Subjectivity and Perspective in Truth-Theoretic Semantics

Peter Lasersohn
University of Illinois

Version: 5/14/2015

Preface

This book has been a long time in the making. The fundamental question it addresses — how to analyze sentences which are about matters of opinion rather than matters of fact — is one that occurred to me when I first encountered the idea of truth-conditional semantics as a graduate student, back in the early 1980's. I am sure it has occurred to many other beginning semantics students as well. But modern semantic theory is so complex and technical that we must spend a good deal of time and effort learning to "think in the formalism" before we can really understand it; and by the time we have acquired the necessary habits of thought, questions like this no longer occur so naturally to us. We forget the things that looked strange to us at the beginning, or write them off as due to our early misunderstandings. So it was with me, at any rate. It takes another long time to learn *not* to think in the formalism again, and by that point, most researchers are so specialized that novel questions outside their own narrow area of expertise still do not occur very naturally.

Luckily, we have beginning students to remind us of sticky points in the foundations of our field. Students trained in the tradition of Generative Linguistics often take for granted that everything mentioned in a linguistic theory is supposed to be some sort of mental representation. This can lead to a systematic misunderstanding of semantic theories which use concepts of truth and denotation. So, like many other semantics instructors, I often emphasized near the beginning of introductory courses that the denotation of an expression is normally the thing "out in the world" it corresponds to, not a representation of that thing in the mind of a language user; and that whether or not a sentence is *true* normally depends on how things really, objectively are in the external world, not just on language users' mental representations of the world. Occasionally, students responded to these points with the same kind of concern I had had as a student: How then to deal with sentences whose truth or falsity seems to be a matter of opinion rather than a matter of fact? Sometimes the examples they had in mind in raising this sort of question were aesthetic claims, like *That painting is beautiful*; sometimes they were moral claims, like *Abortion is wrong*; and sometimes they were what I came to call "personal taste" claims, like *Roller coasters are fun*.

As I later recounted in the opening paragraphs of Lasersohn (2005), I developed a set of routine responses to such questions, designed to resolve the issue quickly and allow the class to proceed to the topics it was really about: quantification, tense, anaphora, intensionality, etc. I usually sidestepped the aesthetic and moral examples on the grounds that it was not completely clear whether these really were matters of opinion rather than fact. The personal taste examples were harder to dismiss in this way, so I suggested that they could be analyzed as true relative to one person and false relative to another, in much the same way as a sentence like *My name is Peter Lasersohn* is true if I say it, but false if someone else does.

Sometime in the early 1990's I sat in on a seminar on demonstratives offered by David Braun at the University of Rochester. I had previously read Kaplan (1989) and related work such as Montague (1970), but had never really thought about indexicality in much detail. As I began to devote more attention to the topic, I became increasingly dissatisfied with my quick answer to the question of how to deal with personal taste examples. If they really involved first-person indexicality on the model of *My name is Peter Lasersohn*, then a speaker who said "Roller coasters are fun" would not be disagreeing with a speaker who said "Roller coasters are not fun," any more than I would be disagreeing with my friend Bill, if I said "My name is Peter Lasersohn" and Bill

said "My name is not Peter Lasersohn." But intuitively, the speakers of the two sentences about roller coasters *do* disagree with each other, while Bill and I do not.

At first I took this as an argument that taste (and aesthetic and moral) sentences really did make objective claims, whose truth was a matter of fact — not opinion — despite our initial intuitions to the contrary. With no first-person element to their meanings, *fun* (and *beautiful* and *wrong*) would presumably express ordinary properties, the same for everyone, which hold of their arguments absolutely, objectively and factually. But this idea never sat quite right with me, especially as it concerned the taste examples. At that point, however, I could see no way to resolve the matter, and set it aside to work on other topics.

Some years later, conversations with linguists who advocated various forms of "Cognitive Semantics" prompted me to think again about how a truth-theoretic semantic theory could deal with sentences which were about matters of opinion. Not that these linguists raised that particular issue to me — but their apparent reluctance to acknowledge any role for real-world truth and denotation in semantic theory led me to reconsider what, from my own perspective, seemed like problematic examples for truth-theoretic semantics; and sentences about personal taste and other matters of opinion were high on that list. It seemed to me that there must be some way to treat such sentences in a truth-and-denotation-based semantic theory, without portraying them as though they made simple claims of objective fact.

About 2001 I began to look in earnest at the range of options available for the analysis of taste sentences. There were serious technical or empirical problems with every method I could think of, except the one which seemed philosophically most bizarre: claiming that taste sentences expressed the same content for everyone, but nonetheless varied in truth value from person to person. I wrote up a paper outlining the problems and proposing this solution, and gave oral presentations of it at the University of Illinois and the University of Chicago in 2001 and 2002, respectively.

Family matters required me to stop working on the paper for about two years. In retrospect, this was probably a good thing; on resuming work in 2004 I was able to make a number of improvements that had not occurred to me two years earlier. Most important of these was a formalization which allowed easy direct comparison to Kaplan (1989) — a feature which I think contributed significantly to the later success of the paper.

About this time I also became aware of recent philosophical work on related topics. I was surprised to discover that Kölbel (2002) had just published an entire book arguing for very similar conclusions to those of my paper. At first I worried that this meant that my paper could not be seen as making an original contribution, but eventually I decided that my arguments, emphasis, and style of presentation were distinct enough from Kölbel's to have their own merits. It was comforting, too, to realize that I was not alone in pursuing what historically had been an extremely unpopular philosophical position.

I posted my paper publicly in mid-2004 and was contacted almost immediately by John MacFarlane. His (2003) paper had argued that the truth values of future contingent sentences should be relativized to contexts of assessment (not just contexts of use) and since then had written several papers — at that point still unpublished — exploring the general idea of assessment-sensitivity, and applying it to various other sorts of examples, including taste sentences. John directed me to that work and also sent some very insightful comments on my paper.

Circumstances again forced me to stop work for several months. When I returned to the paper, I realized that although I had been arguing that taste predicates should not be analyzed on the model of first-person pronouns, my analysis still tied them to the speaker in a problematic way. I revised the analysis to fix that problem by distinguishing between "formal" contexts and concrete situations of use. I communicated this idea to MacFarlane in a response to his earlier comments, and in the exchange which followed it became clear that my position had more-or-less converged with his. I was still not using the terminology of "context of assessment" or the associated formalism, but the concept was unmistakably implied.

I submitted the paper to *Linguistics and Philosophy*, and it was published in late 2005 as 'Context Dependence, Disagreement, and Predicates of Personal Taste'.

I soon found that I had caught a much bigger wave than I had anticipated. Aside from Kölbel's book and MacFarlane's papers (which were now appearing in quick succession), similar analyses were being suggested by Richard (2004) for scalar predicates and by Egan, Hawthorne, and Weatherson (2005) for epistemic modals. Many more papers followed in response. Improbably, relativism about truth was becoming a popular position in analytic philosophy.

I followed up my (2005) paper with a series of shorter articles developing the analysis further, and responding to counterarguments (Lasersohn (2007), (2008), (2009b), (2011), (2013)). But I felt that the research program I was now engaged in would benefit from sustained presentation in a single document, and began work on this book in 2008.

As I originally conceived it, the book was to include a more detailed exploration and justification of the kind of analysis presented in Lasersohn (2005) and my later articles; reviews of related literature, including all the major subsequent developments and counterarguments by others; and my own responses to those counterarguments. As time went on and the body of related literature continued to expand, it became clear that this was not a realistic goal, and that any serious attempt at it would only result in a massive, unreadable catalog of responses and counter-responses. With some reluctance but also a sense of relief, I decided instead to devote my attention to those features of the analysis which were most interesting to me, without attempting to respond to every alternative proposal.

The primary goal of the book is to demonstrate that a truth-theoretic semantic theory is capable of accounting for the meanings of sentences which intuitively concern matters of opinion rather than fact — and can do so with only relatively minor adjustments to the standard formal apparatus of such theories, and without discounting the intuition of subjectivity such sentences elicit. A secondary goal is to argue for a specific formal technique to be used in the analysis of such sentences, specifically a model of semantic contents as determining extensions only relative to parameters whose values are left indeterminate in the context of utterance. Because this technique requires a definition of truth as relative to non-standard parameters, and because such a relativization of truth has until recently been an unpopular position in analytic philosophy, another major goal of the book is to motivate and defend such a relativized conception of truth and demonstrate its coherency and plausibility. Finally, another important goal is to draw out the implications of this technique for the analysis of specific linguistic phenomena, including control of null-subject complement clauses, presupposition, and propositional attitude predicates. Because much of the existing literature on relativist semantics is more philosophical than linguistic in focus, many of these implications have been relatively unexplored.

In my (2005) paper, I couched the analysis in a Kaplan-style semantics, and illustrated it

using a simple logical language which was based directly on Kaplan's language LD. My thinking was that the analysis would be clearer and easier to follow if it were integrated into a familiar formal system, and that presenting the analysis in the context of a standard theory of indexicality like Kaplan's would bring out more clearly the differences between the kinds of context-sensitivity exhibited by predicates of personal taste and the kinds exhibited by ordinary indexicals. This style of presentation served its purpose very well, but also had the perhaps unfortunate effect of keeping the analysis a step removed from the details of natural language grammar. As a linguist, I think these details deserve careful exploration; so with the comparison to Kaplan already accomplished in my earlier paper, I have taken the opportunity in this book to give syntactic and semantic rules for a fragment of English rather than a simple language like LD. This grammar is still somewhat simplified, but at least can be construed as a simplified model of *English* grammar rather than as a wholly artificial logical language. The semantics also differs from Kaplan's in numerous ways not connected specifically with the analysis of personal taste sentences, both in its general architecture and in its treatment of specific words and constructions. I have tried to explain the reasons for setting up the semantic system in the way I do, but not having taken Kaplan (1989) as my stepping-off point (as I did in Lasersohn (2005)), I have not treated every departure from Kaplan's approach (or from my own earlier presentation) as requiring specific justification. It is my hope that the parts of the book which deal with tense, locatives, infinitives, PRO, etc., will be of interest to linguists and philosophers working on these topics, even independently of their implications for personal taste examples and other sentences about matters of opinion.

Portions of this book — in varying states of completion and "draftiness" — were read and discussed by the Philosophy and Linguistics Reading Group at the University of Illinois at Urbana-Champaign, and by the Subjectivity in Language and Thought research group at the University of Chicago. Thanks to both groups for helpful comments and advice.

Thanks also to Sharon Haworth and Nathan Lasersohn for love, patience and support during the long process of writing this book, and to my colleagues and students at the University of Illinois for a congenial work environment. This book is dedicated with love and gratitude to my father, William Bock Lasersohn, M.D.

Table of Contents

Prefacei
Chapter 1 : Subjectivity, Disagreement and Content 1
1.1. The goal: a truth-theoretic semantics for sentences expressing subjective judgment . 1
1.2. Matters of fact and matters of opinion 3
1.3. Subjectivity as relative truth 5
1.4. Context, content and denotation 7
1.5. Homomorphic interpretation and differentiation of content 12
Chapter 2 : Dismissing the Easy Alternatives 16
2.1. Indexical analyses 16
2.2. Quantificational analyses 22
2.3. Absolutism and ignorance 23
2.4. Expressivism 25
2.5. Metalinguistic and metacontextual conflict 28
Chapter 3 : Ordinary Indexicality 31
3.1. Syntactic assumptions 31
3.2. Pronouns, names and anaphora 34
3.3. Common nouns, quantification and binding 37
3.4. Tense 46
3.5. Spatial deixis 53
3.6. Negation and contradiction 60
3.7. Intensionality 63
Chapter 4 : Basic Relativist Semantics 67
4.1. The judge parameter 67
4.2. Revising the grammar 70
4.3. Extension to taste expressions 73
4.4. Truth in a relativist semantics 75
4.5. Derelativization in the object language 77
Chapter 5 : “Hidden” and “Disguised” Elements 81
5.1. Phonological reduction 81
5.2. Null elements in syntax 83
5.3. Unarticulated constituents 90
5.4. Constructional indexicality 91
5.5. Sublexical and compound indexicality 92
5.6. Indexical and quantificational interpretations of hidden elements 97
5.7. Conclusion 99
Chapter 6 : Pragmatics of Truth Assessment 101
6.1. Contexts of use and parameter values 101
6.2. Contexts of assessment and parameter values 104
6.3. Adopting a stance 106

6.4. Truth assessment and the adicity of *true* 108
6.5. Exocentricity and indexicality 109
6.6. Acentric stances 113
Chapter 7 : Attitude Predicates in Relativist Semantics 114
7.1. Stance and belief 114
7.2. Factives, relativism and speaker commitment 117
7.3. Infinitival clauses and time reference 119
7.4. Centered attitudes and control of infinitival subjects 123
7.5. Effective centering without PRO 134
7.6. *De se* belief and autocentric truth assessment 141
7.7. Against non-indexical contextualism 142
Chapter 8 : Assertion and Other Speech Acts 150
8.1. Assertion, Norms, and Portrayal as True 150
8.2. Assertion, conversation, and context change 153
8.3. Why assert? 155
8.4. Questions 157
Chapter 9 : Between Fact and Opinion 159
9.1. Aesthetic judgment and refinement of taste 159
9.2. Contingent futures 162
9.3. Epistemic modality 166
9.4. Non-taste candidates for relativism: scalar cut-offs, sufficiency, and derogation 168
9.5. Conclusion 173
Chapter 10 : Reliability, Imagination, and the Functional Motivation for Relativism 174
10.1. An evolutionary fable 174
10.2. Space and logical space 179
10.3. Reliability and traversable dimensions 181
10.4. Reliability-based content 185
10.5. From reliability to truth 187
References 191

Chapter 1 : Subjectivity, Disagreement and Content

1.1. The goal: a truth-theoretic semantics for sentences expressing subjective judgment

For some time now, the dominant approach to semantic theory has been truth-theoretic: we explain the meanings of sentences in terms of their truth conditions. Depending on our views, we may do this by directly identifying meaning with truth conditions, or by modeling meaning as truth conditions, or perhaps simply by explaining meaning in such a way that the truth conditions of sentences are systematically derivable from their contents; but in any case, the notion of *truth* plays a central role in our explanation of meaning. This has been an extremely worthwhile and productive strategy, and while it has had (and no doubt will continue to have) its fair share of critics, it is hard to deny the advances it has brought to our understanding, both of particular grammatical constructions and of the workings of language in general.

Despite its proven value, truth-theoretic semantics has always seemed an odd fit for certain kinds of sentences. Intuitively, a sentence is true if it conforms to the facts, accurately representing things as they really, objectively are. But some sentences don't seem to be about matters of fact — at least not entirely. A sentence like *This licorice is tasty* has the form of an assertion, but intuitively seems more to express a matter of personal taste or opinion than to make an objective claim. What are we to say about its truth conditions? Does it make sense to think of such a sentence as even *having* truth conditions, or to try to explain its meaning in terms of truth and falsity?

Sentences expressing personal taste are far from the only ones for which this issue arises. Just how rich does John have to be to make the sentence *John is rich* true? Most of us, I think, share the intuition that there is no "fact of the matter" that resolves this question; and while there are well-developed arguments to the contrary,[1] it is precisely because such arguments run sharply against our intuitions that we find them exciting. Similarly, many people have the intuition that there is no fact of the matter whether a future contingent sentence like *There will be a sea battle tomorrow* is true; if the future is authentically open, the facts have not yet been determined. Depending on our views we might also claim that there are no objective moral facts, or even that all evaluative sentences — whether moral, aesthetic, or some other category — concern non-factual matters. Derogatory epithets seem to be descriptive, but the applicability of the description often seems to be a matter of opinion rather than fact. Certain kinds of modal sentence express our epistemic states, perhaps in the same way as sentences like *This licorice is tasty* express our personal tastes; if so, will objective facts always suffice to determine a truth value for *John may be in the next room*?

Each of these questions involves its own, complex set of issues, and it may well be that they should not all be answered in the same way. Nor, indeed, is it clear that they should all be regarded primarily as semantic questions, since some of them carry significant implications for the theories of ethics, aesthetics, and other areas. But the way we answer these questions is bound to affect any semantic theory which includes a truth definition; and particularly if we claim that some examples are only "subjectively" true or false, the implications for the general form of a semantic theory may be substantial.

[1] See, e.g., Williamson (1994).

In this book, I will develop and defend a semantic theory which respects the intuition that there may be no "fact of the matter" which determines truth values for certain kinds of sentences. The theory nonetheless takes truth and falsity as central to explaining meaning, and uses familiar techniques from modern (formal, truth-conditional, logical) semantic theory. It will account for non-factual meaning by relativizing the truth values of sentence contents to parameters whose values are not always objectively or factually determinable. The resulting theory may reasonably be called "relativist," and much of the book will be devoted to presenting and defending a kind of relativism about truth.

Relativism about truth has until recently been such an unpopular position in analytic philosophy that it may be difficult to see the presentation of a relativist semantic theory as anything other than a defense of relativism. Yet when I began work in this area, my purpose was not so much to promote relativism as to defend logically-oriented, truth-theoretic semantics against the claim that it could provide no natural account of sentences which express subjective judgment, as opposed to making objective claims. If truth-theoretic semantics really has nothing to say about such examples, they might be seen as motivation for displacing the concept of truth from its central explanatory role in semantic theory, and lend support to the kinds of theories which treat truth conditions as a matter of only peripheral — or even illegitimate — interest to semantics.

I believe that the theory of truth can and should be central to semantics, and that the standard techniques of modern formal semantics can be adapted to deal with examples which are only subjectively true or false. This adaptation will require some revision, of course; but from a formal standpoint, the revisions will be fairly minor. Even from a conceptual standpoint, I think the revisions are not as radical as some people might expect. Such examples do not require a thoroughgoing reconstruction of the foundations of semantics on non-truth-theoretic grounds; nor does a defense of truth-theoretic semantics require that we deny the existence of sentences whose contents are true or false only subjectively, treating them instead as though they made purely factual claims.

Our focus throughout the book will be on semantic issues. If there are deeper philosophical connections to be drawn, we will studiously avoid drawing them, in order to keep our attention on semantics. This will be easier if our primary examples are of relatively little philosophical interest *aside* from the evidence they provide for subjective truth. For that reason, our initial focus will not be on words like *good* or *beautiful* or *know*, whose analysis inevitably carries implications for central areas of philosophical inquiry, but on rather more mundane examples such as *fun* and *tasty*.[2] After developing a reasonably detailed analysis of such examples, we will have the opportunity to consider its extensibility to a wider range of cases — still limiting attention as much as possible to semantic issues.

