red'.[19] And since ordinary speakers' intuitions about the truth values of sentences are generally reliable, their intuitions about the truth value of 'Melon is red' vary across contexts.

If this explanation is clearly the best explanation for the phenomenon, Rothschild and Segal can rightly argue that Cappelen and Lepore fail to respect some good evidence for semantics. However, Cappelen and Lepore could argue that it isn't the best explanation along the following lines. First, it is not clear that ordinary speakers are generally reliable about

[18] It is not clear whether ordinary speakers do uniformly judge 'Melon is red' to be true in the Greengrocer context and false in the Artist's Studio context, or whether their judgments remain the same if we modify the ways in which the contexts are presented to them. Here I assume for the sake of argument that it is possible to infer a particular distribution of truth values over different contexts from the ordinary truth value intuitions solicited by Travis cases like Rothschild and Segal's.

[19] Here I assume that the variations in the extension of 'red' are large enough to cause variations in the truth value of 'Melon is red' across contexts.

the truth values semanticists assign to sentences. For example, ordinary speakers tend to think that maximal standard absolute adjectives, such as 'flat', can be true of many objects (e.g. blackboards, Holland). But, according to Kennedy (2007)'s theory of gradable adjectives, every maximal standard absolute adjective is true of an object only if that object possesses the property denoted by that adjective to the maximal degree, which means that maximal standard absolute adjectives are false of most objects, contrary to ordinary intuitions. The reason why Kennedy can justifiably allow a discrepancy between the truth values he assigns to sentences and ordinary truth value intuitions is that his theory doesn't derive its empirical adequacy from a good agreement between sentences' truth values and ordinary truth value intuitions, but from its ability to explain or predict phenomena that are suggested by those intuitions, such as the entailment between 'The glass is FULL' and 'The glass can't be fuller'.

Second, we should note that Cappelen and Lepore's overall theory, unlike Rothschild and Segal's, isn't meant to apply only to color adjectives, but to every predicate in natural language. Since they think that it is easy to construct Travis cases based on any given predicate, and that it is implausible that most predicates in natural language are indexicals, they would argue that Rothschild and Segal's explanation is unattractive on the grounds that it is liable to making every predicate that can be featured in a Travis case an indexical.

So Cappelen and Lepore are in a good position to answer Rothschild and Segal's charge that they fail to respect some good evidence for semantics. Perhaps Rothschild and Segal's best line of objection is to demonstrate that their account has some theoretical virtue that Cappelen and Lepore's lacks. Here is a suggestion about how their objection could have gone. Since both sides feel compelled to explain the phenomenon that ordinary intuitions about the truth value of 'Melon is red' tend to vary across contexts, Rothschild and Segal can object to Cappelen and Lepore's overall theory on parsimony grounds: Whereas Rothschild and Segal can explain that phenomenon by directly appealing to the nature of indexicality resolution, Cappelen and Lepore need to supplement their semantics fragment with a pragmatic explanation for that phenomenon, and to explain away the appearance that the (context-insensitive) semantic value they assign to 'Melon is red' play no essential role in their explanation of that phenomenon.

That said, this objection is only compelling if we restrict ourselves to color adjectives. If we focus on every predicate that can be featured in a Travis case instead, and think as Cappelen and Lepore do that it is implausible that most predicates in natural language are indexicals, then it becomes much harder to decide which account is preferable.

3.4 Indexical Predicates vs. Hidden-Indexicals

Before we assess Rothschild and Segal's case against Szabo's account, let me explain the motivation of Szabo's account, and contrast it with Rothschild and Segal's. As I said in the introduction, according to Szabo, each color adjective has a comparison class variable, which asks for the relevant comparison class, and a part variable, which asks for the relevant part at which an object has the color referred to by the adjective. The reason he posits the part variable is that he wants to account for the truth values of 'the leaves are green' in the two contexts presented by Travis in a compositional way: Since the part variable can pick out different locations at which the leaves are green in different contexts (e.g. the surface of the leaves vs. the leaves without the paint), the extension of 'green' is allowed to vary across contexts. This means that he can explain the change in the sentence's truth value solely in terms of the extensions of its parts and the way they are put together syntactically.