[2] Of course this is not to say that such examples carry no independent philosophical interest at all; but they do not seem to present the same temptation for distraction from semantic issues.

1.2. Matters of fact and matters of opinion

What does it mean to say that a sentence is not concerned exclusively with "matters of fact," but is instead concerned at least partly with "matters of opinion"? Much of this book will be concerned with developing a precise answer to this question, but it will be useful at this early stage to say at least a little bit to help delineate our topic. The point here is not to provide a definition, much less a "test" by which matters of opinion and matters of fact can be distinguished, but simply to help readers build up an informal, pretheoretic sense of what kinds of sentences we will be dealing with.

To a large extent I think we all share intuitions about what a matter of opinion is, and share the intuition that matters of opinion are different in some way from matters of fact. If someone asks whether beets are tasty, or bowling is fun, or whether a particular joke is funny, or a class was interesting, we intuitively recognize these questions as concerning matters of opinion, and as different in some important way from questions like whether beets are cruciferous, or whether the joke made John laugh, or whether it is now three o'clock, which are concerned solely with matters of fact.

Of course there may be examples about which our intuitions are unclear, or vary from person to person; and certainly as theorists we may differ radically in our claims about what distinguishes the two classes of examples. But it is a significant fact that we do "come equipped," so to speak, with intuitions which pretheoretically identify certain kinds of sentences as concerning matters of opinion and others not. This immediately closes to us the option of constructing an analysis which does not draw the distinction in any way. There must be *some* kind of difference between the two classes of sentences, or they would not provoke different intuitions. This is not to claim that our intuitions are always reliable, or that the mere intuition that some sentences are only subjectively true or false is enough to establish that they are. We all know of areas of inquiry where initial intuitions turn out on inspection to be wrong or even incoherent. But in the study of language, intuitions are data, and a theory which is incapable of explaining how differences in intuition might arise has clearly missed something.

Let us assume, therefore, that some sort of real distinction underlies the intuitive difference between matters of fact and matters of opinion. In what sorts of cases does the intuition that a sentence concerns a matter of opinion arise?

First, we should note we may have (and express) opinions about all kinds of things which we would not ordinarily describe as "matters of opinion." For example, if my friends ask me whether seahorses are vertebrates, I may offer my opinion on the matter; but this does not mean that it is a matter of opinion whether seahorses are vertebrates; this is a matter of fact. The phrase *matter of opinion* implies, as the simple word *opinion* does not, that opinions are somehow crucial to settling the issue — that facts alone will not suffice to determine an answer.

That having been said, it is noteworthy that matters of opinion are ordinarily encoded in our language in a manner which is virtually identical to the manner in which matters of fact are encoded. Both are prototypically expressed in ordinary declarative indicative sentences. There is no major difference in syntactic structure between sentences like *That roller coaster is fun* and sentences like *That roller coaster is 400 feet tall*, as there is between these sentences on the one hand and interrogative or imperative sentences like *Is that roller coaster 400 feet tall?* and *Be fun!* on the other.

Likewise, clauses expressing matters of opinion have virtually the same external syntax as clauses expressing matters of fact. They appear as complements to the same predicates: *John*

believes/knows/hopes/says/claims/denies that the roller coaster is fun is no less natural than *John believes/knows/hopes/says/claims/denies that the roller coaster is 400 feet tall*. They are joined by the same connectives: *The roller coaster is fun and/or/if/because the park was designed by John*. They may be modified by the same adverbs: *Usually/hopefully/surprisingly, roller coasters are fun/400 feet tall*.

The fact that sentences expressing matters of opinion appear as complement clauses to the same predicates as sentences expressing matters of fact gives us at least *prima facie* evidence that we may stand in the same mental attitudes to the contents of both kinds of sentence. Unless we are prepared to claim that large parts of our everyday discourse are literally false, we must count the contents of sentences like *Roller coasters are fun* among the things that can be believed, known, hoped, etc. The same sort of consideration suggests that the contents of sentences expressing matters of opinion, like the contents of sentences expressing matters of fact, are among the things one can say, claim, deny, or use similarly in other speech acts.

Sentences expressing matters of opinion also seem to bear the same systematic syntactic and semantic relations to corresponding sentences in other moods or sentence types as do sentences expressing matters of fact. Interrogative and subjunctive clauses such as *whether the roller coaster is fun* and *that the roller coaster be fun* seem to relate to the indicative clause *that the roller coaster is fun* in precisely the same way — both syntactically and semantically — as *whether the roller coaster is 400 feet tall* and *that the roller coaster be 400 feet tall* do to *that the roller coaster is 400 feet tall*. Again, this is suggestive of some basic similarity between clauses expressing matters of fact and clauses expressing matters of opinion; otherwise, how will we be able to analyze mood and sentence type in such a way that they apply uniformly to both kinds of clause?

Finally, it should be noted that sentences expressing matters of opinion appear no less naturally as premises or conclusions of logical arguments than do sentences expressing matters of fact. The conclusion in (1)a. seems to follow from the premises in precisely the same way as the conclusion in (1)b.:

(1) a. If the ride was fun, John will go on it again.
The ride was fun.
Therefore, John will go on it again.

b. If the ride cost less than $5, John will go on it again.
The ride cost less than $5.
Therefore, John will go on it again.

It appears that in their syntactic structure, their contribution to compositional semantics, their speech act function, and their role in reasoning, sentences expressing matters of opinion are fundamentally similar to sentences expressing matters of fact. One may well wonder whether there is really any need to draw a distinction between the two at all.

There is at least one major apparent difference, however: sentences expressing matters of opinion present the possibility of what (since Kölbel (2004)) has come to be known as *faultless disagreement*.[3]

[3] I am not completely comfortable with this term, for several reasons; see especially fn. 17 and p. 147 below.

Suppose John and Mary are riding the roller coaster together. Mary — a real thrill-seeker — is enjoying herself thoroughly, but all the sudden dips and turns are making John feel dizzy and nauseated. The following short conversational exchange takes place:

(2) *Mary:* This is fun!
John: No, it isn't!

Here, John seems to be referring to the same thing in his use of the pronoun *it* as Mary does in her use of *this*; so there is something which Mary claims is fun and John claims is not fun. We therefore regard them as disagreeing with one another. Yet neither one of them seems to be making an error of fact. We may regard each of them as entitled to his or her own opinion — as fully justified in adopting and asserting that opinion — even though this places them in direct contradiction to one another.

In contrast, in disagreement over factual matters, such as whether or not it costs five dollars to ride the roller coaster, one of the two parties in the disagreement must be right and the other one wrong:

(3) *Mary:* This costs $5.
John: No, it doesn't.

John and Mary may both reach their conclusions based on apparently sound observations, testimony and/or reasoning, but if they contradict each other as in (3), one of them must be making a mistake — one of their assertions is "faulty."

As a rough initial characterization, then, let us say that *a sentence expresses a matter of opinion if it is declarative in syntactic form, but gives rise to faultless disagreement when contradicted*. No doubt this characterization will need some fine-tuning, but it at least gives us a rough-and-ready way of identifying our topic. As we continue, we will obviously need to provide some explanation of how it is possible to disagree faultlessly, and of what distinguishes faultless from non-faultless disagreement. It will also be necessary to show how this explanation can be integrated with an explicit semantic theory, explaining how the meanings of sentences expressing matters of opinion can be derived, and how they contribute to the interpretations of larger sentences of which they form a part.

1.3. Subjectivity as relative truth

A wide variety of techniques present themselves as options for the analysis of sentences expressing matters of opinion, and for drawing the distinction between such sentences and those which express matters of fact. We shall have to explore several of these in some detail in order to justify a particular choice. It may be useful, in reviewing our options, to have a clear sense in advance of the particular position this book will be arguing for, and of some preliminary reasons for thinking that it may be a productive line to pursue.

However, it does provide a convenient and intuitive label for the kinds of examples which form the primary motivation for a relativist semantics and is too well established to try to replace.

The semantic theory to be presented here will define a notion of semantic *content*, and present principles for the assignment of contents to what I will call *uses* of linguistic expressions. Content is the level at which logical notions such as contradiction and entailment are fundamentally defined. Content is distinct from denotation, but relates to it systematically. In particular, each content will have various potentially distinct denotations, depending on the values of certain parameters or indices, including a possible world index and one or more additional indices encoding variation in personal tastes, perspective, epistemic state, or other factors which produce variation in subjective judgment concerning matters of opinion.

One particular world will be privileged as actual. In cases where a content deals with purely factual, objective matters, there will be no variation in denotation according to non-world indices. For such contents we may define a monadic, unrelativized, non-parametric denotation assignment. Truth values will directly recoverable from denotations, so the truth values of contents dealing purely with matters of fact will be assigned absolutely. In this case we may speak of the sentence as expressing a content which is true or false *simpliciter*, not just relative to particular values for the indices.

But in cases where an expression deals with matters of opinion, the relativization of the denotation assignment will be ineliminable, so that no absolute truth values can be assigned, only relative, parameter-dependent, truth values. Because possible worlds are "total," resolving all matters of fact, ineliminable variation in truth value according to non-world parameters amounts to dependency of truth value on non-factual matters. In this case the content expressed by a sentence is neither true nor false *tout court*, but only relative to particular values for these non-world indices — that is, only relative to a particular way of resolving parameters which the facts of the world leave unresolved.

The result is a system containing several distinct notions of truth, including a simple, monadic concept of truth *simpliciter*, which applies only to the contents of sentences dealing purely with matters of fact, and a parameterized notion of truth-relative-to-indices which applies both to the contents of sentences dealing with matters of fact, and to the contents of sentences dealing with matters of opinion.

I claim that this second, relativized notion of truth is needed in order to capture the many similarities which sentences expressing matters of opinion share with sentences expressing matters of fact; that it is sufficient to this purpose; and that the first, unrelativized notion of truth must not be applied to such sentences, in order to capture the differences between the two types of sentence, especially with regard to faultless disagreement. The theory to be presented here thus contrasts both with theories which claim that such sentences are so radically different from statements of fact that no notion of truth and falsity can apply to them, and with theories which deny the difference altogether, claiming instead that even sentences like *This licorice is tasty* and *The roller coaster is fun* describe matters of objective fact, the same for everyone.

Why adopt such a theory? It will take the rest of the book to argue for this in detail, but the initial motivation for pursuing the idea may be summarized like this: We need to employ some sort of truth assignment to the contents of sentences expressing matters of opinion in order to account for their role in logical argumentation, their service as objects of certain mental attitudes, their use in assertoric speech acts, and their participation with other declarative sentences in syntactic and morphological alternations such as mood and interrogative marking — in other words, in order to account for the many ways in which sentences expressing matters of opinion function linguistically as though they had truth values. Yet we cannot apply the same simple, monadic,

"absolute" truth predicate to them as we apply to sentences expressing matters of fact, because then we would not be able to maintain the distinction between matters of fact and matters of opinion, and particularly would not be able to maintain the "faultlessness" of faultless disagreement. Employing two distinct but systematically related notions of truth will allow us to capture both the similarities and the differences, without overstating or understating either.

1.4. Context, content and denotation

Some of the arguments to be given in what follows will depend on considerations having to do with the nature of semantic content, how it relates to truth and denotation, and how contents are individuated. It is therefore worth clarifying some basic assumptions about these matters.

I assume, with Kaplan (1989) and many others, that semantic content can vary with pragmatic context. Most obviously, the content of a sentence containing a personal pronoun or other indexical expression will vary depending on its reference, which is determined in context. For example, the sentence *I am drinking coffee* expresses a different content in a context where I am the speaker than it does in a context where someone else is the speaker; whoever the speaker may be, the sentence expresses the content that that person is drinking coffee (at the time of utterance) — different contents for different speakers.

Of course, this does not in any way conflict with the claim that this sentence has a meaning which is the same no matter who the speaker is; we simply need to distinguish more than one notion of "meaning." Again following Kaplan, we may use the term *character* for meaning in the sense that the meaning of *I am drinking coffee* does not vary with pragmatic context, as distinct from content, which may.

Why do we need the level of content in addition to character? As Kaplan points out, appealing to a separate notion of content greatly simplifies the characterization of the distinction between necessary and contingent truth. Intuitively, it is a contingent fact that I am located where I am right now; I could have been somewhere else. Yet in any circumstance where someone asserts *I am here now*, the assertion must be true.[4] We cannot say that a sentence expresses a necessary truth if it is true no matter when it is uttered, or we will get the wrong results here. But if we say that when this sentence is used by Peter Lasersohn in Urbana, Illinois at 11:43 a.m., May 10, 2011, it expresses a content which is true if and only if Peter Lasersohn is in Urbana, Illinois at 11:43 a.m., May 10, 2011, then we can easily account for the intuition that this is a contingent truth, for I could have chosen to go to Chicago on that date.

More important for our current purposes than the characterization of necessary truth is the definition of *contradiction*. The level of content, not character, is the natural locus for identifying when two claims contradict each other. If I say *I am drinking coffee*, and you say (at the same time) *I am not drinking coffee*, you are not contradicting me, even though you utter a sentence which is the syntactically negated version of my sentence. It is easy to see why there is no contradiction here: your utterance expresses a content which is true iff it is false that *you* are drinking coffee; the content you are semantically negating is distinct from the content I am asserting. On the other

[4] Setting aside examples where the speaker points at a map as he or she says *here*, or other similar cases involving reference to individuals, times and locations other than the speaker and the time and location of utterance. Of course such examples exist, but they do not detract from the fundamental theoretical point being made here.

hand, if you were to utter *You are not drinking coffee* to me as I utter *I am drinking coffee*, you *would* be contradicting me — negating the very content which I assert — even though the sentence you use is not the syntactically negated version of my sentence (or any other sentence expressing the same character). This is straightforward to account for if we claim that *You are drinking coffee* expresses the same content when addressed to an individual *x* at time *t* as *I am drinking coffee* expresses when uttered by *x* at *t*, but not so straightforward if we assume that character is the only significant level of interpretation.

I will also assume that content plays a specific kind of pragmatic role, namely that it serves as the object of direct illocutionary acts. If I make a direct statement, ask a direct question, or give a direct order, then what I directly assert, ask, or order is the content of the sentence I use in performing that speech act. For example, if I say "I am drinking coffee," what I directly assert is the content which the sentence *I am drinking coffee* has in the context of my assertion. Of course in so doing, I might also perform additional illocutionary acts indirectly: I might assert that I would not like tea, or invite you to have some coffee, or perform any number of other speech acts. It may in fact happen that indirect acts of this kind are much more central to my communicative intent in making the utterance than is the direct act, but they are indirect nonetheless. No doubt the objects of these other acts are things of the same type as the contents of sentences, but they are not the content of the sentence *I am drinking coffee* (even in context), in the sense of the term *content* intended here. I will assume that ordinarily, each sentence has exactly one content in context, aside from special cases like puns and deliberate double-entendres.[5]

Perhaps the best known arguments *against* a level of content like that adopted here are presented in Lewis (1980). Lewis argues separately against two different versions of this notion: one due to Stalnaker (1970) and the other to Kaplan (1989).

In Stalnaker's version, a sentence in context expresses a proposition, which can be construed as a set of possible worlds. Propositions also serve as the objects of mental attitudes and of various kinds of speech acts. Lewis argues that we cannot identify propositions in this sense with the "semantic values" of the sentences expressing them, because such propositions are not assigned compositionally: The proposition expressed by *Somewhere, the sun is shining* in a given context *c* is not determined by the proposition expressed by its constituent sentence *The sun is shining* in *c*.[6]

As Lewis notes, this argument is only as strong as the assumption that the clause *the sun is shining* in the larger sentence *Somewhere, the sun is shining* is syntactically identical to the free-standing sentence *The sun is shining*. If, for example, we took the constituent clause *the sun is shining* to contain a free variable ranging over locations, bound in the larger sentence by the quantifier *somewhere*, while the free-standing sentence *The sun is shining* contains instead a hidden indexical expression whose value is fixed by the context, the argument would hold no force at all. Lewis somewhat mockingly calls this move the "schmentencite" strategy, dismissing it as a "cheap and pointless." He does concede that "If the coordinates of indices were homogeneous in kind and unlimited in number — which they are not — it might be handy to use variables as a

[5] Even in such cases, it might be preferable to assume that each *use* of an expression (in the sense developed on p. 13 ff. below) has a single content, but allow more than one use of the same expression to occur simultaneously in the same context.

[6] The assumption seems to be that *The sun is shining* as used in *c* expresses the proposition that the sun is shining in the location of *c* — an assumption I think is highly questionable, for reasons to be discussed on p. 69 below. But for the sake of argument, let us temporarily grant the point.

device for keeping track of exactly how the truth of a schmentence depends on the various coordinates."

A detailed defense of the kind of analysis that Lewis dismisses in this way would take us too far afield, but I would like to suggest briefly that Lewis was wrong in claiming that we need to track only a limited number of coordinates at a time. The English language clearly provides the resources for tracking an arbitrary number of locations in the same sentence, through the use of the locative demonstrative *there*. ("Don't put it there; put it there, there or there.") Of course *The sun is shining* does not contain an overt demonstrative, but sentences without overt demonstratives which nonetheless require multiple locations to be tracked simultaneously are not hard to construct:

(4) Wherever it rains, it snows nearby and sleets in between.

The need to track multiple values is not limited to location parameters, but has been argued for times and worlds as well; see Cresswell (1990) among many others. The use of variables or some equivalent device rather than, or in addition to, a small, fixed set of indices is not at all pointless, but necessary.

Let us move now to Lewis' response to Kaplan's version of content. As Lewis presents it, Kaplan's primary motivation for positing the level of content is to model "what is said" in using a sentence: If John says "I am hungry" and Mary says to John "You are hungry," we regard them as having said the same thing; hence we analyze the sentences they use as expressing the same content in their respective contexts. Likewise if John says "I am hungry" and Mary says "I am hungry," they have said different things, even though they used the same words; hence we analyze these words as expressing different contents in the contexts of the two utterances. Lewis argues that this sort of argument is not compelling:

> In every [such] case, the proper naive response is that in some sense what is said is the same for both sentence-context pairs, whereas in another — equally legitimate — sense, what is said is not the same. Unless we give it some special technical meaning, the locution 'what is said' is very far from univocal. It can mean the propositional content, in Stalnaker's sense (horizontal or diagonal).[7] It can mean the exact words. I suspect it can mean almost anything in between. True, what is said is the same, in some sense... So what, unless the sense in question is more than one among many?