The following shows the meaning he would assign to 'red', which we can contrast with Rothschild and Segal's proposal:

(11) a. $[\![red(C)(p)]\!]^{\sigma}$ = the set of objects that are red in part $\sigma(p)$ relative to comparison class $\sigma(C)$, where the value of σ is determined by a context (Szabo).

b. $[\![red]\!]^{\sigma} = \sigma$('red') = the extension of 'red' as determined by σ, where the value of σ is determined by a context (Rothschild and Segal).

I should emphasize that what distinguishes Szabo's proposal from Rothschild and Segal's isn't so much the variables he posits,[20] but the additional semantic or type-theoretic constraints he puts on the possible extensions of color adjectives: Whereas the extension of Szabo's 'red' at least depends on the value of the comparison class argument C, the extension of Rothschild and

[20] It is possible to recast Szabo's proposal using Jacobson's variable-free semantics (1999, 2014). So what matter aren't the variables, but the type-theoretic constraints on the possible extensions of color adjectives they entail.

Segal's 'red' is solely determined by the assignment function given by a context. This difference between their account is important, because we are going to see that the comparison class argument in Szabo's account, with a slight modification, makes his account more predictively adequate than Rothschild and Segal's.

We are now ready to assess Rothschild and Segal's objections to Szabo's account. One of their main contentions is that the syntax Szabo posits for 'red' is more complicated than the one they posit, and unnecessarily so. To argue for that claim, they take great lengths to argue that the comparison class variable and the part variable on Szabo's account can't be bound and are hence not well justified. The followings are some of their examples:

(12) Only one collection of fruit$_i$ was next to exactly one red$(c_i)(p)$ colour tile.

(13) Every face$_i$ says the cube is red(c)(p$_i$).

I agree fully with their assessment that the comparison class argument isn't bound in (12). And I think that the comparison class argument itself can never be bound. But they shouldn't be surprising: We usually only expect individual variables to be bindable. Since the comparison class variable asks for a set rather than an individual, it isn't clear we should expect it to be bindable, any more than we should expect the relation variable in 'John's book' to be bindable.[21]

I also agree fully with their assessment that the part variable can't be explicitly bound. This does seem to suggest that Szabo owes us an explanation why that variable is unlike the typical individual variables (e.g. the implicit argument in 'local'), which are bindable.[22] Here is an additional problem with Szabo's part variable: It can't handle every Travis case based on color adjectives. For example, we can construct a Travis case based on the observation conditions like this:[23]

(14) A 50-page white book and a yellow book are illuminated by red light and blue light respectively. So the white book looks red and the yellow book looks green. Consider two scenarios. In the first scenario, both the speaker and the hearer don't know the true

[21]See Stanley (2000, p.429).
[22]See Partee (1989).
[23]See Hansen (2011) for further examples of Travis cases that can't be handled by Szabo's part variable.

colors of the books. So the speaker refers to the books' colors as they are perceived by her and her hearer. And she appears to speak truly if she utters 'the red one is 50-page long'. In the second scenario, both the speaker and the hearer know the true colors of the books, and the hearer is interested in getting a short book of the right color to decorate her Christmas tree. So the speaker appears to speak falsely if she utters the same sentence.

Szabo's account is unable to handle this Travis case, because the change in truth value of the sentence isn't due to a shift in the relevant part at which the 50-page book has its color. So Szabo may need to posit an additional variable to handle this case. But Rothschild and Segal would complain that, since Szabo's 'red' has at least two individual variables and their 'red' has none, other things being equal, their account is preferable to Szabo's on parsimony grounds.

Let's now assess whether both accounts have the same explanatory and predictive power. Rothschild and Segal acknowledge that their indexical account is unable to explain the bound readings of 'red' if they exist.[24] So, in order to hold on to their claim that their account has sufficient theoretical strength to rival Szabo's account, they have taken even greater lengths to convince us that 'red' never has a bound reading. This claim is our focus for the rest of this section.

Rothschild and Segal claim that the following example doesn't show that 'red' is bindable:

(15) Every kind of animal in the zoo has a member that is red.

They acknowledge that what is red enough for a goldfish is different from what is red enough for a squirrel. But they argue that they can account for that reading without positing a bindable element in 'red', because, thanks to the minimal semantic constraints they put on its possible extensions, 'red' can pick out the set of objects each of which is 'canonically red' for its kind (p.479). Their idea is that 'red' can pick out the property expressed by 'x is canonically red for x's kind', and that the kind goldfish has a member that satisfies that property, and so does the

[24]'In this paper we give a simple indexical semantics for predicates like 'red' and 'hexagonal'. If these predicates exhibit something analogous to the bound usages we find in (1) ['Every man likes his mother'] then our semantics will not be able to account for that' (p.476).

kind squirrel.

However, there are two issues with this argument, which we will discuss one by one. The first issue is that it is faced with the *Bo Jackson problem* discussed by Stanley (2007b) in response to a similar explanation for the bound readings of 'old'. The second issue is that it fails to predict that 'red' picks out that particular extension rather than any possible extensions we can think of, and that Szabo's account, with a slight modification, is actually in a better position to predict that.

This is the Bo Jackson problem. Consider the sentence 'Every sports team has a member who is old',[25] which is structurally similar to Rothschild and Segal's example. Since what is old for equestrian sports is different from what is old for gymnastics, Rothschild and Segal might say that 'old' picks out the set of persons that are canonically old for her sport. However, talented athlete such as Bo Jackson can play two sports. If he is assessed for his relative oldness based on the standards of his football team and his baseball team, it won't do to ask whether he is old for his *sport* (since he plays two sports). It won't do to ask whether he is old for his *sports* either, because he could be too old for one sport without being too old for the other. So we need a separate standard of oldness for each of the sports teams he is in. Rothschild and Segal may have deliberately chosen an example which makes it difficult for us to construct a parallel counterexample, because an animal rarely belongs to two kinds. But since we can easily construct an example where an object is evaluated for its relative redness against two different groups (e.g. Every artist has found a model whose favorite dress is red enough), I don't think that they can avoid the Bo Jackson problem.

Hawthorne (2007) has suggested a solution to the problem that is friendly to Rothschild and Segal's account: The common noun 'member' in Stanley's example, he suggests, can pick out *guises* of individuals, rather than individuals. The thought is that 'member' can double-count Bo Jackson by including in its extension both Bo Jackson the football player and Bo Jackson the baseball player. Since each of these ghostly individuals plays only one sport, we can obtain the intended reading of Stanley's example by having 'old' pick out the set of guises that are canonically old for her sport.

[25] Stanley (2007b), p.217.

There is no need to dispute whether 'member' in Stanley's example can pick out guises (let's assume that it can), because we can exorcise them by modifying Stanley's example as 'Every sports team has an old member who plays in multiple teams'. Since no guise plays in multiple teams by definition, and the sentence need not be read as a contradiction, 'member' must pick out normal persons this time. But since the modified sentence can still be read as saying that Bo Jackson is old by both the standard of his football team and the standard of his baseball team, the Bo Jackson problem can't be solved by invoking guises.

Let's now move on to the second issue with their argument. Upon arguing that 'red' in their zoo-animal sentence (15) has the reading they claims it has, they argue against Szabo's account on the grounds that it fails to deliver a similar reading of 'tall':

(16) John is tall and Jenny is too. (p.483)

(17) Children who are tall are generally well nourished. (p.484)

The context of their first example is that John is a four-year-old and Jenny is a fourteen-year-old. It has the reading that John is tall for his own age group, and Jenny is tall for her own age group. Their second example has a similar reading on which the children are tall for their own ages. We can notice that the relevant reading of 'tall' in these examples is just like their intended reading of 'red' in the zoo-animal sentence (15): Whereas the extension of 'red' is the set of objects each of which is canonically red for *its* kind, the extension of 'tall' is the set of persons each of whom is tall for *her* age. What is shared by these readings is that the objects that satisfy the gradable adjective varies with the individuals to which that adjective applies (which strongly suggests that the readings involve binding). Rothschild and Segal contend that these 'readings' are compatible with their account, but not with Szabo's account, because, since the comparison class argument in Szabo's account only asks for a single comparison class, it appears that he can't allow the set of objects that satisfy a gradable adjective to vary with the individuals to which that adjective applies.