Lewis is certainly right that the pretheoretic locution 'what is said' admits of a wide variety of interpretations, including propositional content and the specific words uttered (though I am not sure what there is "in between" these two notions). I would distinguish content from other notions not by identifying it with "what is said" in this pretheoretic sense, but by appealing to content as the natural level at which to define contradiction, necessary truth and related semantic notions, and by assigning contents a theoretical role as the objects of direct illocutionary acts, as just outlined. Lewis develops a notion of "intension" which is more closely analogous to Kaplanian character

[7] The reference is to Stalnaker (1970), Stalnaker (1979).

than to Kaplanian content; but defining contradiction, necessity, etc., at such a level is markedly more complicated. Moreover, following Frege, I take what would ordinarily be the content of an expression to serve in opaque contexts as its denotation, as discussed in more detail in Section 1.5, below,[8] but no such role appears to be available for character.

I conclude that Lewis' arguments do not provide compelling reasons to reject a notion of content intermediate between character and denotation, forming the natural locus for defining contradiction and related notions, and serving in the various other ways just discussed.

Certain kinds of contents are the sort of thing that can be true or false. In order not to prejudge our main issue, let us assume that at least the contents of ordinary declarative sentences dealing with matters of fact have truth values, and leave open for now whether the contents of other sorts of expressions — particularly sentences dealing with matters of opinion — have truth values as well.

In order to use contents to account for the distinction between necessary and contingent truth as discussed above, we will employ the familiar technique of assigning them truth values relative to possible worlds. I will ultimately argue that we need to relativize the truth values of sentence contents to additional indices as well, but for now, let us assume that the only parameter relative to which sentence contents are assigned truth values is a possible world parameter. More generally, contents of linguistic expressions will be assigned denotations relative to worlds.

It will be useful to adopt some notational and terminological conventions, so that we may talk about contents and denotations in a precise and compact manner. As already mentioned above, our theory will assign contents to *uses* of linguistic expressions. We write '$\text{USE}(\alpha, \boldsymbol{\alpha})$' to mean that α is a use of expression $\boldsymbol{\alpha}$. The intended sense of "use" here is one where each use (in each possible world where it occurs) is tied to a particular time and place, a particular speaker, etc. — use-tokens, rather than general ways in which an expression may be used.[9]

The same use of a sentence may occur in more than world, and have different characteristics in the worlds where it occurs. For example, suppose in the actual world, John uses the sentence *I see you now* at 11:55 a.m. on July 9, 2014, addressing Mary, thereby expressing a particular content. A different content would be expressed if this use had occurred at 11:56 a.m. instead, or if it had been addressed to Bill, or if the conventions of the English language had been significantly different. Uses will therefore be assigned contents relative to worlds.

We may consider the time, place, speaker, etc., with which a use is associated in a given world as together forming the *context* of that use in that world. Each use has a unique context in any world where it occurs. Because the context of a use varies by world, the assignment of

[8] In connection with this it is worth noting Montague (1970) takes the distinction between "meanings" (which are closely analogous to Kaplanian characters) and "senses" (which are closely analogous to Kaplanian contents) to lie precisely in the fact that senses sometimes serve as denotations, while meanings do not.

[9] We could as easily use terms like *token* or *occurrence* in place of *use*, but I prefer *use* because I find that these other terms — especially *token* — are too often understood in such a way that a single orthographic inscription of an expression counts as a single token, no matter how many times it is used; and this is not the notion I have in mind. If I have a card inscribed with the word *yes*, and show it to you to answer one question on Tuesday and different question on Thursday, this counts as two uses of the word, with different contents.

contents to uses relative to worlds amounts to an assignment of contents to expressions which varies with the context.[10]

Let us use boldface lower-case Greek letters as variables over linguistic expressions, plain lower-case Greek letters as variables over uses of such expressions, and plain upper-case Greek letters as variables over contents. We relate contents to uses via "double-round" brackets, and denotations to contents via "double-square" brackets:

(5) Notation for contents and denotations:
   a. $⦇\boldsymbol{\alpha}⦈^{u}$: the content of use $\boldsymbol{\alpha}$ in world of use u.
   b. $[\![A]\!]^{w}$: the denotation of content A relative to world w.
   c. $[\![\boldsymbol{\alpha}]\!]^{u,w} = [\![⦇\boldsymbol{\alpha}⦈^{u}]\!]^{w}$

Note that possible worlds are used in two distinct ways here. Worlds are used in the assignment of contents to uses, and again in the assignment of denotations to contents. This is essentially the technique of Stalnaker (1979) and related work; because we are assigning contents to uses rather than expression-types (as, for example, in Kaplan (1989)), fixing a world suffices for the determination of content, because uses are by their nature contextually bound (on a world-by world basis).

Employing a familiar technique, we let the denotations of declarative sentence contents be two objects called **truth** and **falsity**, and give preliminary definitions of the predicates 'true relative to' and 'false relative to' in terms of these:

(6) Φ is true relative to w iff $[\![\Phi]\!]^{w}$ = **truth**; Φ is false relative to w iff $[\![\Phi]\!]^{w}$ = **falsity**.

We may notate the actual world as $w_{@}$ and define a monadic, unrelativized truth predicate for as in (7):

(7) Φ is true iff Φ is true relative to $w_{@}$.

Some readers may object to this definition in that it defines simple, unrelativized truth in terms of truth relative to a world — even though, as many people intuitively feel, unrelativized truth is in some sense the more basic or primitive notion, from which truth relative to worlds is an abstraction. This objection is misguided. The direction of definition in a semantic theory should not be construed as implying that the *definiens* is cognitively or metaphysically more primitive than the *definiendum*. An analogy may be useful here to the definition of multiplication in arithmetic. At a very basic level, multiplication is simply iterated addition: 3×5 is what you get by starting with 0 and adding 5 to it 3 times. Nearly all of us, I think, first learned multiplication in terms of addition as children using something like this definition, and many of us, no doubt, still

[10] I treat uses as formal primitives, but if desired, the analysis could easily be reformulated to model uses as pairs of an expression with a partial function from worlds to contexts. From here, it is a short step to recast the analysis as assigning contents to expressions relative to such functions, rather than assigning them to uses. I do not regard the choice between a theory which assigns contents to uses/tokens/occurrences and theory which assigns contents to expressions relative to contexts or other parameters as particularly significant.

think of multiplication as basically nothing more than iterated addition. Yet this conception of multiplication will not work in the general case. For example, it will not extend to fractions, because it makes no sense to iterate an operation a fractional number of times. Therefore, we move from the cognitively more primitive definition in terms of iterated addition to a more abstract, sophisticated, algebraic definition of multiplication, as an operation obeying certain laws (commutativity, associativity, distributivity over addition, etc.) Given such a definition, we may identify the operation which reduces to iterated addition as a special case. In the same way, we may acknowledge that the notion of truth relative to a world (or any other index) is an abstraction from the more cognitively primitive notion of truth *simpliciter*, yet identify truth *simpliciter* as a special case of this more general concept. The definition in (7) may be understood as identifying it in this way.

1.5. Homomorphic interpretation and differentiation of content

The assignment of truth values to sentence contents, together with the identification of contents with the objects of illocutionary acts, has the effect of constraining how content relates to syntactic structure — at least if we wish to respect certain intuitions about how truth relates to reference. Intuitively, truth depends on reference in the following way, which I will put at first very informally:

(8) If you say something true, then swap out part of the sentence and replace it with something else, and the result is that you say something false, then you must not have been talking about the same thing in the part that you swapped out as you were in the part that you replaced it with. (And likewise if you say something false, then swap out part of the sentence and say something true.)

This principle seems to me to be quite obvious, but also obviously needs some clarification and tightening up if we expect to get much mileage from it in a formal linguistic theory. As stated in (8), it appeals to the notions of "saying something" and "talking about," but these phrases can be construed in several different ways, which we must keep separate in a precisely stated analysis. I would suggest that "saying" here should be identified with *directly asserting*, so that (under the assumptions outlined above) what is "said" in this sense may be identified with the content of the sentence used in making the utterance. Similarly, the notion of "talking about" which is relevant in (8) is one in which what is "talked about" in using a particular part of a sentence (in context) is the contribution of that part of the sentence to determining the truth value of the content of the sentence — in other words, what we would ordinarily call *denotation* or *semantic reference*. The notion of "parts" of a sentence also requires some clarification, and a great deal might be said in this regard, but all that is necessary for the moment is an assumption that uses of sentences are built up from uses of more basic expressions through successive application of syntactic operations, and that the "parts" of a sentence use φ are the uses of more basic expressions to which these operations apply in deriving φ.

Under this interpretation (and maintaining our identification of the truth values of sentences with their denotations) it is straightforward to show that the substitutivity claim in (8) is equivalent to a claim that semantic operations can be defined corresponding to the syntactic operations so that the function mapping each expression use onto its denotation is a

homomorphism with respect to the two sets of operations (see, e.g., Lasersohn (2009a)). An initial formulation of this claim might be the following:

(9) For each syntactic operation *F* there is a corresponding semantic operation *G* such that for all $\alpha_1,...,\alpha_n$, *u*, *w*: $[\![F(\alpha_1,...,\alpha_n)]\!]^{u,w} = G([\![\alpha_1]\!]^{u,w},...,[\![\alpha_n]\!]^{u,w})$

This principle will be familiar to many readers as a version of the "Principle of Compositionality," but some caveats are necessary if we are to adopt this terminology. The term *compositional*, in its sense from modern semantic theory, derives as far as I know from Katz and Fodor (1963), who appeal to the idea of compositionality to explain how speakers can understand novel sentences. It is by no means clear that Katz and Fodor intended this terminology as meaning that interpretation is assigned *homomorphically*, and there is nothing in this article to suggest that their argument is supposed to apply at the level of denotation. (Their concern is with formal feature-marking.) As I have argued elsewhere (Lasersohn (2009a)), a principle of homomorphic denotation assignment does not provide a reasonable account of speaker's abilities to understand novel sentences; nonetheless, it is important theoretically because it captures the dependence of truth on reference, as just discussed.

The principle of homomorphic denotation assignment as given in (9) hides some complexity which may not be apparent, particularly if one is used to thinking of compositionality in terms of expressions rather than uses, or thinking of context as held constant throughout a complex expression (as for example in Kaplan (1989)). In (9), the relevant syntactic operations are ones that apply to uses, not expressions.[11] The left-hand world index (notated *u*), which fixes the content for each use, is held constant — but there is no reason to suppose that each use among $\alpha_1,\ldots,\alpha_n$ will have the same context, even in a given world. This is as it should be — time passes during an utterance, new objects may become salient, and the conversational participants may move in space; these developments amount to changes in the context over the course of an utterance. Such changes to the context during the utterance of a complex expression can affect the denotations of that expression's parts — particularly the interpretation of indexicals like *now*, *here*, *this*, *that*, and so on. Moreover, the *linguistic* context of each part of a complex use is different, and we will want to allow linguistic context as well as non-linguistic pragmatic context to affect denotation assignment (particularly in the case of a Fregean reference shift, as will be discussed shortly).

For our purposes much of the significance of the idea of homomorphic denotation assignment comes from the fact that it motivates a *Fregean reference shift* in certain grammatical constructions, and from the fact that this shift has implications for the identification and differentiation of content. We need hardly review the reasons for positing such a shift in much detail: In certain syntactic positions, substitution of one term for what would ordinarily be another with the same denotation does not necessarily preserve truth value:

(10) a. John said that Mark Twain wrote *Huckleberry Finn*.

[11] This does not deny the theoretical importance (much less the existence) of syntactic operations defined at the level of expressions, of course. Operations at the various levels must relate to each other systematically, but a detailed exploration of this would take us too far afield.

b. John said that Samuel Clemens wrote *Huckleberry Finn*.

The fact that (10)a. can be true while (10)b. is false would be problematic for the principle of homomorphic denotation assignment, if we took the denotations of *Mark Twain* and *Samuel Clemens* to be identical, as they normally seem to be. Rather than abandon such a principle, we analyze these names as in these sentences denoting something other than what they normally denote. Analogous examples could be constructed using any expressions in the subordinate clauses of (10)a. and (10)b.; therefore we analyze all constituents of these clauses (including the clauses themselves) as undergoing a similar shift in denotation. We are already committed to the idea that the objects of illocutionary acts are the sorts of things that can serve as contents of sentences (see p. 8, above); (10)a. and (10)b. describe John as performing illocutionary acts, the objects of which appears to be the contents of the subordinate clauses; we therefore analyze these clauses as denoting in these examples something that might otherwise serve as their contents. Content in our sense thus plays the role of *sense* in Frege (1892).

The preceding argument establishes the shift of denotation to what would ordinarily be content only for constituents of clauses serving as complements to speech act verbs. However, a similar pattern occurs in the complements of verbs of mental attitude and in many other constructions, so we will need to posit some sort of reference shift in these syntactic positions as well. Several considerations point to the necessity of treating many if not all of these constructions as shifting denotation in the same way, namely to what would ordinarily be content. For example, the argument in (11) is valid, and this is most straightforwardly explained if we analyze the subordinate clause *Mark Twain wrote Huckleberry Finn* as denoting the same thing in the third line as it does in the second:

(11) John believes whatever Mary says.
Mary says that Mark Twain wrote *Huckleberry Finn*.
Therefore, John believes that Mark Twain wrote *Huckleberry Finn*.

We will also need to treat this clause as denoting the same thing when serving as complement to mental attitude verbs as it does when serving as complement to speech act verbs if we are to maintain the idea of homomorphic denotation assignment for examples like (12), where a single occurrence of this clause serves as the object of both a speech act verb and a mental attitude verb:

(12) John says and believes that Mark Twain wrote *Huckleberry Finn*.

If we accept that some uses of a complex expression denote what for other uses of the same expression would be the content, and further that when a complex expression occurs in a context which induces a Fregean reference shift, so do its constituent parts, it follows from (9) that not just denotations, but *contents* may be assigned homomorphically. We may therefore assume (13):

(13) For each syntactic operation F there is a corresponding semantic operation G such that for all $\alpha_1,...,\alpha_n$, u: $⟨F(\alpha_1,...,\alpha_n)⟩^u = G(⟨\alpha_1⟩^u,...,⟨\alpha_n⟩^u)$

The use of an expression's content as its denotation in specific syntactic positions provides us with a heuristic for identifying whether two occurrences of a phrase have the same content or

different contents in cases where this may be unclear. For example, a popular syntactic analysis of infinitival phrases like *to win the election* in *John wants to win the election* treats them as full clauses, with a "hidden" or phonologically empty subject interpreted as a pronoun anaphoric to *John*. This syntactic analysis might be taken as suggesting a semantic analysis in which the infinitival phrase expresses the same content as the sentence *John wins the election*. We can evaluate the plausibility of this semantic claim by considering how the infinitival phrase is interpreted in constructions where (i) we expect a Fregean denotation shift, so that the infinitive denotes what would otherwise be its content, and (ii) we have an independent reason to believe that a phrase in that position contributes its denotation only once in the derivation of a sentence. For example, the "Right Node Raising" construction involves a phrase which is interpreted with respect to two conjuncts, as in (14):

(14) John wants, and Mary expects, this piece of pie.

Here, the phrase *this piece of pie* in some sense "distributes" over the conjuncts *John wants* and *Mary expects*. The denotation of *this piece of pie* is the same with respect to both conjuncts: according to this sentence, the piece of pie that John wants is the same one as the piece of pie that Mary expects. The phrase *this piece of pie* contributes a single denotation to the sentence, not two different denotations to serve as arguments to the two verbs. An examination of additional examples would show that this is a general pattern for this construction, not a quirk of this particular sentence. When we turn to examples with an infinitival phrase, we find that we cannot maintain this generalization and also maintain our hypothesis about the interpretation of the hidden subject of the infinitive:

(15) John wants, and Mary expects, to win the election.

This sentence does not mean that Mary expects John to win the election, or that John wants Mary to win the election; rather, John's want is directed toward himself, and Mary's expectation is directed toward herself. We know from examples like (14) that the distributed phrase has the same denotation with respect to both conjuncts; in this example the conjuncts contain intensional verbs, so that the distributed phrase denotes what would ordinarily be its content; therefore the (ordinary) content of this phrase cannot be identical to the content either of *John wins the election* or of *Mary wins the election*.

We shall return in more detail to the issue of the contents of infinitival phrases in Chapter 7. For the present, the more important point is that our claims about content in general are constrained by our assumptions that content is the appropriate level for defining contradiction, serves as the object of speech acts and mental attitudes, is assigned homomorphically, and functions as denotation in certain syntactic positions. Maintaining these assumptions will limit our choices as theorists in deciding what kinds of contents we can analyze particular expressions as having, and in particular in deciding when expressions do or do not differ in content.

Chapter 2 : Dismissing the Easy Alternatives

In this chapter, we review some of the more obvious possible analyses for sentences about matters of opinion, and outline problems for these analyses as applied to "personal taste" sentences such as *Licorice is tasty* or *Roller coasters are fun*.[12] By showing the problems which these analyses face, I hope to provide some initial motivation for exploring a more unusual alternative, namely a relativist semantics in which some sentences express contents which have truth values only relative to parameters whose values are left indeterminate by matters of fact. None of the analyses presented in this chapter will be explored in complete detail, and some of them, perhaps, could be used as starting points for developing more sophisticated analyses that might avoid the problems I point out. My purpose here is not to present an argument by elimination, in which a relativist semantics is the only theory left standing after all other approaches have been debunked, but just to show that accounting for the full range of data in a non-relativist semantics is not as easy or as straightforward as one might at first suppose. Once we have set aside these oversimple analyses, the way will be clear to develop a different approach, and defend it against more sophisticated alternatives.