However, Rothschild and Segal have overlooked an easy repair for Szabo's account. As we discussed in the second chapter, hidden variables, such as the domain variable in quantifiers, can be functional. So Szabo could have posited in 'red' a functional comparison class variable that maps an individual to its relevant comparison class, and modify the lexical meaning of 'red' in the following way:

(18) $[\![red\, C(x)]\!]^{\sigma}$ = the set of objects that are red relative to comparison class $\sigma(C)(\sigma(x))$, where the value of σ is determined by a context.

This slight modification improves the theoretical strength of Szabo's account considerably. He can now respond to Rothschild and Segal's objection by analyzing their (16) along the following lines:

(19) a. John x is [tall $C(x)$] and Jenny y is [tall $C(y)$].

b. Let A be a function that maps a person to people of his/her age. Both John and Jenny are members of the extension $\{y$: y is tall relative to the comparison class $A(y)\}$.

The idea is that is that we have 'John' and 'Jenny' bind the individual variable in 'tall',[26] and to have the audience (or a context) supply the function A, which maps a person to people of his/her age. We obtain the reading intended by Rothschild and Segal because both John and Jenny fall into the set each of whose member is tall relative to people of his/ her age. A similar explanation applies to their (17).

The modified account isn't vulnerable the Bo Jackson problem, because the relevant comparison class for a given individual can vary with the teams when the individual variable is bound by the generalized quantifier 'every sports team':

(20) Every sports team x has a member who is [old $C(x)$].

Let me spell this out: Let's suppose the variable C is saturated by the function M, which maps a team to its members. When Bo Jackson is evaluated for his relative oldness based on the

[26]'John' can denote the set of sets each of which contains John. The sentence 'John and every man loves Jenny' suggests that 'John' can have the same semantic and syntactic type as the generalized quantifier 'every man'. So it isn't surprising that the meaning of 'John' can bind an individual variable.

standard of his football team a, we ask whether he is old relative to the comparison class $M(a)$. Likewise, when he is evaluated for his relative oldness based on the standard of his football team b, we ask whether he is old relative to the comparison class $M(b)$. So we guarantee a separate standard of oldness for each of his teams.

Szabo's modified account is also in a better position to explain why 'red' in Rothschild and Segal's zoo-animal sentence (15) has as its extension the set of objects each of which is canonically red for *its* kind. This shouldn't be surprising, because Rothschild and Segal's examples (16) and (17) are based on a reading of 'tall' that is just like their intended reading of 'red' in (15). This means that Szabo could explain the intended reading of 'red' in the same way that he explains the intended reading of 'tall'. To be more specific, he could explain why 'red' has Rothschild and Segal's intended reading along the following lines:

(21) Every kind of animal in the zoo x has a member that is [red $C(x)$].

'Red' can plausibly pick out the set of animals each of which is canonically red for its kind when C is saturated by the function that maps a kind of animal to its canonically red members, because each member is now assessed for its relative redness against the canonically red members of its kind.

The most crucial point here isn't that the modified account delivers precisely the reading Rothschild and Segal have in mind; it is that the semantic or type-theoretic constraints on the possible extensions of 'red' enable Szabo to predict that it can have the class of readings on which the relevant set of red objects varies with the objects ranged over by a generalized quantifier preceding the adjective, rather than any possible reading we can think of. Since Rothschild and Segal have only put minimal semantic or type-theoretic constraints on the possible extensions of 'red', it is unable to predict that 'red' has that class of readings. This lack of predictive power casts doubt on their claim that their account has at least as much theoretical strength as Szabo's account.

Perhaps the best option for Rothschild and Segal is to split the difference between their account and Szabo's: They could posit that 'red' has both an indexical element just as they suggest, and a functional comparison class argument, which predicts the class of bound readings we just discussed. The following lexical item shows how this idea can be implemented:

(22) $[\![\mathit{red}\, C(x)]\!]^{\sigma} =$ the set of objects that are members of both σ('red') and the comparison class $\sigma(C)(\sigma(x))$, where the value of σ is determined by a context.