2.1. Indexical analyses

One family of analyses treats personal taste sentences as involving some sort of indexical reference. For example, if John says "Licorice is tasty," he would typically be expressing his own taste or opinion, so we might take his utterance as involving indexical reference to John himself, or perhaps to his "standard of taste." This style of analysis was in fact my own first impulse when I began thinking about such examples some time ago; I took predicates of personal taste to involve first-person indexical reference, essentially on the model of pronouns like *I* or *me*. On this view, John's utterance "Licorice is tasty" is an assertion that licorice is tasty for him. If Mary does not like licorice, she might well say "Licorice is not tasty," meaning that it is not tasty for her.

The idea that sentences expressing matters of opinion involve indexicality accounts for at least two notable features of sentences expressing personal taste or similar matters of subjective judgment or opinion. First, it captures the intuition, shared by many speakers, that the truth values of sentences like *Roller coasters are fun* or *Licorice is tasty* vary from person to person. In eliciting judgments whether sentences like these are true or not, one often finds (alongside simple answers like "Yes" and "No") that speakers frequently give responses like "It depends who you ask." If we analyze such sentences as semantically similar to (16)a. and b., where X is a pronoun-like indexical expression — perhaps equivalent to *me* or *us* — our analysis will represent these sentences as exhibiting an interpersonal variation in truth value, according to the value of X:

(16) a. Roller coasters are fun for X.
b. Licorice is tasty for X.

[12] This chapter roughly follows the presentation in Lasersohn (2005) sec. 4.

Second, this analysis captures very nicely the "faultlessness" of faultless disagreement cases. Suppose again that John and Mary are riding the roller coaster, Mary says "This is fun!" and John says "This is not fun!" These two sentences are analyzed as in (17), but with no requirement that X be fixed to the same value in both utterances:

(17) a. This is fun for X.
b. This is not fun for X.

The analysis might claim that the natural interpretation for Mary's utterance is one where X is fixed to Mary herself, so that she means that the ride is fun for her, while the natural interpretation for John's utterance is one where X is fixed to him, so that he means that the ride is not fun for him. In this way, we can regard John and Mary as "both right," with neither one making an error.

In considering this approach to personal taste sentences, it will be useful to have a basic theory of indexicality in place. Only against a background of some reasonably explicit assumptions about how indexical expressions in general are interpreted does it make sense to evaluate the success of an indexical analysis of personal taste sentences. In this short section I will give only a very brief and simplified account of indexicality, saving the task of a more detailed analysis to Chapter 3 and Chapter 5. But even a brief and simplified account should be enough to make clear the main challenges an indexical account must face.

In one way or another, nearly all analyses of indexicality relativize the denotations of indexical expressions to features of pragmatic contexts — a time, location, speaker, addressee, etc. The denotations of larger expressions containing indexicals, including the truth values of sentences, are relativized in the same way, with the denotation of a complex expression relative to a context being derived systematically from the denotations of its parts relative to their contexts. We might, for example, say that relative to any world *u* where it occurs, each use α of an expression **α** involves a speaker $speaker_{u,\alpha}$, an addressee $addr_{u,\alpha}$, a time of use $time_{u,\alpha}$, location of use $loc_{u,\alpha}$, and whatever other parameters seem necessary. We could then stipulate that relative to *u*, a use α of the pronoun *I* denotes $speaker_{u,\alpha}$, a use β of the pronoun *you* denotes $addr_{u,\beta}$, a use γ of the adverb *now* denotes $time_{u,\gamma}$, and a use δ of the adverb *here* denotes $loc_{u,\delta}$.[13]

Early formal semantic treatments of indexicality (Lewis (1972), Montague (1968)) typically allowed denotation to vary with context in this way, but did not employ a notion of *content* that also varied with context; rather, the content (or intension) of an expression α was modeled as a function mapping any context *c* (or *n*-tuple of parameter values supplied by *c*) onto the denotation of α relative to *c*. As already discussed in Section 1.4 above, this greatly complicates the characterizations of necessary truth, contradiction, and related logical notions. Therefore, we follow Kaplan (1989) (and also earlier authors such as Montague (1970)) in positing a level of content which varies with the denotations of indexicals.

[13] Simple stipulations like these mask a good deal of complication in how the denotations of indexicals are determined in context; see, e.g., Nunberg (1993). At this introductory stage a certain amount of simplification seems reasonable, however.

In Section 1.4 above, we adopted an assumption that contents are assigned denotations relative to possible worlds. We can get the desired variation in content very easily, and assign denotations at the same time, by making lexical stipulations roughly as follows:

(18) For any use α and worlds *u*, *w*:
  a. If USE(α, *I*), then $[\![\alpha]\!]^{u,w} = speaker_{u,\alpha}$
  b. If USE(α, *you*), then $[\![\alpha]\!]^{u,w} = addr_{u,\alpha}$
  c. If USE(α, *now*), then $[\![\alpha]\!]^{u,w} = time_{u,\alpha}$
  d. If USE(α, *here*), then $[\![\alpha]\!]^{u,w} = loc_{u,\alpha}$

Recall that $[\![\alpha]\!]^{u,w} = [\![\text{⟨}\alpha\text{⟩}^u]\!]^w$, so these rules indirectly constrain contents in addition to assigning denotations. Note that as defined here these words express contents which denote rigidly — *w* is absent from the right side of the final "equals" sign in all these rules, so that the only variation according to world is in the left-hand world parameter, which determines which content is expressed, not in the right-hand world parameter, which determines what that content denotes.

In contrast, ordinary content words like verbs typically express contents whose denotations do vary with worlds. We might constrain the contents of *sleep* and *see* as in (19), for example:[14]

(19) For any worlds *u*, *w*:
  a. If USE(α, *sleep*), then $[\![\alpha]\!]^{u,w} = [\lambda x\,.\,[\lambda t\,.\,x \text{ sleeps at } t \text{ in } w]]$
  b. If USE(α, *see*), then $[\![\alpha]\!]^{u,w} = [\lambda y\,.\,[\lambda x\,.\,[\lambda t\,.\,x \text{ sees } y \text{ at } t \text{ in } w]]]$

In these rules, *w* does appear on the right side of the final "equals" sign, so that denotation of each content will vary from world to world. Note however that *u* does not appear on the right side of the "equals" sign, so the context plays no role in determining content — a debatable point to which we shall return in Sec. 5.5. below.

Rules like those in (18) and (19) are compatible with, but do not require, an identification of contents with functions from possible worlds to denotations — so that the content of a use α of the pronoun *I* in a world *u*, for example, would be identified with a constant function mapping each possible world *w* onto the speaker of α in *u*, essentially as in Montague (1970) or the formal part of Kaplan (1989).[15] However, if we do not make this identification, then an expression may vary in content from context to context even if its denotation does not: nothing prevents a case where $\text{⟨}sleep\text{⟩}^{u_1} \neq \text{⟨}sleep\text{⟩}^{u_2}$ but for all *w*, $[\![\text{⟨}sleep\text{⟩}^{u_1}]\!]^w = [\![\text{⟨}sleep\text{⟩}^{u_2}]\!]^w$, for example. This seems quite counterintuitive to me, so I will henceforth identify the content of a use with the function mapping each world *w* onto that use's denotation relative to *w*:

[14] The λ-operator should be interpreted in the usual way: '[λ*v* . *A*]' denotes that function mapping any possible value *a* for the variable *v* onto whatever *A* denotes when *a* is fixed as the value of all free occurrences of *v* in *A*. If *A* is a declarative sentence, we assume that it denotes **truth** or **falsity**. In informal presentation, the possible values for variables will sometimes be indicated rather inexplicitly by the choice of letter: *x* or *y* for individuals, *w* for worlds, *t* for times, etc. Brackets will usually be dropped when no confusion can result. A λ-operator may be prefixed to an expression which already contains another λ-operator at the front; in this case we often suppress the dot which separates the operator from its scope.

[15] To put it in more Kaplanian terms, we would identify the content of the pronoun *I* relative to a context *c* as a constant function mapping each possible world *w* onto the speaker of *c*.

(20) For any use α: $\text{⟨}\alpha\text{⟩}^u = \lambda w\,.\,[\![\alpha]\!]^{u,w}$

This immediately raises the problem of how to differentiate the contents of uses of *different* expressions to whose denotations are identical in all worlds — such as *sleep or not sleep* and *be equal to itself* — something we apparently must do in order to account for the fact that such phrases are not intersubstitutable *salva veritate* in intensional contexts. I believe such examples can be adequately dealt with using existing techniques such as diagonalization more or less in the style of Stalnaker (1999) and Section 7.5. below. Readers who are skeptical on this point are invited not to make the suggested identification, but to substitute whatever their favorite theory of semantic content may require. Most of the analysis to be presented here will be compatible with a variety of views on this issue, as long as contents may be assigned denotations relative to worlds.

Now let us turn to predicates of personal taste, such as *tasty* or *fun*. Such predicates sometimes appear with an overt prepositional phrase headed by *for* or *to*, indicating the individual relative to whom the subject is tasty or fun:[16] *The licorice is tasty for/to John*; *Roller coasters are fun for Mary*. It would appear natural therefore to analyze *tasty* and *fun* as two-place predicates, so that for any use α of *tasty* and any use β of *fun*, and any individual *x*, *y* and worlds *u*, *w*: $\langle x, y\rangle \in [\![\alpha]\!]^{u,w}$ iff *x* is tasty for *y* in *w*, and $\langle x, y\rangle \in [\![\beta]\!]^{u,w}$ iff *x* is fun for *y* in *w*.

When a predicate like *tasty* or *fun* appears without an overt prepositional phrase, an indexical analysis may claim that its second argument place is implicitly filled by some unpronounced expression $\boldsymbol{\pi}$, interpreted in much the same way as an indexical pronoun. If we take $\boldsymbol{\pi}$ to be first person and singular, then for any use π of $\boldsymbol{\pi}$ and worlds *u*, *w*, $[\![\pi]\!]^{u,w} = speaker_{u,\pi}$. An alternative would be to claim that *tasty* and *fun* are lexically ambiguous between two-place predicates defined as above, and one-place predicates defined so that for all α, *x*, *u*, *w*: if USE(α, *tasty*), then $x \in [\![\alpha]\!]^{u,w}$ iff *x* is tasty in *w* for $speaker_{u,\alpha}$, and if USE(α, *fun*), then $x \in [\![\alpha]\!]^{u,w}$ iff *x* is fun in *w* for $speaker_{u,\alpha}$. Either way, the sentence *Licorice is tasty* is interpreted essentially like *Licorice is tasty for me*.

More specifically, it should be evident that at this point that if φ and ψ are two uses of the sentence *Licorice is tasty*, where $speaker_{u,\varphi}$ = John, but and $speaker_{u,\psi}$ = Mary, then $\text{⟨}\varphi\text{⟩}^u$ = [λ*w* . Licorice is tasty for John in *w*], while $\text{⟨}\psi\text{⟩}^u$ = [λ*w* . Licorice is tasty for Mary in *w*]. That is, if John says "Licorice is tasty," his use of this sentence has as its content the function which maps any world *w* onto **truth** iff licorice is tasty for him in *w*, while if Mary says "Licorice is tasty," her use of this sentence has as its content the function which maps any world *w* onto **truth** iff licorice is tasty for her in *w*. If, in the actual world $w_@$ licorice is tasty for John but not for Mary, then $\text{⟨}\varphi\text{⟩}^u$ will map $w_@$ onto **truth**, but $\text{⟨}\psi\text{⟩}^u$ will map $w_@$ onto **falsity**. The sentence is in this way analyzed as true when John says it, but false when Mary says it, capturing the subjective nature of the meaning of this sentence — or so the analysis claims.

Unfortunately, this analysis has an obvious problem — namely that it fails to account for the intuition that if John says "Licorice is tasty" and Mary says "Licorice is not tasty," they are contradicting each other. The natural characterization of contradiction in a possible worlds semantics is one in which two contents Φ, Ψ contradict each other iff there is no world *w* such that

[16] For reasons I do not fully understand, such prepositional phrases sound more natural to some speakers with *fun* than they do with *tasty*. However, it is not hard to find naturally occurring examples with both predicates; see Lasersohn (2005) fn. 3.

$\Phi(w)$ = **truth** and $\Psi(w)$ = **truth**. But as we just saw, it might well be the case that $\text{⟨}\varphi\text{⟩}^{u}(w_{@})$ = **truth** while $\text{⟨}\psi\text{⟩}^{u}(w_{@})$ = **falsity**, where φ and ψ are two uses of the sentence *Licorice is tasty*. Presumably, negation reverses the truth value of the negated content, so if χ is a use of *Licorice is not tasty* it could easily happen that $\text{⟨}\varphi\text{⟩}^{u}(w_{@})$ = **truth** while $\text{⟨}\chi\text{⟩}^{u}(w_{@})$ = **truth**. But then there is no contradiction. Under this analysis, if John says licorice is tasty, and Mary says licorice is not tasty, the content she is denying is not the same content as he is asserting, so she is not contradicting him. But this seems wrong — if John says that licorice is tasty, and Mary says it isn't, her claim is in direct contradiction to his.

To put it slightly differently, although this analysis captures well the "faultlessness" of faultless disagreement cases, it does not capture the "disagreement."[17] There is nothing in the analysis which predicts any difference between this sort of example, and one where John makes a factual claim using an overt first-person singular pronoun, while Mary makes a claim using a negative version of the same sentence:

(21) *John:* I am a doctor.
Mary: I am not a doctor.

But in this sort of example, we have no intuition of disagreement or contradiction at all. There must be something besides ordinary first-person singular reference involved in the personal taste examples, or our intuitions would be the same for them as for other examples involving first-person singular reference.

We could easily change the analysis so that John's and Mary's assertions *are* treated as contradicting each other, by giving up the assumption that the hidden indexical is first-person singular. Instead, we might treat it as first-person plural — equivalent to *us* — or perhaps as freely fixable to any pragmatically relevant group or individual. This allows us to make sense of some examples that were problematic under the view that the hidden indexical is always first-person singular.

For example, assume again that John and Mary are riding the roller coaster together, Mary says "This is fun!" and John says "No, it isn't!" If we take Mary's utterance as equivalent to *This is fun for us* and John's as equivalent to *It isn't fun for us* — where *us* refers to the group consisting just of John and Mary themselves, or perhaps to a group containing John, Mary, and other people like John and Mary in relevant respects — then their assertions are in direct contradiction to one another.

Moreover, we can make good pragmatic sense of John's response, at least if we assume that something can be fun for a group only if it is fun for (all) the members of the group. Since John is riding the roller coaster, he knows first-hand that it is not fun for him; it follows that it is not fun for any group of which he is a member; hence he is in a perfect position to refute Mary's claim that it is fun for the group.

Unfortunately, this explanation will not extend to the full range of examples. Consider a case where the order of John's and Mary's assertions is reversed:

[17] Here is the first of several ways in which the term "faultless disagreement" is not fully satisfactory. There are many ways two individuals can disagree with one another, not all of which involve making mutually contradictory assertions. But the problem with the current example is not just that we intuitively regard John and Mary as disagreeing with each other, but that their assertions directly contradict each other.

(22) *John:* This is not fun!
Mary: Oh, yes it is!

Here, as before, we feel that John and Mary are contradicting each other. We can account for that intuition by claiming that both their utterances involve indexical reference to the same group, which contains both John and Mary as members. But then it becomes very hard to make sense of Mary's response. John is riding the roller coaster; hence he is in a perfect position to know that the roller coaster is not fun for him. Under the assumption that something is fun for a group only if it is fun for all members of a group, it follows directly that the roller coaster is not fun for any group containing John, including whatever group is implicitly referred to in John's and Mary's utterances. Mary is therefore in no position to refute the claim that the roller coaster is not fun for the group. What justification could she have for contradicting John's claim? By contradicting him, she must be acting irrationally, or ignoring what he said, or claiming to know his own mind better than he does himself, or something similar. But in fact she does not seem to be doing any of these things; she is simply expressing disagreement with John's claim that the roller coaster is not fun.

Her response would be easy to make sense of if there were some sort of context shift between the two utterances, so that the hidden indexical referred to a different group or individual in Mary's utterance than it does in John's utterance. But then, of course, the two utterances would not contradict each other; we would be back where we started.

Perhaps the problem is with our assumption that something can be fun for a group only if it is fun for all members of the group. We might claim instead that it suffices to be fun for *almost* all of the members of the group (or even just for *most* of the members of the group, or for *normal* members, or *representative* members, etc.; a wide variety of options present themselves). This predicts that Mary's response in (23) should make sense:

(23) *John:* Was the roller coaster fun for the boys?
Mary: Yes, but some of them hated it.

To my ear, however, Mary's response sounds self-contradictory.[18]

Even if we grant that something can be fun for a group without being fun for all members of the group, however, we still face a problem. If John and Mary are riding the roller coaster, and he is having a terrible time, then he is perfectly justified in asserting *This is not fun!*, and Mary is very unlikely to convince him that it is fun just by showing him the results of a survey. But on the proposal in question, displaying the survey results should be an airtight argument: *This is fun!* is taken to mean nothing more than that the referent of *this* is fun for enough of the right members of the relevant group.

Intuitively, when Mary says "This is fun!", and John says "No, it isn't!", they aren't arguing about the prevailing view of some group, so that their dispute would be resolved if they made enough observations of enough different people riding the roller coaster. Rather, the dispute is an expression of John's and Mary's own conflicting opinions of the roller coaster. Each of them asserts

[18] See Lasersohn (1999) for related discussion, including a non-truth-conditional account of the greater tolerance for exceptions displayed by *the* than by *all the*.

his or her own personal perspective over and against that of the other.[19] An indexical analysis does not seem to capture this aspect of the dispute.

There is some potential for misunderstanding here, so I must emphasize that the central problem here is not one of explaining the conversational structure of disputes.[20] In fact, the whole idea of a conversational dispute is incidental to the fundamental problem which personal taste sentences (and other sentences about matters of opinion) present to a semantic theory. The underlying problem is that if one person asserts that licorice is tasty, or that a particular roller coaster ride is fun, and another person asserts that licorice is not tasty, or that the ride is not fun, their assertions are in direct contradiction to each other. This is so whether or not the two speakers are engaged in a conversational dispute.[21] If John, in Urbana, Illinois, asserts that licorice is tasty, and Mary, a stranger he has never met, in Santa Cruz, California, asserts that licorice is not tasty, their assertions contradict each other, even though they are not engaged in conversation with each other. Indeed, they need not even be aware of each other's existence. We might further say that John and Mary disagree with each other about the tastiness of licorice, even if neither of them makes an assertion, so long as John believes licorice is tasty and Mary believes it is not. As Cappelen and Hawthorne (2009) emphasize, there are two distinct senses of the word *disagree*: one in which it denotes a certain kind of social interaction — roughly, a conversational dispute — and another sense in which two people may be said to disagree even if they do not interact in any way.[22] The problem of faultless disagreement is by no means limited to the first of these senses.

2.2. Quantificational analyses

Occasionally, semanticists have claimed that personal taste sentences involve some hidden argument to predicates like *fun* or *tasty*, but rather than treating this argument as semantically like an indexical pronoun, have treated it as a quantified variable. For example, Chierchia (1984) suggests a formula essentially like the following for the sentence *Dancing is fun*:[23]

(24) $\exists x[\mathbf{fun'}(x)(\cap\mathbf{dance'})]$

In this analysis, *Dancing is fun* is true iff dancing is fun for at least one individual. The existence of such an individual is presumably a fixed fact for everyone, so the truth value will not vary from person to person. Thus, if Mary says *Dancing is not fun*, she will be directly contradicting John's assertion of *Dancing is fun*.

---

[19] Of course I do not mean *assert* here in the technical sense from speech act theory, in which the object of an assertion should be the sort of thing which can serve as the content of a declarative sentence. Rather, this is the ordinary, pretheoretic use of *assert*, in which one can assert one's interests, values, perspective, etc.

[20] Contrary to what Snyder (2013) and Barker (2013), for example, seem to assume.

[21] *Pace* Glanzberg (2007).

[22] Cappelen and Hawthorne describe this in part as an aspectual difference, with *disagree* serving as an activity verb under the first reading and a stative verb under the second reading; but aspectual classification is not really crucial to the argument.

[23] I have made a few minor adjustments to Chierchia's formula to increase clarity, omitting the generic tense operator, expanding his "ADROP" operator, and reducing the formula by λ-conversion. The ∩-symbol is a nominalization operator corresponding to the gerundive *–ing* suffix in *dancing*.

However, the truth conditions assigned on this analysis are simply much too weak. Suppose John doesn't like roller coasters at all. Then he can sincerely say *This is not fun* as he rides a roller coaster, even if he knows that many other people do enjoy them. But under this analysis, he could not say that, at least not sincerely. So it is implausible simply to existentially quantify the hidden argument.

If existential quantification is too weak, we might consider using some stronger kind of quantification, such as universal or generic quantification. Suggestions to this effect may be found in Epstein (1984) or Bhatt and Izvorski (1995), for example. Under this analysis, we would be taking *This is fun* to be semantically equivalent to *This is fun for everyone*, or to *This is fun for people in general*, or something along these lines.

Like other forms of quantification, the quantification in this case is presumably over some pragmatically restricted domain, so perhaps it would be better to paraphrase the reading assigned as *This is fun for every relevant individual*, or *This is fun for the relevant individuals in general*. But now we are back to an analysis we considered and rejected just a few pages ago! If we analyze *This is fun* as meaning it is fun for all the relevant individuals, it is hard to make sense of examples like (22), in which John asserts that the ride is not fun, and Mary contradicts him, saying that it is. And if we say that *This is fun* means that it is fun for the relevant individuals in general, then (23) should sound fine, and the issue of whether something is fun should be resolvable by taking a survey. But this is wrong; you can sincerely assert that something is fun as long as you believe it is fun for *yourself*, even if you know that other people don't enjoy it. If John enjoys himself while compiling a catalog of his paper clip collection, he might sincerely say "This is fun!" even though he is aware that most people would find this activity very tedious. If sentences like *This is fun* involve generic quantification over individuals, this sort of assertion should not be licensed. Moreover, sentences like (25)a. should make perfect sense, while (25)b. should be self-contradictory:

(25) a. This is not fun at all, although I'm having fun doing it.
b. This is fun, but most people would hate it.

The actual pattern we observe is the opposite: (25)a. sounds internally incoherent, while (25)b. is completely natural.

2.3. Absolutism and ignorance

Another possibility is that predicates like *tasty* and *fun* in sentences like *Licorice is tasty* or *This is fun* are simple one-place predicates, with no hidden arguments, indexicality, or context-dependence at all, so that these sentences express propositions which are true or false absolutely, the same for everyone. This approach could perhaps be regarded as the analog in the area of predicates of personal taste to Williamson's (1994) analysis of vague predicates. Williamson argues that the meanings of apparently vague predicates actually have sharp boundaries, so that, despite appearances, there is always a definite fact of the matter as to whether someone is thin or not, for example; and we would likewise be claiming that there really is a definite fact of the matter as to whether roller coasters are fun or not.

Williamson's analysis depends on the claim that even though vague predicates have sharp boundaries, we have no way of knowing, or even discovering, where they are. This state of uncertainty is the source of our intuition of indeterminate boundaries. Likewise, one could claim that

there is a definite fact of the matter whether licorice is tasty, or roller coasters are fun, but we simply have no way of discovering it. This provides an interesting analysis of faultless disagreement cases: even though one of the two parties to the disagreement must be making an error of fact, we would in principle never have any way of discovering which one it is. Hence the disagreement would never be resolvable, and the party who makes a factual error could not be held responsible for it. In that sense, neither party is at fault, so the disagreement is faultless despite the factual error.

There is some appeal to this analysis, but it should be carefully noted that there is one important respect in which issues of taste are disanalogous to issues of where to draw the boundary with vague predicates. In vagueness examples, we feel a good deal of hesitation to assert or deny claims about borderline cases. Suppose, for example, that John is neither clearly thin, nor clearly not thin, but falls into that portion of the thinness scale between the clearly thin and the clearly not thin. In this case, we hesitate to assert either (26)a. or (26)b.:

(26) a. John is thin.
b. John is not thin.

This is as expected in a Williamson-style treatment of vagueness, since in such a case, according to the analysis we have no way of knowing whether John is thin or not thin, and therefore would not have adequate grounds for making an assertion.

If we treat taste sentences in an analogous way, claiming that we are in a state of ignorance about those issues of taste in those cases where we observe variations in judgment from person to person, and that we have no way of discovering the facts, then we should expect a similar degree of hesitation to make assertions and denials in such cases. However, we frequently feel no hesitation at all in making assertions of tastiness or fun based on our own judgments, even if we know that other people make the opposite judgment. If you have tasted licorice, enjoyed it, and remember the experience,[24] you are licensed to say that it is tasty, even if you know that others have the opposite reaction. If you ride the roller coaster and hate it, you may confidently deny that it is fun, even as you see other riders enjoying themselves.

I suggest that our willingness to make these sorts of assertions is because in the taste cases, unlike the vagueness cases, we speak from a position of epistemic *privilege*, not a position of inescapable ignorance. In fact, this is an even stronger privilege than we get even from direct observation. If I observe a car, perceive it to be red, and make an assertion to that effect, I could still be in error, for example if I am color-blind. But even I have an unusual tongue defect which makes me experience flavors differently from most people, if I try the licorice and like it,[25] I am justified in saying it is tasty.

None of this conflicts with the fact that our tastes can change, or with the fact that we sometimes reconsider our own taste judgments after talking with others or observing their reactions.

[24] We might add a few more clauses to deal with special cases: if you feel confident that the licorice you tasted was typical of licorice in general, and that the circumstances under which you tasted it did not involve unusual changes to your taste perception, and that you attended properly to the sensation of taste as you ate the licorice, and that your memory is accurate, etc. None of this detracts from the point that one may make assertions about tastiness based purely on one's own experience.

[25] Again, if I am confident that I experienced the flavor as it would normally taste *for me*, and that I was attending to it properly, etc.

As our tastes change, the taste claims we are justified in asserting naturally change as well. Of course if our taste judgments do change, then we disagree with our previous judgments, just as we may disagree with the judgments of other people. This is just faultless disagreement with one's former self.

Nor does the claim that we make taste assertions from a position of epistemic privilege conflict with the fact that we *do* sometimes hesitate to make assertions of fun or tastiness in cases where we expect that someone else — who may not share our tastes — will use our assertion as the basis for deciding their own course of action. In such cases, more than the simple truth of the assertion is at issue; the practical consequences for the addressee are also a concern. In general one hesitates to make assertions which may lead to unpleasant consequences for one's addressees, even when the assertions are true; this case is no different.

There is a second, more semantic problem with the idea that predicates like *fun* and *tasty* have no hidden arguments, indexicality or relativization at all, so that they express fully objective properties. If we claim this, we still must account for examples in which these predicates appear with a prepositional phrase headed by *for* or *to*, as in *Roller coasters are fun for Mary* or *Licorice is tasty to John*. The preposition presumably must combine with an individual argument (the object of the preposition) to form a phrase which can take a one-place predicate as an argument to form a complex one-place predicate. We should be able to specify its denotation by filling in the question marks in (27):

(27) If USE(α, *for*) then $[\![\alpha]\!]^{u,w} = \lambda x \lambda P \lambda y$. ????

In terms of Montague-style type theory, it seems clear that *for* must denote a function of type $\langle e, \langle\langle s, \langle e, t\rangle\rangle, \langle e, t\rangle\rangle\rangle$, but I, for one, am at a loss as to how to say which one, except disquotationally as in (28):

(28) If USE(α, *for*) then $[\![\alpha]\!]^{u,w} = \lambda x \lambda P \lambda y$. *y* is *P* for *x* in *w*

But this leaves it a mystery as to how *for* operates on the values of *P* and *x* to derive the complex predicate *P for x*. It does not provide us with anything like an algorithm or system for determining which things are *P* for *x* in *w* in terms of *P*, *x* and *w*. The analysis presumably treats *for* as a non-logical constant, whose denotation varies arbitrarily from model to model. It would be preferable, if we could, to treat *for* as a logical operator, with a systematic semantic effect. Only if the analysis makes clear predictions, independently of our intuitions about the word *for*, as to how phrases of the form *P for x* relate to *P* and to *x*, can we test whether those predictions conform to our intuitions about what *for* actually means. The analysis I give below in Section 4.5. will make such predictions, but it does not appear that an analysis that takes *fun* and *tasty* as simple, objective, one-place predicates, with no hidden arguments, indexicality or other relativization, can do so.

2.4. Expressivism

Another possible approach to taste sentences is to claim that they lack truth values entirely. If the

truth conditions assigned under an indexical analysis, a quantificational analysis, and an absolutist analysis all seem problematic, perhaps it is because we were wrong to assume that sentences like *Roller coasters are fun* or *Licorice is tasty* even have truth values to begin with. Certainly, the intuition is strong that truth is a matter of conformity to the facts, and if there are no facts of the matter whether roller coasters are fun or licorice is tasty, this would suggest that categories like truth and falsity should not apply to such sentences.

Perhaps the most obvious way to work out this position is to claim that these sentences lack truth values because they are used to perform some other sort of direct illocutionary act besides making assertions, and that the assignment of truth values and truth conditions only makes sense for assertive speech acts and not for other kinds.

If such sentences are not used to make direct assertions, what sort of illocutionary act is involved? One obvious possibility is that they simply express some aspect of the speaker's affective or emotional state.

To make this idea concrete, consider some examples which less controversially serve merely to express affective state — including expressing a sense of fun or tastiness:

(29) a. Whee!
b. Mm-mm.
c. Oh, boy!

Utterances like these do not make assertions; they serve as conventional expressions of what the speaker is feeling. In claiming that sentences like *Roller coasters are fun* or *Licorice is tasty* likewise are used to perform acts of affective expression rather than make assertions, we would be assimilating them more-or-less to examples like those in (29).

No one would deny that sentences like *This is fun!* or *This is tasty!* can be used to perform acts of affective expression. (Nor, for that matter, need an expressivist deny that such sentences are used, in part, to make statements.) The issue is whether they do so *directly* in the sense discussed in Chapter 1, p. 8. An utterance may involve the simultaneous performance of several speech acts, including both assertive and non-assertive acts. But we have been assuming that one of these acts has special status as the "direct" illocutionary act performed in making the utterance, and that the compositionally assigned semantic content of the uttered sentence serves as the object of this direct illocutionary act. To this I might add that the type of speech act which a sentence is used to perform directly seems itself to be fixed by the syntactic form of the sentence as part of the grammatical conventions of the language — assertions for declarative sentences, questions for interrogatives, etc.[26] In the analysis under consideration in this section, it is claimed that sentences like *Roller coasters are fun* or *Licorice is tasty* do not make direct assertions in this sense; their compositionally assigned semantic contents are not asserted, and the syntactic forms of such sentences do not fix assertion as the illocutionary act they are used to perform as part of the grammatical conventions of the language.

This style of analysis faces a number of challenges. First, there is the fact that people do, in

[26] Perhaps one could maintain that different sentence types are assigned different types of semantic content, each serviceable as the object of a different class of illocutionary acts, rather than claiming that the sentence types themselves are directly associated with different kinds of illocutionary acts; but this is an open question.

fact, apply the predicates *true* and *false* to taste sentences (or their contents). There is nothing at all unnatural about the dialog in (30), for example:

(30) *Mary:* This is fun!
John: That's not true! This isn't fun at all!

In contrast, examples like (31) sound completely unnatural:

(31) *Mary:* Whee!
*John: ??*That's not true! This isn't fun at all!

John's response is anomalous even though Mary's utterance strongly implicates that whatever activity she is engaged in is fun. This pattern provides at least *prima facie* evidence that sentences like *This is fun* have the kind of semantic contents which can be true or false.

Another concern is one classically raised by Geach (1965): Sentences like *This is fun* or *This is tasty* can appear embedded under truth-functional connectives and other logical operators, and participate in the usual logical consequence relations which such embeddings give rise to. One would like to preserve the idea that (32) is an ordinary example of Modus Ponens, for example:

(32) If there is a loop, the roller coaster is fun.
There is a loop.
Therefore, the roller coaster is fun.

But it is quite hard to see how to maintain this idea if sentences like *The roller coaster is fun* do not have truth values. An argument is logically valid if the truth of the premises provides a certain sort of guarantee of the truth of the conclusion. If the conclusion of (32) is not the sort of thing that can have a truth value, then (32) is not a valid argument — probably we should not even call (32) an argument at all on this view. But this seems very strange.[27]

We might attempt to preserve the claim that (32) is a valid argument by reconstructing the whole field of logic around some concept other than truth, but one rather suspects that then we are simply applying the word *logic* to a different field than usual — not that we are giving an alternative theory of the same topic.

In addition to the concern that taste sentences appear to function in logic as though they had truth values, we may add the concern that they function *grammatically* just like sentences which are

[27] An expressivist might claim that (32) is a valid argument because the first premise is not the sort of thing that can have a truth value; hence there is can be no case where the premises are true but the conclusion false. But saying this does not solve the more general problem. Just as we want our theory to portray (32) as valid, we want it to portray the following argument as *in*valid:

(i) If there is a loop, the roller coaster is fun.
There is a loop.
Therefore, the roller coaster is not fun.

But if (32) is vacuously valid because there can be no case where the premises are all true, then (i) must be vacuously valid in the same way.

used to make direct assertions about matters of fact. As already discussed on pp. 3–4, sentences expressing matters of opinion have the same internal syntactic structure as sentences expressing matters of fact, have the same external syntactic distribution, show the same alternations in mood and sentence type, appear as complements to the same speech act and attitude verbs — in fact, as far as the grammatical system of the language is concerned, they appear to behave in every way like ordinary truth-value-bearing clauses. A linguistic analysis which treats them as fundamentally different would seem to be missing a key point.

Finally, if we claim that these sentences are neither true nor false, it becomes very unclear how to capture one of the original intuitions that opened the whole inquiry, namely, the intuition that if John says something is fun, and Mary says it's not fun, they're contradicting each other. If the sentences aren't even the kinds of things that can have truth values, it is hard to see in what sense a contradiction could arise.

2.5. Metalinguistic and metacontextual conflict

Perhaps we could maintain the idea that taste sentences have truth values, and still account for the intuition that two people disagree if one of them says *Roller coasters are fun* and the other says *Roller coasters are not fun*, if we were to claim that the disagreement was not really about whether roller coasters are fun, but about something else. At first blush this may seem like a strange option, given the forms of these two sentences; but in fact any sentence may be used to convey all kinds of information other than its usual semantic content, and it is quite possible that our two speakers convey mutually conflicting information of some other kind.

One way we might work out this idea in more detail is by noting that in addition to their ordinary semantic content, sentences can be used to convey metalinguistic information about the meanings of the words involved. For example, if Mary says "This is fun" while riding the roller coaster, part of what she is saying is a comment on the roller coaster, but her utterance also clearly conveys some metalinguistic information as well, namely that she considers the word *fun* to have a meaning which applies to the roller coaster. If you already had a clear sense of what the roller coaster was like, but were unsure what the word *fun* meant — or in particular what Mary meant by *fun* — her utterance would clarify the meaning of the word for you, at least to some extent.

Taking the analysis a step further, if John then objects, "This is not fun," we could take him not to be arguing with Mary about the roller coaster at all, but instead about the interpretation of the word *fun*. We could then explain the intuition that John and Mary disagree as due to the fact that John and Mary take opposite sides in this conflict over the meaning of the word, rather than appealing to a contradiction between the ordinary semantic contents of their utterances.

A related approach would exploit the fact that if the interpretation of a sentence varies from context to context, then using the sentence may convey not just its ordinary semantic content, but also information about the context itself. This point is interestingly exploited by Barker (2002), for example: If one of the functions of the pragmatic context is to specify boundaries for vague predicates like *tall*, then uttering a sentence like *Feynman is tall* in a situation where it is clear exactly how tall Feynman is will not tell us anything about Feynman's height, but will tell us something about where the boundary between the tall and the non-tall is — in other words, it will give us information about what pragmatic context we are in. Especially since some aspects of interpretation may depend on features of the context which are not obvious or publicly accessible to all the discourse participants, such as speaker intentions or background assumptions, it is not

surprising that linguistic utterances may sometimes serve more to clarify the context than to convey their ordinary semantic content.

With this in mind, we might claim that predicates like *fun* and *tasty* have a hidden indexical argument whose value is fixed in context in such a way that its referent is not always obvious; *Licorice is tasty* would be treated as meaning something like "Licorice is tasty for the relevant person or persons," where it may not always be clear to the discourse participants which people are relevant. This uncertainty amounts to uncertainty about the context of utterance, so we may take disputes about whether licorice is tasty to be disputes about what context the discourse participants are in — "metacontextual" disputes.