The resulting account handles Travis cases in the same way they intend, and it has at least as much theoretical strength as their original account and Szabo's modified account. But it doesn't suggest that Rothschild and Segal can now declare victory over Szabo, because, since the comparison class argument is derived from Szabo's account, Szabo is equally entitled to this hybrid account.

3.5 How to and How Not to Respond to Travis Cases

Our plan in this section is to propose a new response to Travis cases in light of our discussion above. The upshot of our response is that the challenge to truth-conditional semantics from Travis cases evaporates once we clarify the sense in which ordinary intuitions provide good evidence for semantics.

As we discussed in §3, the idea that ordinary intuitions provide good evidence for semantic theory shouldn't be understood as the doctrine that semanticists should ensure a good agreement between the truth values they assign to sentences and ordinary truth value intuitions, because, as my toy indexical account demonstrates, a good agreement between the two doesn't by itself lend support to a semantic fragment. Ordinary intuitions provide good evidence for a given semantic fragment, I argued, only when they suggest some linguistic phenomenon that can be explained or predicted well by that fragment.

Once we understand 'good evidence' in the explanation and prediction sense rather than the agreement sense, we can now proceed to disarm Travis cases. The Travis cases based on 'Melon is red' plausibly suggest two phenomena: The first is a certain distribution of truth values over contexts (e.g. the Greengrocer, the Artist's Studio); the second is the weaker phenomenon that ordinary intuitions about its truth values tend to vary across contexts. The problem with Szabo's part variable is instructive of why both friends and enemies of truth-conditional semantics can't hope to explain or predict the first phenomenon. While Szabo has isolated the relevant part at which an object has its color as one of the factors affecting ordinary truth value intuitions, there are likely to be various other factors, such as the observation conditions, the interests and the goals of the discourse participants, and what they assume each other to know, etc. Exhausting those factors and mapping out how they interact to affect ordinary truth value intuitions seems to be a daunting, if not hopeless, task. So, as far as explaining or predicting that phenomenon is concerned, the friends and the enemies of truth-conditional semantics don't seem to have an advantage over each other.

With the second phenomenon, we have seen in our previous discussion that there isn't a shortage of explanations for it, and none of them requires giving up the assumption that the truth value of a sentence is compositionally derived from its parts' extensions. We can either adopt a pragmatic explanation such as Cappelen and Lepore's, or an indexicality-based explanation such as Rothschild and Segal's — but the latter should be suitably modified to handle the bound and the unbound readings of gradable adjectives. Both explanations have its drawbacks: Cappelen and Lepore's explanation is less parsimonious than Rothschild and Segal's because the truth value of 'Melon is red' plays no essential role in their explanation. Rothschild and Segal's explanation, if applied to every predicate that can be featured in a Travis case, multiplies indexicals in natural language greatly, and it may rely on the controversial assumption that ordinary speakers have reliable intuitions about sentences' truth values (i.e. their semantic values). But the good news for us is that, unless the enemies of truth-conditional semantics can propose a potentially better explanation for the phenomenon, there seems to be no urgency in deciding between, or making improvements on, the pragmatic explanation and the indexicality-based explanation.

Here is a further reason why we can postpone our explanation for the second phenomenon. As we saw in our discussion about the bound readings of 'red' and 'tall' (§4), the real action of semantics lies in explaining or predicting such phenomena as ambiguity and binding,[27]

[27] Here I do not assume that binding is a primitive phenomenon, because it can be understood in terms of entailment (consistency). For example, we can verify the existence of the bound reading of 'Everyone loves his mother' by asking ordinary speakers whether, on accepting that sentence to be true, they should accept 'John loves John's mother' to be true as well.

rather than the two phenomena suggested by Travis cases. Those phenomena allow us to decide between different semantics fragments (e.g. Rothschild and Segal's vs. Szabo's), but also between the program of truth-conditional semantics and its denial. For example, other things being equal, a semantics fragment (research program) that can explain and predict the bound and the binding-free readings of the following sentences is more empirically adequate than the one that can't.

(23) a. Every kind of animal in the zoo has a member that is red.

b. John is tall and Jenny is too.

c. Children who are tall are generally well nourished.

So here is my overall response to Travis cases. As we just discussed, the enemies of truth-conditional semantics have no real advantage over its friends in explaining or predicting the two phenomena suggested by Travis cases. But every good semantic explanation or prediction of linguistic phenomena puts pressure on the denial of truth-conditional semantics, because the burden is on the enemies of truth-conditional semantics to show how they can explain or predict the same phenomena without truth-conditional semantics. So we can in fact effectively defend truth-conditional semantics against the challenge from Travis cases by identifying phenomena for which we can provide good explanations or predictions — that is, by doing semantics — rather than focusing solely on the narrow range of phenomena suggested by Travis cases.

3.6 Conclusion

Much of the interest in Travis cases is fueled by the rather pervasive assumption that ordinary intuitions should by default be treated as intuitions about various theoretical constructs of semantics, such as semantic values and truth conditions. I hope our discussion contributes to resisting that assumption by emphasizing the explanatory and predictive nature of semantic theorizing. I argued that Rothschild and Segal's methodological objection to Cappelen and Lepore's account fails, because, in the absence of semantic explanation or prediction, a good agreement between semantic values and ordinary truth value intuitions doesn't justify a semantic theory. I've also argued that Rothschild and Segal's objection to Szabo's account fails, because Szabo's comparison class variable, if properly generalized, can do non-trivial explanatory and predictive work. Finally, I argued that, once we reorient our attention from ensuring agreement between semantic values and ordinary intuitions to explaining and predicting phenomena that are suggested by those intuitions, the challenge to truth-conditional semantics from Travis cases evaporates.

Bibliography

Anand, P. (2006). *De de se*. MIT Doctoral Dissertation.

Anand, P. and Nevins, A. (2004). Shifty operators in Changing Contexts. In *Proceedings of Semantics and Linguistic Theory*, volume 14.

Bach, K. (1994). Conversational Impliciture. *Mind and Language*, 9(2):124–162.

Bach, K. (2000). Quantification, Qualification and Context A Reply to Stanley and Szabó. *Mind & Language*, 15(2-3):262–283.

Bach, K. (2005). Context ex Machina. In Szabo, Z., editor, *Semantics Versus Pragmatics*, pages 15–44. Oxford University Press.

Barker, C. (2002). The Dynamics of Vagueness. *Linguistics and Philosophy*, 25(1):1–36.

Barker, C. (2013). Negotiating taste. *Inquiry*, 56(2-3):240–257.

Bezuidenhout, A. (2002). Truth–conditional Pragmatics. *Noûs*, 36(s16):105–134.

Bierwisch, M. (1989). The semantics of gradation. *Dimensional adjectives*, 71:261.

Breheny, R. (2003). A Lexical Account of Implicit (Bound) Contextual Dependence. In *Proceedings of SALT*, volume 13, pages 55–72.

Burnett, H. (2014). A Delineation Solution to the Puzzles of Absolute Adjectives. *Linguistics and philosophy*, 37(1):1.

Camp, E. (2006). Contextualism, metaphor, and what is said. *Mind & language*, 21(3):280–309.

Cappelen, H. and Lepore, E. (2004). A Tall Tale: In Defense of Semantic Minimalism and Speech Act Pluralism. *Canadian Journal of Philosophy*, 34(sup1):2–28.

Cappelen, H. and Lepore, E. (2005). *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*. Blackwell.

Carston, R. (2002). *Thoughts and Utterances*. Blackwell.

Chierchia, G. and McConnell-Ginet, S. (2000). *Meaning and grammar: An introduction to semantics*. MIT press.

Cooper, R. (1979). The Interpretation of Pronouns. *Syntax and Semantics*, 10:61–92.

Cresswell, M. (1977). The Semantics of Degree. In Partee, B., editor, *Montague Grammar*, pages 261–292. Academic Press.

Cruse, D. A. (1980). Antonyms and gradable complementaries. *Perspektiven der lexikalischen semantik: Beitraäge zum wuppertaler semantikkolloquium vom*, pages 2–3.

Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*, volume 7. Springer Science & Business Media.