Whether we take faultless disagreement to be metalinguistic or metacontextual, a useful technique for identifying the content on which the disagreement turns is *diagonalization* in the sense of Stalnaker (1999). This technique relies on the fact that uncertainty about the content of a sentence amounts to uncertainty about what world one is in. If one is unsure whether *fun* (or some particular person's use of *fun*) means "exciting" or "amusing" or "relaxing," that amounts to uncertainty whether one is in a world w_1 where it means "exciting," or w_2 where it means "amusing" or w_3 where it means "relaxing." Likewise, if we take *fun* to have a constant meaning (in the sense of character) but different contents in different contexts depending on the value of some hidden indexical element, then uncertainty about the value of this indexical amounts to uncertainty about the world: if Mary utters "Roller coasters are fun" and we are unsure about whether the value of the indexical is Group A, Group B or Group C, this amounts to uncertainty whether we are in a world w_1 where Mary's utterance took place in a context where Group A is relevant, or a world w_2 where it took place in a context where Group B is relevant, or a world w_3 where it took place in a context where Group C is relevant. In this case, the sentence will express different contents in w_1, w_2 and w_3.

In either case, the content expressed by *Roller coasters are fun* in w_1 may itself be evaluated for truth or falsity relative to w_2 and w_3, and not just w_1; and likewise with the contents expressed in w_2 and w_3. In fact, this is the normal conception of truth conditions; we hold the content of a sentence constant as we evaluate it from world to world. However, Stalnaker suggests, under the right pragmatic circumstances, we may reinterpret a sentence as expressing its *diagonal* proposition — the proposition which is true at a world *w* iff the content which the sentence normally expresses at *w* is true in *w*. For example, if the content expressed at w_1 is true at w_1 and w_2 but not w_3, and the content expressed at w_2 is true at w_1 and w_3 but not w_2, and the content expressed at w_3 is true at w_2 but not w_1 or w_3, then the diagonal proposition will be true at w_1 but not w_2 or w_3. This is illustrated in (33):

(33)

		Worlds as points at which content is evaluated		
		w_1	w_2	w_3
Worlds as determinants of content	w_1	T	T	F
	w_2	T	F	T
	w_3	F	T	F

Now if Mary says "This is fun!" and John says "No, it isn't!" we may analyze the disagreement as based around a reinterpretation of her utterance as expressing its diagonal proposition — which is true iff whichever proposition the

sentence *This is fun!* would ordinarily express in the world they are in is in fact true in whichever world that is. If, for example, Mary takes them to be in world w_1 where *fun* means "exciting," but John takes them to be in world w_2, where *fun* means "amusing," he may justifiably deny this proposition, and express his denial with *No, it isn't!* Viewed as expressing diagonals in this way, their utterances contradict each other. The same effect can be obtained if we take them to disagree not about the meaning of *fun* but about whether Group A or Group B is the relevant set of individuals.

The technique of diagonalization is a useful one which we will have occasion to use later in this book. (See especially Sec. 7.5.) However, it is problematic to use it — or any analysis based on metalinguistic or metacontextual conflict — in order to account for faultless disagreement over taste. The reason is simply that metalinguistic disagreement presumes a shared language, and metacontextual disagreement presumes a shared context. But neither of these is required for faultless disagreement over matters of taste.

As we have already seen, two people may disagree even if they are geographically separated and unaware of each other's existence: If John, in Urbana, Illinois, asserts that licorice is tasty, and Mary, a stranger he has never met, in Santa Cruz, California, asserts that licorice is not tasty, their assertions contradict each other, even though they are not engaged in conversation with each other and do not presume each other to share a conversational context. Their disagreement cannot be over the nature of "the context."

Likewise, John might be a monolingual speaker of English, and Mary a monolingual speaker of German. If John says *Licorice is tasty* and Mary says *Lakritze ist nicht lecker*, they are disagreeing — contradicting each other, in fact — even though they are not in conflict over the meanings of any English or German words.

In sum, their disagreement seems to be about licorice and whether or not it is tasty — not about the language, or who the relevant people are.

To summarize this chapter, whether we take words like *fun* and *tasty* to involve hidden indexicality, or quantification over the individuals for whom something might be fun or tasty; or take them to express completely objective properties, with no relativization, or as purely expressive, with no associated truth conditions; or analyze disagreements over taste as metalinguistic or metacontextual, we face significant challenges. None of these ideas provides a straightforward, unproblematic account of the semantics of sentences about matters of taste. I suggest it would be worth our while to consider an alternative approach.

Chapter 3 : Ordinary Indexicality

In this chapter, I develop an analysis of indexical expressions in English, to use as background to the relativist treatment of predicates of personal taste to be presented in Chapter 4. There will not be much in this chapter which is original or surprising, I think; nor will it serve directly to advance the arguments for relativism. The point, rather, is to have a more detailed theory of context-dependent expressions in place, with an explicit syntax and semantics, so that it will be clear how a relativist treatment of taste examples fits into the larger picture. The highly simplified treatment of indexicality already presented in Sec. 2.1. is inadequate to this purpose.

3.1. Syntactic assumptions

For the syntactic part of the analysis, I will use an *ad hoc* and simplified theory of syntax, but one which should be easy to follow and reasonably clear how to relate to more sophisticated syntactic frameworks. Each expression will be assigned to a set of syntactic categories, including a "major category" corresponding to the traditional part-of-speech classification, and potentially some minor categories or "features." To begin with, let us assume the following basic categories; though we will add to the list as the analysis proceeds:

(34) *Basic categories:*
   a. Major: N, V
   b. Minor: SG, PL, NOM, ACC, PROPER, PRO, 1P, 2P, 3P, MASC, FEM, NEUT, REFL, PAST, PRES, INF

Here N is the category of noun phrases, V of sentences (assumed here to be headed by the verb). SG is the category of singular nouns; PL of plural nouns; NOM of noun phrases with nominative case; ACC of noun phrases with accusative case; PROPER of proper names; PRO of personal pronouns; 1P, 2P and 3P of first-person, second-person, and third person noun phrases, respectively; MASC of masculine pronouns, FEM of feminine pronouns, NEUT of neuter pronouns; REFL of reflexive pronouns; PAST of past tense verb phrases, PRES of present tense verb phrases, and INF of infinitival clauses.

We will label each occurrence[28] of an expression of category N in φ with a positive integer which we will call its *referential index*, and each occurrence of an expression of category PRES or PAST with a positive integer which we will call its *temporal index*. These indices will be used to track binding and other anaphoric relations.[29]

An expression which takes arguments will be labeled with a derived category, using slash notation in the manner of Categorial Grammar:[30]

[28] *Occurrences* in this sense should not be confused with uses. The sentence *The man ate the sandwich* contains two occurrences of the word *the*, but of course there may be any number of different uses of this sentence, each of which contains two uses of *the*.

[29] Indices in this sense are simply syntactic labels, and must not be confused with indices in the sense of parameters relative to which truth values are assigned. This ambiguity in the term *index* is unfortunate, but both usages are too well entrenched to permit substitute terminology.

[30] Parentheses will often be dropped when no ambiguity results.

(35) *Derived categories:*
   a. If X is a category and Y is a category, then (X/Y) is a category.
   b. If X is a category and Y is a category, then (Y\X) is a category.

Here, (X/Y) is the category of an expression which may combine with an expression belonging to category Y on its right to form an expression belonging to category X, and (Y\X) is the category of an expression which may combine with an expression belonging to category Y on its left to form an expression belonging to category X. Formally, we state this as in (36), treating each category simply as the set of its members:

(36) a. If $\alpha \in (X/Y)$ and $\beta \in Y$, then $[\alpha\beta] \in X$.
   b. If $\alpha \in (Y\backslash X)$ and $\beta \in Y$, then $[\beta\alpha] \in X$.

We will later add some additional derived categories and combinatorial rules.

Here are some sample lexical items with category assignments. Homophonous items (for example singular and plural *you*) are distinguished with diacritics, such as subscripted *sg* or *pl*:

(37) For all positive integers *i*:
- $I_i \in$ N ∩ PRO ∩ 1P ∩ SG ∩ NOM
- $me_i \in$ N ∩ PRO ∩ 1P ∩ SG ∩ ACC
- $myself_i \in$ N ∩ PRO ∩ 1P ∩ SG ∩ ACC ∩ REFL
- $we_i \in$ N ∩ PRO ∩ 1P ∩ PL ∩ NOM
- $us_i \in$ N ∩ PRO ∩ 1P ∩ PL ∩ ACC
- $ourselves_i \in$ N ∩ PRO ∩ 1P ∩ PL ∩ ACC ∩ REFL
- $you_{sg_i} \in$ N ∩ PRO ∩ 2P ∩ SG ∩ NOM ∩ ACC
- $yourself_i \in$ N ∩ PRO ∩ 2P ∩ SG ∩ ACC ∩ REFL
- $you_{pl_i} \in$ N ∩ PRO ∩ 2P ∩ PL ∩ NOM ∩ ACC
- $yourselves_i \in$ N ∩ PRO ∩ 2P ∩ PL ∩ ACC ∩ REFL
- $he_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ NOM ∩ MASC
- $him_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ ACC ∩ MASC
- $himself_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ ACC ∩ MASC ∩ REFL
- $she_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ NOM ∩ FEM
- $her_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ ACC ∩ FEM
- $herself_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ ACC ∩ FEM ∩ REFL
- $it_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ NOM ∩ NEUT ∩ ACC
- $itself_i \in$ N ∩ PRO ∩ 3P ∩ SG ∩ ACC ∩ NEUT ∩ REFL
- $they_i \in$ N ∩ PRO ∩ 3P ∩ PL ∩ NOM
- $them_i \in$ N ∩ PRO ∩ 3P ∩ PL ∩ ACC
- $themselves_i \in$ N ∩ PRO ∩ 3P ∩ PL ∩ ACC ∩ REFL
- $John_i \in$ N ∩ PROPER ∩ 3P ∩ SG ∩ NOM ∩ ACC
- $Mary_i \in$ N ∩ PROPER ∩ 3P ∩ SG ∩ NOM ∩ ACC

- *The Giant Dipper*$_i$ ∈ N ∩ PROPER ∩ 3P ∩ SG ∩ NOM ∩ ACC
- *sleep*$_{inf}$ ∈ (ACC\(V∩INF))
- *sleep*$_{1sg_i}$ ∈ ((NOM∩1P∩SG)\V) ∩ PRES
- *sleep*$_{2sg_i}$ ∈ ((NOM∩2P∩SG)\V) ∩ PRES
- *sleeps*$_i$ ∈ ((NOM∩3P∩SG)\V) ∩ PRES
- *sleep*$_{pl_i}$ ∈ ((NOM∩PL)\V) ∩ PRES
- *slept*$_i$ ∈ (NOM\V) ∩ PAST
- *see*$_{inf}$ ∈ ((ACC\(V∩INF))/ACC)
- *see*$_{1sg_i}$ ∈ (((NOM∩1P∩SG)\V)/ACC) ∩ PRES
- *see*$_{2sg_i}$ ∈ (((NOM∩2P∩SG)\V)/ACC) ∩ PRES
- *sees*$_i$ ∈ (((NOM∩3P∩SG)\V)/ACC) ∩ PRES
- *see*$_{pl_i}$ ∈ (((NOM∩PL)\V)/ACC) ∩ PRES
- *saw*$_i$ ∈ ((NOM\V)/ACC) ∩ PAST
- *don't*$_{1sg_i}$ ∈ (((NOM∩1P∩SG)\V)/(ACC\(V∩INF)) ∩ PRES
- *don't*$_{2sg_i}$ ∈ (((NOM∩2P∩SG)\V)/(ACC\(V∩INF)) ∩ PRES
- *doesn't*$_i$ ∈ (((NOM∩3P∩SG)\V)/(ACC\(V∩INF)) ∩ PRES
- *don't*$_{pl_i}$ ∈ (((NOM∩PL)\V)/(ACC\(V∩INF)) ∩ PRES
- *didn't*$_i$ ∈ ((NOM\V)/(ACC\(V∩INF)) ∩ PAST
- *necessarily* ∈ V/V

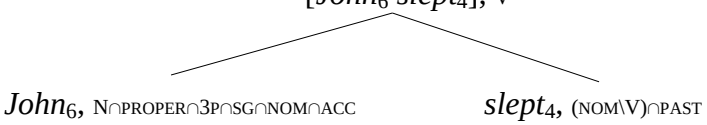
We consider a sentence to be well-formed only if the following conditions are met:[31]

(38) 1. If **α** is of category REFL, there is some **β** ≠ **α** c-commanding **α** in the minimal expression of category V containing **α**, such that REF(**α**) = REF(**β**).
2. If **α** is of category PRO but not of REFL, there is no **β** ≠ **α** c-commanding **α** in the minimal expression of category V containing **α**, such that REF(**α**) = REF(**β**).

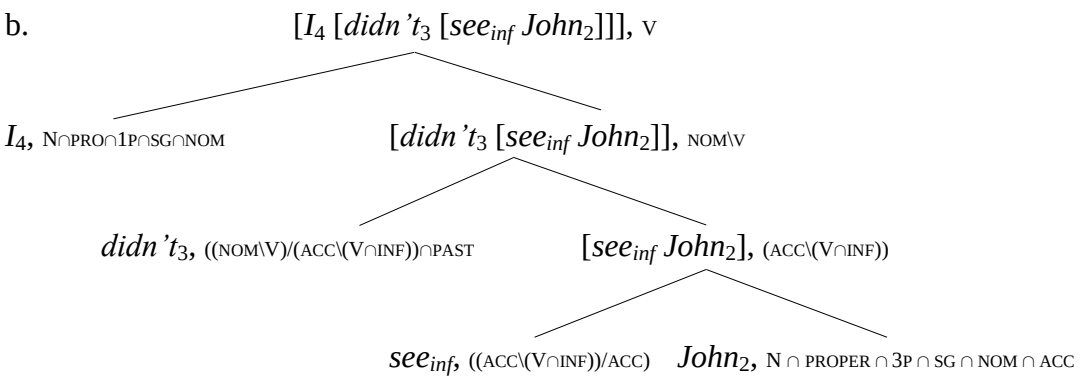
Agreement for gender and number of pronouns with their antecedents will be enforced semantically rather than by syntactic stipulation.

The following diagrams show syntactic structures of the sample sentences *John slept* and *I didn't see John*, in Montague-style tree notation:

(39) a.



[31] These conditions will be familiar to many readers as a versions of "Principle A" and "Principle B" of Chomsky's (1981) Binding Theory. We say that α c-commands β in φ iff φ contains some expression γ consisting (entirely) of α and δ, and δ contains β.



3.2. Pronouns, names and anaphora

Personal pronouns provide perhaps the most obvious examples of indexical expressions. The referent of a pronoun may vary from context to context, so we must regard the context as providing, in some way or another, a referent for each indexically interpreted pronoun. In general, the referents of indexical pronouns must be pragmatically relevant and "familiar" in the context — they must be things which the participants in the discourse regard as having some bearing on the matters under discussion, and presume each other to be acquainted with. Something may be familiar because it has been pointed out by the speaker or some other participant in the discourse, because it has been mentioned previously in the discourse, because it is visually or auditorially salient in the location where the discourse takes place, because it forms part of the general background knowledge the discourse participants presume each other to have by virtue of shared culture or experience, or perhaps for other reasons. We will not concern ourselves much here with the details of how an object becomes relevant or familiar, but simply assume that relative to a world of use *u*, each use α of an expression **α** is associated with a set $Relevant_{u,\alpha}$ of relevant objects, which has a subset $Familiar_{u,\alpha}$ of familiar objects.[32]

I will assume that over the course of a single use of an expression, nothing ceases to become relevant or familiar, and that the things which are relevant or familiar in the context of a complex expression are just those which are relevant or familiar in the contexts of its parts. That is:

(40) Where α has syntactic constituents $\alpha_1,\ldots,\alpha_n$ (and $1 \leq i,j \leq n$):
   a. If $time_{u,\alpha_i}$ precedes $time_{u,\alpha_j}$ then $Relevant_{u,\alpha_i} \subseteq Relevant_{u,\alpha_j}$ and $Familiar_{u,\alpha_i} \subseteq Familiar_{u,\alpha_j}$.
   b. $Relevant_{u,\alpha} = Relevant_{u,\alpha_1} \cup \ldots \cup Relevant_{u,\alpha_n}$ and $Familiar_{u,\alpha} = Familiar_{u,\alpha_1} \cup \ldots \cup Familiar_{u,\alpha_n}$.

The assignment of referents to pronouns is complicated by the fact that pronouns may be used anaphorically, so that their interpretation depends in some way on the interpretation of the

[32] I understand "object" here broadly enough to include individuals, groups, times, etc.

antecedent. Rather than treating demonstrative and anaphoric uses of pronouns completely independently, it would be preferable to give a unified analysis.

We are assuming (in common with a great deal of syntactic work) that each pronoun is labeled syntactically with a referential index. To represent the referential intentions of the speaker, let us also us assume that the pragmatic context of a use α (in a world *u*) will normally provide a function $Ref_{u,\alpha}$ mapping positive integers from some finite set onto elements of $Familiar_{u,\alpha}$. We regard $Ref_{u,\alpha}(i)$ as the denotation of any constituent of α bearing referential index *i*. For example, if $Ref_{u,\varphi}(1)$ = John, $Ref_{u,\varphi}(2)$ = Mary and $Ref_{u,\varphi}(3)$ = Bill, then in φ (as φ occurs in *u*) John is the denotation of me_1, Mary is the denotation of her_2, and Bill is the denotation of $himself_3$. I assume that such values remain fixed through the course of a single use of an expression. That is, (construing functions as sets of pairs):

(41) Where α has syntactic constituents $\alpha_1, \ldots, \alpha_n$ (and $1 \leq i,j \leq n$):
   a. If $time_{u,\alpha_i}$ precedes $time_{u,\alpha_j}$ then $Ref_{u,\alpha_i} \subseteq Ref_{u,\alpha_j}$.
   b. $Ref_{u,\alpha} = Ref_{u,\alpha_1} \cup \ldots \cup Ref_{u,\alpha_n}$

Now we need to give rules assigning denotations to pronouns and other lexical items. I will do this in two steps. First we define a function *Lex* assigning each use of lexical item a "customary" denotation relative to each pair of possible worlds. Then for each use of an expression α and pair of worlds *u*, *w* we define the denotation of α in *u* relative to *w*, notated $[\![\alpha]\!]^{u,w}$, based ultimately on what *Lex* assigns to its lexical constituents. If **α** is a lexical item and α is an *extensional* use of **α** (that is, if α is in an extensional context), then $[\![\alpha]\!]^{u,w}$ will be identical to *Lex*(α, *u*, *w*), as stipulated in (42); but if α is intensional, they may differ. If **α** is syntactically complex rather than a lexical item, then *Lex*(α, *u*, *w*) will not be defined at all.

(42) For all lexical items **α** and worlds *u*, *w*: if USE(α, **α**) and α is extensional, then $[\![\alpha]\!]^{u,w}$ = *Lex*(α, *u*, *w*).

As yet we do not have a definition of which uses are extensional, but let us delay that for a little bit. All our examples until Section 3.7. will involve extensional uses only.

Next we stipulate that pronouns denote in accordance with the assignment provided by the context, as long as person, number and gender restrictions are respected:

(43) For any use α of a pronoun $\boldsymbol{\alpha_i}$, and any worlds *u*, *w*: *Lex*(α, *u*, *w*) = $Ref_{u,\alpha}(i)$, provided that $Ref_{u,\alpha}$ is a *feature-respecting assignment* relative to *u*, α (undefined otherwise).

(44) Where α is a use of some expression $\boldsymbol{\alpha_i}$ and *r* is any function mapping integers into $Familiar_{u,\alpha}$, we say that *r* is a feature-respecting assignment relative to *u*, α iff the following conditions are met:

   a. If $\boldsymbol{\alpha_i}$ is of category 1P, then *r*(*i*) includes $speaker_{u,\alpha}$;
   b. If $\boldsymbol{\alpha_i}$ is of category 2P, then *r*(*i*) includes $addr_{u,\alpha}$;
   c. If $\boldsymbol{\alpha_i}$ is of category SG, then *r*(*i*) is an individual;

d. If $\boldsymbol{\alpha_i}$ is of category PL, then $r(i)$ is a group;[33]
e. If $\boldsymbol{\alpha_i}$ is of category MASC, then $r(i)$ is male in u;[34]
f. If $\boldsymbol{\alpha_i}$ is of category FEM, then $r(i)$ is female in u;
g. If $\boldsymbol{\alpha_i}$ is of category NEUT, then $r(i)$ is sexless in u.

If a pronoun is indexed so that its person, number or gender restrictions do not match the properties of the intended referent, the denotation of the pronoun is undefined. We may consider this a case of presupposition failure, leaving truth-valueless the sentence in which the pronoun appears. Of course it is possible for a speaker to convey meaning in some sense even if he or she uses the wrong pronoun, but I take it this is because of the availability of pragmatic repair strategies, not because the grammar of the language actually gives a defined interpretation in such cases.

The rule in (43) will assign denotations to pronouns when used in a purely demonstrative fashion, without linguistic antecedents. But it can also be used to account for simple cases of anaphora. We assume antecedence relations are shown in standard fashion by syntactic co-indexation. If the antecedent of one pronoun is another pronoun, we need say nothing more; our rules already enforce coreference. Thus in (45), *she* must denote the same individual as *her*, etc.:

(45) *John*$_1$ *saw*$_3$ *her*$_2$. *She*$_2$ *slept*$_4$.

If the antecedent is a proper name, we may account for the coreference in much the same way. We label names with positive integers just as we do with pronouns. The index on a name will be used to keep track of which individual the name is used to refer to, and any pronoun with the same index will refer to the same individual. This requires some rules to fix the denotation of a name to the individual assigned to its index in context:

(46) For any worlds u, w:
a. If USE(α, *John*$_i$) then *Lex*(α, u, w) = $Ref_{u,\alpha}(i)$, provided that $Ref_{u,\alpha}(i)$ is named *John* in u;
b. If USE(α, *Mary*$_i$) then *Lex*(α, u, w) = $Ref_{u,\alpha}(i)$, provided that $Ref_{u,\alpha}(i)$ is named *Mary* in u;
c. If USE(α, *The Giant Dipper*$_i$) then *Lex*(α, u, w) = $Ref_{u,\alpha}(i)$, provided that $Ref_{u,\alpha}(i)$ is named *The Giant Dipper* in u.
(Undefined otherwise)

Note that these rules do not make it part of the *content* of the names *John*, *Mary*, and *The Giant Dipper* that their referents[35] are named *John*, *Mary* and *The Giant Dipper*; the content of a given use

[33] I assume for simplicity and concreteness here that plural pronouns denote groups. Ultimately, it may be more desirable to adopt a system in which plurality is property of the denotation relation itself, rather than the object denoted, as in Boolos (1998), Schein (1993) and related work.

[34] It is perhaps debatable whether one's sex can vary from world to world; but if it can, it seems to be the "context world" u which is used in determining the gender of the pronoun, not the "modal evaluation world" w. It seems more natural (at least to my ear) to say *If John were female, he would wear a dress*, for example, than *If John were female, she would wear a dress*. I set aside here many complex and interesting questions which come up when we recognize that genetic sex, gender identity, and gender presentation all sometimes diverge from one another and that some of these may shift over time. Thanks to Lori Moon for discussion. If we construe "male" and "female" in (43) to refer to impermanent characteristics, we would need add a temporal index to pronouns and rephrase clauses e. and f. to include reference to the value of this index.

[35] *The Giant Dipper* is the name of a roller coaster on the Beach Boardwalk in Santa Cruz, California.

of *John* in a world *u* will denote the same individual in all possible worlds *w*, even though that individual might be named differently in the different worlds, and likewise for *Mary* and *The Giant Dipper*. Obviously, a great deal might be said to clarify what it means for someone to be named *John*, or named *Mary*, etc.; here we might appeal to a causal-chain theory in the manner of Kripke (1980) or whatever other theory of names seems appropriate and compatible with our other rules; (46) is not intended as an alternative to such theories.[36]

Our rules will now correctly assign *she* in (47) the same referent as *Mary*:

(47) *John$_1$ saw$_3$ Mary$_2$. She$_2$ slept$_4$.*

3.3. Common nouns, quantification and binding

Quantification is of interest to the study of indexicality in at least a couple of ways. First, many indexical expressions, including paradigmatic examples like personal pronouns, can be bound by quantifiers. Any adequate theory of indexicality, therefore, must clarify the connection between indexicality and bindability, and suggest a plausible reason why the same expressions so often seem to exhibit both. Second, quantification is normally over a restricted domain, determined at least partly on a pragmatic basis, and this pragmatic effect on interpretation is often analyzed as a form of indexicality. In this section, we limit attention to the first of these issues, setting the second aside until Sec. 5.5.

One initially plausible explanation for the connection between indexicality and binding is that variable binding involves quantification over contexts. We could regard each context *c* as providing an assignment of referents Ref_c (analogous to our current $Ref_{u,\alpha}$); we might analyze *Every man sees Mary* as consisting syntactically of *Every man* and something like x_1 *sees Mary*, and as true in context *c* iff x_1 *sees Mary* is true in every context *c* ′ such that $Ref_{c'}(1)$ is a man and which is otherwise just like *c*. This is essentially the technique of Montague (1970), which models pragmatic contexts directly as assignments of values to variables.

However, it is not in general possible to alter just one element of a context, keeping everything else the same, and have the result also be a possible context — a point argued forcefully by Lewis (1980). If, for example, we are in a context *c* in which John is $Ref_c(1)$, having been made salient by virtue of my pointing at him, there can be no context *c* ′ exactly like *c* except that Bill is $Ref_{c'}(1)$, since in *c* I am pointing at John, and Bill might not be salient at all. There might be some other possible context *c* ′ such that Bill is $Ref_{c'}(1)$, but this would have to be a context in which I am pointing at Bill (or in which he is made salient in some other way) — which involves changing more features of *c* than just $Ref_c(1)$: there must be some difference in what direction I am pointing, or in who is present in that direction, etc., and these differences amount to a requirement that the world and/or time of *c* ′ is different from that of *c*. But ordinary quantification over individuals does not seem to automatically involve simultaneous quantification over worlds and times.

Lewis' solution to this problem is to sharply distinguish contexts from (tuples of) index values. Operators are interpreted so that the denotation of a complex expression consisting of an operator and its scope is dependent on the denotations of that scope relative to various other values

[36] Barwise and Perry (1983) also stress that the denotation of a name must be someone named with that name.

for the indices — but these index values need not represent a possible context of use at all. Adapting this idea to our current problem, we might treat reference assignments not just as elements of the context, used in assigning contents to sentences, but also as indices relative to which sentence contents are assigned truth values. Then we could analyze *Every man sees Mary* as expressing a content Φ which is true in *w* iff $[\![\Phi]\!]^{w,Ref}$ = **truth** for every function *Ref* such that *Ref*(1) is a man and which is otherwise just like Ref_c.

At least a couple of objections can be raised against this approach. First, it makes the assignment of values to variables a parameter relative to which sentence *contents* are assigned truth values. This renders content neutral with respect to the referents of pronouns; pronouns do not get their values in the assignment of contents to linguistic expressions, but in the assignment of denotations to contents. If we were to treat indexical expressions in general this way, there would be no character/content distinction at all; the two notions would collapse. Yet we have already seen substantial motivation for a notion of content distinct from character in characterizing the distinction between necessary and contingent truth, in characterizing contradiction and other logical properties, and in modeling the objects of assertion and other speech acts.

Second, this approach violates our principle of homomorphic denotation assignment (9). This is partly for the trivial reason that (9) was formulated under the assumption that the only index involved in the assignment of denotations to contents was a possible world index, but also for the more substantial reason that it makes the truth value of a sentence (relative to some series of index values) dependent not on the denotations of its parts (relative to those same index values), but instead dependent on the denotations of its parts relative to a whole series of additional index values. Specifically, the truth value of a sentence in a given context *c* depends not just on the denotations of its parts in their respective contexts, but also on their denotations relative to functions which disagree with Ref_c in specific ways. This technique undermines either the principle that truth depends on what one talks about, or the identification of an expression's denotation relative to *c* with what that expression is used to talk about in *c* (in the relevant sense of "talk about").

This problem is easily solved if we are willing to regard the denotations of linguistic expressions as sometimes a little more abstract than what we intuitively think of as their referents — specifically, if we regard them sometimes as functions from assignment functions to intuitive referents, building on the technique in Montague (1968).

Specifically, for any use α and world *u*, let us use the set of (finite) extensions of $Ref_{u,\alpha}$ as assignments of values to variables. If α contains a free variable, it will denote a function whose domain is the set of all finite extensions of $Ref_{u,\alpha}$.[37]

In the cases where the referential and temporal indices (if any) on a use α are in the domain of $Ref_{u,\alpha}$, we do not employ this technique, but continue as before. For example, if USE(α, he_1) and $Ref_{u,\alpha}(1)$ = John, then $[\![\alpha]\!]^{u,w}$ = John. But if the index on α is outside the domain of $Ref_{u,\alpha}$, its denotation will be a function mapping assignments onto a variety of different values: Suppose USE(α, he_4) and $Ref_{u,\alpha}$ is the function which maps 1 onto John, 2 onto Mary and 3 onto Bill (and does not provide any other mappings). Let *r* be the function just like $Ref_{u,\alpha}$ except that it also maps 4 onto

[37] An *extension* of a function *f* is a function which is a superset of *f*, regarded as a set of ordered pairs. The point here is to give a semantics for variable binding that respects the principle of homomorphic interpretation. While I regard this as an important theoretical principle with a solid intuitive basis, readers who find that basis less compelling than intuition that *John* denotes John himself rather than a function mapping assignment functions onto John are welcome to recast the analysis using a more conventional semantics for binding.

Fred, and let *r′* be the function just like $Ref_{u,\alpha}$ except that it also maps 4 onto Ralph. Then $[\![\alpha]\!]^{u,w}$ will be a function mapping *r* onto Fred, mapping *r′* onto Ralph, etc. — but undefined for $Ref_{u,\alpha}$, since $Ref_{u,\alpha}$ does not provide a mapping for 4.

To get these effects we now replace our pronoun rule in (43) with the following, where '$Ref_{u,\alpha} \preccurlyeq r$' means that *r* is a finite extension of $Ref_{u,\alpha}$:[38]

(48) For any use α of a pronoun $\boldsymbol{\alpha_i}$, and any worlds *u*, *w*:

a. If $Ref_{u,\alpha}(i)$ is defined and $Ref_{u,\alpha}$ is a feature-respecting assignment with respect to *u*, α, then $Lex(\alpha, u, w) = Ref_{u,\alpha}(i)$.

b. If $Ref_{u,\alpha}(i)$ is not defined, then $Lex(\alpha, u, w) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r$ and *r* is a feature-respecting assignment with respect to *u*, α . $r(i)]$

For certain expressions, including proper names and first and second person pronouns, we stipulate that their indices must be included in the domain of $Ref_{u,\alpha}$, so their denotations will always be individuals or groups, not functions from extensions of $Ref_{u,\alpha}$ to individuals or groups.[39]

Personal pronouns in general in English cannot be left as free variables; they must either be indexical or anaphoric. For this reason, we stipulate the following constraint, which may be understood as a kind of "familiarity" condition:

(49) For any use φ of a sentence **φ**: If **φ** contains a pronoun **α** with referential index *i*, and there is no **β** ≠ **α** c-commanding **α** in **φ** also with referential index *i*, φ is felicitous in *w* only if *i* is in the domain of $Ref_{u,\varphi}$.

Because we are now treating pronouns as sometimes denoting individuals or groups, and sometimes denoting functions from extensions of $Ref_{u,\alpha}$ to individuals or groups, we must formulate our rules to interpret complex phrases so that they take both possibilities into account. We distinguish four cases:

(50) Where γ is a use of a complex expression consisting of α and β:

- Case A: $[\![\alpha]\!]^{u,w}$ is a function and $[\![\beta]\!]^{u,w}$ is a possible argument of $[\![\alpha]\!]^{u,w}$;
- Case B: For every *r* such that $Ref_{u,\gamma} \preccurlyeq r$, if $[\![\alpha]\!]^{u,w}(r)$ is defined, then $[\![\alpha]\!]^{u,w}(r)$ is a function and $[\![\beta]\!]^{u,w}$ is a possible argument of that function;
- Case C: For every *r* such that $Ref_{u,\gamma} \preccurlyeq r$, if $[\![\beta]\!]^{u,w}(r)$ is defined, then $[\![\alpha]\!]^{u,w}$ is a function and $[\![\beta]\!]^{u,w}(r)$ is a possible argument of that function;

[38] I use the λ-notation of Heim and Kratzer (1998), p. 37, in which the portion between the colon and the dot indicates the domain of the function and the part after the dot specifies what value is assigned to each argument in this domain.

[39] This renders these expressions unbindable — a generalization which is not quite right, since first and second person pronouns can be bound by *only* (Kratzer (2009)): *Only I finished my paper on time*. I believe the present system can be adapted to account for such examples, but to try to do so here would take us too far afield.

- Case D: For every r such that $Ref_{u,\gamma} \preccurlyeq r$, if $[\![\alpha]\!]^{u,w}(r)$ and $[\![\beta]\!]^{u,w}(r)$ are both defined, then $[\![\alpha]\!]^{u,w}(r)$ is function and $[\![\beta]\!]^{u,w}(r)$ is a possible argument of that function.

Interpretation proceeds as follows:

(51) In Case A, $[\![\gamma]\!]^{u,w} = [\![\alpha]\!]^{u,w}([\![\beta]\!]^{u,w})$.
In Case B, $[\![\gamma]\!]^{u,w} = [\lambda r : Ref_{u,\gamma} \preccurlyeq r \text{ and } [\![\alpha]\!]^{u,w}(r) \text{ is defined} . [\![\alpha]\!]^{u,w}(r)([\![\beta]\!]^{u,w})]$.
In Case C, $[\![\gamma]\!]^{u,w} = [\lambda r : Ref_{u,\gamma} \preccurlyeq r \text{ and } [\![\beta]\!]^{u,w}(r) \text{ is defined} . [\![\alpha]\!]^{u,w}([\![\beta]\!]^{u,w}(r))]$.
In Case D, $[\![\gamma]\!]^{u,w} = [\lambda x : Ref_{u,\gamma} \preccurlyeq r \text{ and } [\![\alpha]\!]^{u,w}(r) \text{ and } [\![\beta]\!]^{u,w}(r) \text{ are both defined} . [\![\alpha]\!]^{u,w}(r)([\![\beta]\!]^{u,w}(r))]$.

These rules just do ordinary function-argument application on what would conventionally be the denotations of the constituents, but keep open the "argument slot" for assignment functions in expressions that contain free variables.

We now need to add common nouns to our lexicon:

(52) For all positive integers i, j:
- $man_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{SG}$
- $men_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{PL}$
- $woman_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{SG}$
- $women_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{PL}$
- $puppy_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{SG}$
- $puppies_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{PL}$
- $berry_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{SG}$
- $berries_{i,j} \in \text{N} \cap \text{COMMON} \cap \text{PL}$

Common nouns are labeled with two indices, instead of one: a referential index and a temporal index. By convention we write the referential index first and the temporal index afterwards.

Common nouns are most commonly treated as predicates, hence as denoting sets of individuals (or sets of ordered pairs, in the case of relational nouns like *brother* or *side*). This approach obscures both the name-like nature of common nouns, and the distinction — which is quite sharp in English and many other languages — between nouns on the one hand and clear examples of predicates such as verbs on the other. I will therefore adopt a slightly different approach here, which builds on the semantics for proper names and pronouns we already have in place. The analysis can be easily reformulated to use a more conventional semantics for common nouns if desired.

We are assuming that common nouns, like pronouns and proper names, are syntactically labeled with a referential index. In English, though perhaps not in article-less languages like Latin, the referential index on a common noun should always be "novel" — that is, outside the domain of $Ref_{u,\alpha}$.[40]

[40] Temporal indices are not subject to the novelty condition given in (53).

(53) For any use α of a common noun **α**, α is felicitous in *u* if the referential index on α is outside the domain of $Ref_{u,\alpha}$.