Engdahl, E. (1986). *Constituent Questions: The Syntax and Semantics of Questions with Special Reference to Swedish*. Reidel, Dordrecht.

Hansen, N. (2011). Color Adjectives and Radical Contextualism. *Linguistics and Philosophy*, 34(3):201–221.

Hawthorne, J. (2007). Context-dependency and comparative adjectives. *Analysis*, 67(3):195–204.

Heim, I. (1991). Artikel und definitheit. *Semantik: ein internationales Handbuch der Zeitgenössischen forschung*, pages 487–535.

Heim, I. (2000). Degree operators and scope. In *Semantics and linguistic theory*, volume 10, pages 40–64.

Heim, I. and Kratzer, A. (1998). *Semantics in generative grammar*, volume 13. Blackwell Oxford.

Horn, L. R. (1981). A pragmatic approach to certain ambiguities. *Linguistics and Philosophy*, 4(3):321–358.

Jacobson, P. (1999). Towards a Variable-free Semantics. *Linguistics and philosophy*, 22(2):117–185.

Jacobson, P. (2000). Paycheck Pronouns, Bach-Peters Sentences, and Variable-free semantics. *Natural Language Semantics*, 8(2):77–155.

Jacobson, P. (2006). Direct compositionality and variable free semantics: Taking the surprise out of 'complex variables'. In *Handout for a keynote address delivered at 30th Penn Linguistics Colloquium*.

Jacobson, P. (2014). *Compositional Semantics: An Introduction to the Syntax/semantics Interface*. Oxford University Press.

Kaplan, D. (1989). Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In *Themes from Kaplan*, pages 481–563. Oxford University Press.

Karttunen, L. (1969). Pronouns and variables. In *Fifth Regional Meeting of the Chicago Linguistics Society*, pages 108–115.

Kennedy, C. (1999). *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Routledge.

Kennedy, C. (2007). Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives. *Linguistics and Philosophy*, 30(1):1–45.

Kennedy, C. and McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, pages 345–381.

Kennedy, C. and McNally, L. (2010). Color, Context, and Compositionality. *Synthese*, 174(1):79–98.

King, J. and Stanley, J. (2007). Semantics, Pragmatics, and the Role of Semantic Content. In *Language in Context*, pages 133–181. Oxford University Press.

Klein, E. (1980). A Semantics for Positive and Comparative Adjectives. *Linguistics and Philosophy*, 4(1):1–45.

Krifka, M. (2007). *Approximate interpretation of number words*, pages 111–126. Humboldt-Universität zu Berlin, Philosophische Fakultät II.

Larson, R. and Segal, G. (1995). *Knowledge of Meaning*. Cambridge, MA: MIT Press.

Lasersohn, P. (1999). Pragmatic Halos. *Language*, pages 522–551.

Lasersohn, P. (2005). Context Dependence, Disagreement, and Predicates of Personal Taste. *Linguistics and philosophy*, 28(6):643–686.

Lasersohn, P. (2016). *Subjectivity and perspective in truth-theoretic semantics*, volume 8. Oxford University Press.

Lassiter, D. (2009). Vagueness as probabilistic linguistic knowledge. In *International workshop on vagueness in communication*, pages 127–150.

Lewis, D. (1979a). Attitudes De Dicto and De Se. *The Philosophical Review*, 88(4):513–543.

Lewis, D. (1979b). Scorekeeping in a language game. *Journal of philosophical logic*, 8(1):339–359.

Ludlow, P. (2011). *The Philosophy of Generative Linguistics*. Oxford University Press.

Ludlow, P. (2014). *Living words: Meaning underdetermination and the dynamic lexicon*. OUP Oxford.

McNally, L. (2009). The Relative Role of Property Type and Scale Structure in Explaining the Behavior of Gradable Adjectives.

Morzycki, M. (2011). Metalinguistic Comparison in an Alternative Semantics for Imprecision. *Natural Language Semantics*, 19(1):39–86.

Ninan, D. (2010). De Se Attitudes: Ascription and Communication. *Philosophy Compass*, 5(7):551–567.

Nunberg, G. (1993). Indexicality and Deixis. *Linguistics and Philosophy*, 16(1):1–43.

Partee, B. (1989). Binding Implicit Variables in Quantified Contexts. In *Proceedings of the Chicago Linguistics Society*, volume 25, pages 342–365.

Pinkal, M. (1995). *Logic and lexicon: the semantics of the indefinite*, volume 56. Springer.

Postal, P. M. (1974). On certain ambiguities. *Linguistic Inquiry*, 5(3):367–424.

Predelli, S. (1998). I am not here now. *Analysis*, 58(2):107–115.

Predelli, S. (2011). I am still not here now. *Erkenntnis*, 74(3):289–303.

Recanati, F. (2004). Indexicality and context-shift. In *Workshop on indexicals, speech acts and logophors*.

Rooth, M. (1985). *Association with Focus*. GLSA, Dept. of Linguistics, University of Massachusetts, Amherst.

Rothschild, D. and Segal, G. (2009). Indexical Predicates. *Mind & language*, 24(4):467–493.

Rotstein, C. and Winter, Y. (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural language semantics*, 12(3):259–288.

Russell, B. (1905). On Denoting. *Mind*, 14(56):479–493.

Sauerland, U. and Stateva, P. (2011). Two types of vagueness. In *Vagueness and language use*, pages 121–145. Springer.

Schlenker, P. (2003). A Plea for Monsters. *Linguistics and Philosophy*, 26(1):29–120.

Schlenker, P. (2005). Non-redundancy: Towards a semantic reinterpretation of binding theory. *Natural Language Semantics*, 13(1):1–92.

Sperber, D. and Wilson, D. (2005). Pragmatics. In Jackson, F. and Smith, M., editors, *The Oxford Handbook of Contemporary Philosophy*. Oxford University Press.

Stalnaker, R. (1978). Assertion. Reprinted in *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford University Press.

Stanley, J. (2000). Context and logical form. *Linguistics and philosophy*, 23(4):391–434.

Stanley, J. (2007a). *Language in context*. Clarendon Press.

Stanley, J. (2007b). Semantics in Context. In *Language in Context*, pages 201–230. Oxford: Clarendon Press.

Stanley, J. and Szabó, Z. (2000a). On Quantifier Domain Restriction. *Mind & Language*, 15(23):219–261.

Stanley, J. and Szabó, Z. (2000b). Reply to Bach and Neale. *Mind & language*, 15(2-3):295–298.

Stephenson, T. (2007). Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy*, 30(4):487–525.

Syrett, K., Bradley, E., Kennedy, C., and Lidz, J. (2006). Shifting standards: Children's understanding of gradable adjectives.

Syrett, K., Kennedy, C., and Lidz, J. (2009). Meaning and context in children's understanding of gradable adjectives. *Journal of semantics*, 27(1):1–35.

Szabó, Z. (2001). Adjectives in Context. In Kenesei, I. and Harnish, R. M., editors, *Perspectives on Semantics, Pragmatics, and Discourse*, pages 119–146. Amsterdam/ Phildelphia: John Benjamins Publishing Company.

Szabó, Z. (2006). Sensitivity Training. *Mind & Language*, 21(1):31–38.

Toledo, A. and Sassoon, G. W. (2011). Absolute vs. Relative Adjectives-variance within vs. between Individuals. In *Semantics and Linguistic theory*, volume 21, pages 135–154.

Travis, C. (1985). On what is strictly speaking true. *Canadian Journal of Philosophy*, 15(2):187–229.

Travis, C. (1997). Pragmatics. In Wright, C. and Hale, B., editors, *A Companion to the Philosophy of Language*. Blackwell.

Unger, P. (1978). *Ignorance: A Case for Scepticism*. OUP USA.

Van Benthem, J. (1982). Later than late: On the logical origin of the temporal order. *Pacific Philosophical Quarterly*, 63(2):193–203.

Van Rooij, R. (2011). Implicit versus Explicit Comparatives. In *Vagueness and Language Use*, pages 51–72. Springer.

Von Fintel, K. (1994). *Restrictions on Quantifier Domains*. PhD thesis, University of Massachusetts.

Von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of semantics*, 3(1-2):1–77.