Much as we did with pronouns, we now let common nouns denote functions from extensions of $Ref_{u,\alpha}$ to individuals, so that they function like (restricted) variables:

(54) For any worlds *u*, *w*:
If USE(α, $puppy_{i,j}$), then $Lex(\alpha, u, w) = \lambda r : Ref_{u,\alpha} \preccurlyeq r$ and $r(j)$ is a time and $r(i)$ is a puppy at $r(j)$ in w . $r(i)$

For example, if α is a use of $puppy_{4,3}$ and the domain of $Ref_{u,\alpha}$ is {1, 2, 3} and $Ref_{u,\alpha}(3)$ = July 25, 2011, and *r* is a function like $Ref_{u,\alpha}$ except that $r(4)$ = Fido, then $Lex(\alpha, u, w)$ will be a function which maps *r* onto Fido if, on July 25, 2011, Fido is a puppy in *w*. Similarly if *r′* is like $Ref_{u,\alpha}$ except that $r'(4)$ = Rover, then $Lex(\alpha, u, w)$ will map *r′* onto Rover if on July 25, 2011, Rover is a puppy in *w*, etc. But if Fido is not a puppy in *w* on July 25, 2011, then $Lex(\alpha, u, w)(r)$ is undefined, and if Rover is not a puppy in *w* on July 25, 2011, then $Lex(\alpha, u, w)(r')$ is undefined. The range of $Lex(\alpha, u, w)$ thus corresponds to the extension of *puppy* in a more traditional semantics.

In formulating the semantics for determiners, we will be using some higher-order functions. It will be useful to have a system of *types*, so that we may indicate succinctly the domain and co-domain of a function. For this purpose we use a variant of the standard Montague-style system:[41]

(55) a. e is a type;
b. t is a type;
c. i is a type;
d. If τ is a type then ⟨s, τ⟩ is a type;
e. If τ is a type then ⟨r, τ⟩ is a type;
f. For all τ, υ: if τ is a type and υ is a type, then ⟨τ, υ⟩ is a type.

Each type will be associated with a specific *domain*, defined as follows:

(56) a. $\mathbf{D}_e$ = {*x* | *x* is a group or individual};
b. $\mathbf{D}_t$ = {**truth**, **falsity**};
c. $\mathbf{D}_i$ = {*t* | *t* is a time};
d. If τ is a type, then $\mathbf{D}_{\langle s,\tau\rangle} = \mathbf{D}_\tau^{\{w \mid w \text{ is a possible world}\}}$ (the set of all functions from possible worlds to $\mathbf{D}_\tau$);
e. If τ is a type, then $\mathbf{D}_{\langle r,\tau\rangle} = \mathbf{D}_\tau^{\{r \mid r \text{ is a partial function from positive integers to } \mathbf{D}_e \cup \mathbf{D}_i\}}$ (the set of all functions from assignments of values to variables to $\mathbf{D}_\tau$);
f. If τ is a type and υ is a type, then $\mathbf{D}_{\langle\tau,\upsilon\rangle} = (\mathbf{D}_\upsilon)^{\mathbf{D}_\tau}$ (the set of all functions from $\mathbf{D}_\tau$ to $\mathbf{D}_\upsilon$).

We turn next to determiners, beginning with the indefinite article *a*. It has been clear since Lewis (1975) that indefinite noun phrases must be analyzed as bindable. Therefore, following Heim (1982) and Kamp (1981), I will analyze the existential force often expressed by sentences containing indefinite noun phrases not as coming from the internal semantics of the indefinite article, but from a

[41] See Montague (1970), Montague (1973), or textbooks such as Heim and Kratzer (1998) or Gamut (1991).

more general operation of "existential closure." The indefinite article itself will be treated as semantically trivial, expressing an identity function:

(57) a. $a \in (\text{N} \cap \text{SG} \cap \text{NOM} \cap \text{ACC})/(\text{COMMON} \cap \text{SG})$
b. If $\text{USE}(\alpha, a)$, then $Lex(\alpha, u, w) = \lambda x : x$ is an individual . x

Assuming α is an extensional use of *a puppy*$_{i,j}$, by (51) $[\![\alpha]\!]^{u,w} = [\lambda r : Ref_{u,\alpha} \preccurlyeq r$ and $r(i)$ is a puppy at $r(j)$ in w . $r(i)]$.

In other words, *a puppy*$_{i,j}$, like the bare noun *puppy*$_{i,j}$, denotes (relative to a world w) a function mapping any extension of the reference assignment for its context onto whatever that extension assigns to i, provided it is a puppy at time j in w.

Rather than following Heim and Kamp in treating existential closure as part of the definition of truth, I will assume it is expressed by a hidden operator, which I will notate as *E*. We assign *E* to syntactic category V/V, and interpret it as denoting a function of type $\langle\langle \text{r,t}\rangle,\text{t}\rangle$:[42]

(58) a. $E \in \text{V/V}$
b. $Lex(E, u, w) = \lambda f \in \mathbf{D}_{\langle \text{r,t}\rangle}$. $\exists r[Ref_{u,\alpha} \preccurlyeq r$ and $f(r) =$ **truth**$]$

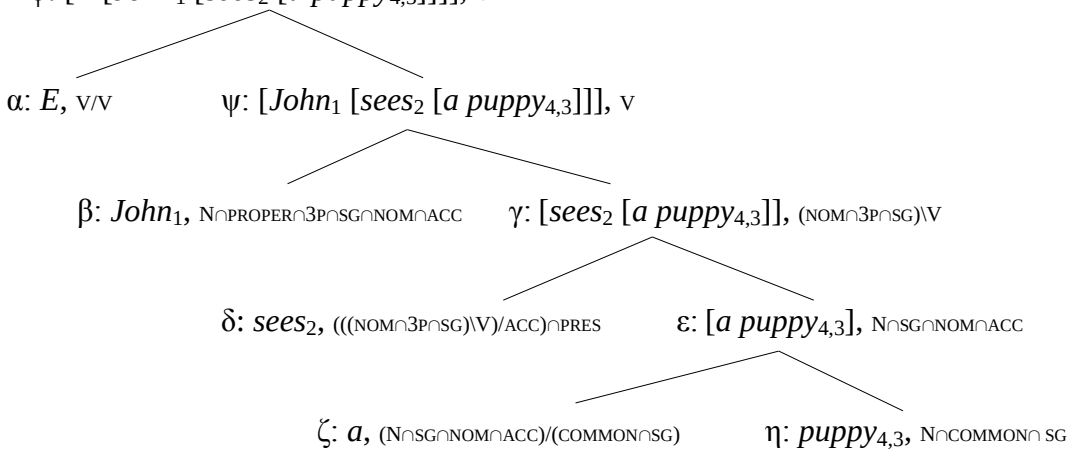
It may be useful to see an example. As yet we have not yet given a semantics for verbs — that will be the topic of the next section. But as a stopgap measure just to allow some illustration, we may assume the following:

(59) If use(α, *sees*$_i$) then $[\![\alpha]\!]^{u,w} = \lambda y \in \mathbf{D}_{\text{e}}$. $\lambda x \in \mathbf{D}_{\text{e}}$. x sees y at $Ref_{u,\alpha}(i)$ in w, where $Ref_{u,\alpha}(i) = time_{u,\alpha}$.

We may now show the syntactic structure for a use of the sentence [*E* [*John*$_1$ [*sees*$_2$ [*a puppy*$_{4,3}$]]]] as in (60). Here, and in subsequent tree diagrams for uses of sentences, I label each node with a Greek letter, to represent the use of the corresponding expression which forms a constituent of the sentence use represented by the whole tree. (This gives us a way to refer to the constituent uses in tree diagrams showing semantic derivations as in (61).)

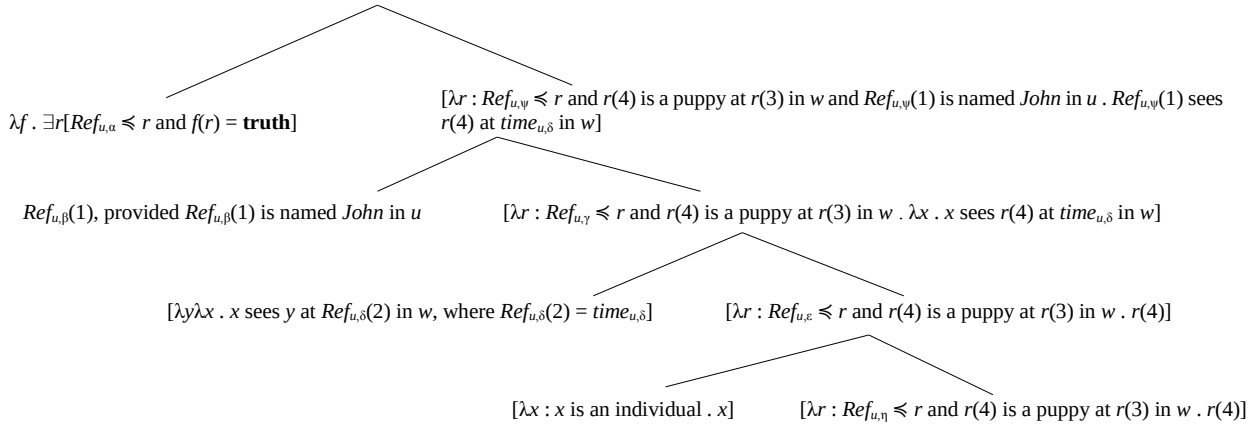
[42] Kamp's and Heim's analyses impose existential closure at the level of whole discourses rather than individual sentences, as I do here. I will not deal in any detail with intersentential anaphora to indefinite antecedents, so a definition at the level of sentences will serve our purposes adequately.

(60) φ: [*E* [*John*$_1$ [*sees*$_2$ [*a puppy*$_{4,3}$]]]], V



We may calculate the truth value of this sentence use φ relative to worlds *u*, *w* as follows:

(61) $\exists r$[Ref$_{u,\varphi} \preccurlyeq r$ and r(4) is a puppy at r(3) in w and Ref$_{u,\varphi}$(1) is named *John* in u and $Ref_{u,\varphi}$(1) sees r(4) at time$_{u,\delta}$ in w]



That is, a use of this sentence will be true iff there is at least one individual which is a puppy at the time picked out by index 3 in the context of the noun *puppy*, which John sees at the time of the context of the verb *sees*.[43]

Next, we add "traces" to our lexicon. These will serve as variables marking the position of a quantifier phrase.

[43] Assuming *Ref*$_{u,\eta}$ assigns a value to 3. If it not, then the sentence will be true iff there is at least one individual *x* and at least one time *t* such that *x* is a puppy at *t*, and John sees *x* at the time of the verb's context. Perhaps some readers will feel intuitively that the time when the relevant individual must meet the condition imposed by the common noun ought to be the same as the time picked out by the tense (in this case, the time of utterance). See Enç (1981) for a multitude of examples showing that this is not in general required.

(62) a. For all positive integers i: $e_i \in \text{N} \cap \text{3P} \cap \text{SG} \cap \text{NOM} \cap \text{ACC} \cap \text{TRACE}$
b. If USE(α, e_i) then $Lex(\alpha, u, w) = \lambda r : Ref_{u,\alpha} \preccurlyeq r.\ r(i)$

Traces are never interpreted demonstratively, so we stipulate that their referential indices must not be assigned values by $Ref_{u,\alpha}$, where α is any expression containing the trace:

(63) If USE(α, **α**) and **α** contains e_i and $e_i \in$ TRACE, then i is outside the domain of $Ref_{u,\varphi}$.

Traces will be replaced by quantifier phrases in the course of syntactic derivation. To accomplish this, we introduce a new sort of derived syntactic category:

(64) If X, Y and Z are categories, then (X|Z>Y) is a category.

(X|Z>Y) is the category of an expression which may combine with an expression of category Y to form an expression of category X by substituting for one of its constituents of category Z. More precisely:

(65) If $\alpha \in$ (X|Z>Y) and $\beta \in$ Y and $\gamma \in$ Z and β contains exactly one occurrence of γ, then $\beta' \in$ X, where β′ is the result of replacing γ in β with α.

For example, we will let [*every* $man_{7,1}$] be of category (V|TRACE>V). This will allow [*every* man_7] to combine with [e_7 $sleeps_2$] (an expression of category V) to form [[*every* $man_{7,1}$] $sleeps_2$], which will also be an expression of category V.

We consider a sentence to be well-formed only if every trace it contains has been replaced by a phrase whose head noun bears the same referential index as the trace.

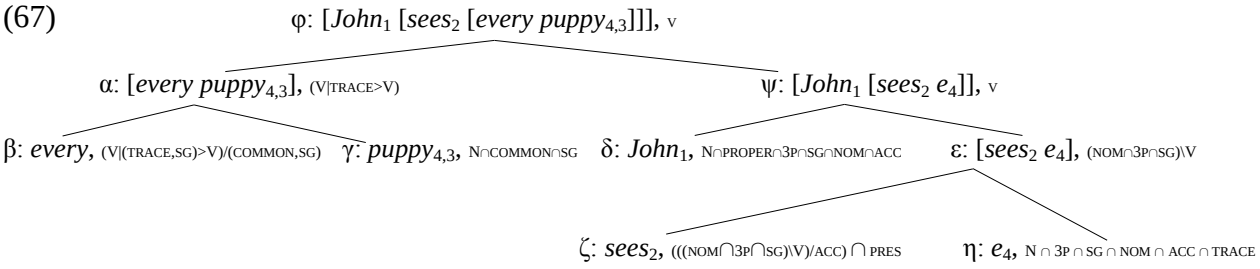
Determiners which bind the variable contributed by their common noun will have an interpretation which allows them to "see" all the values which the common noun denotation maps the various assignment functions onto. We may define *every* as follows:

(66) *every* $\in$ (V|(TRACE,SG)>V)/(COMMON,SG)
If USE(α, *every*) then $Lex(every, u, w) = \lambda f : f \in \mathbf{D}_{\langle r,e\rangle} . \lambda g : g \in \mathbf{D}_{\langle r,t\rangle}$. For all r, if $Ref_{u,\alpha} \preccurlyeq r$ and $f(r)$ is defined, there is at least one r' such that $r \preccurlyeq r'$ and $g(r')$ = **truth**

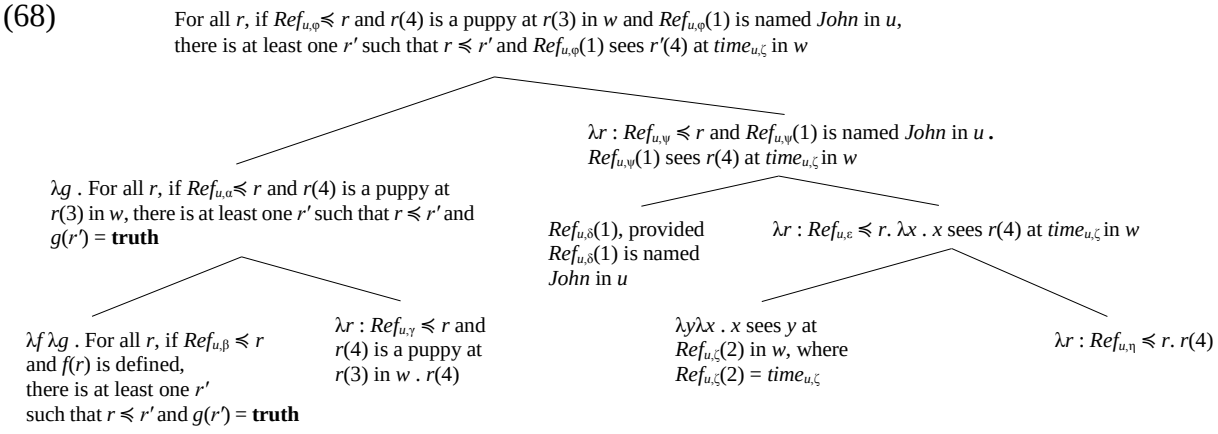
Assuming that α is an extensional use of *every* $puppy_{i,j}$, by (66) $[\![\alpha]\!]^{u,w} = [\lambda g : g \in \mathbf{D}_{\langle r,t\rangle}$. For all r, if $Ref_{u,\alpha} \preccurlyeq r$ and $r(j)$ is a time and $r(i)$ is a puppy at $r(j)$ in w, there is at least one r' such that $r \preccurlyeq r'$ and $g(r')$ = **truth**].

We may now derive the sentence [$John_1$ [$sees_2$ [*every* $puppy_{4,3}$]]] as follows:

(67)



(68)



To put it less formally, a use of *John sees every puppy* is true iff every way of extending its reference assignment so that it assigns a puppy to index 4 can be extended further to satisfy *John sees e*$_4$.[44]

Note that the temporal index 3 on *puppy* may be assigned a value by *Ref*$_{u,\varphi}$, or may not be. In the former case, the sentence quantifies over everything which is a puppy at some particular time provided by the context. In the latter, it quantifies over everything which is ever a puppy at any (relevant) time. Both these options seem available for the interpretation of the English sentence.

Of course quantifiers may also bind pronouns, as the reader should be able to confirm by working through examples such as *Every man*$_{2,1}$ *sees*$_1$ *himself*$_2$. By treating indexical pronouns as receiving their referents via *Ref*$_{u,\alpha}$ and analyzing variable binding by using extensions of *Ref*$_{u,\alpha}$ as assignments of values to variables, we give a single, unified account of both indexical and bound-variable pronouns, rather than treating them as separate, unrelated uses.

Other quantificational determiners such as *most*, *no*, etc. can be analyzed in much the same way as *every*. The definite determiner *the* may also be treated this way, but here I will instead assume a presuppositional analysis, under which definite noun phrases denote particular groups or

[44] The point of extending the reference a second time to satisfy the scope of the quantifier is to existentially bind any indefinites which might be free in that scope; for example, we should assign *Every man*$_{4,3}$ *sees*$_1$ *a puppy*$_{5,2}$ an interpretation in which *a puppy* is existentially bound inside the scope of *every man*. See Heim (1982) for this technique. It may also be that this sentence should be assigned a second interpretation in which *a puppy* takes scope over *every man*, but I forego accounting for this ambiguity here, in the interest of simplicity.

individuals. (Not much in the rest of the analysis depends on this choice.) We add singular and plural versions of *the* to our lexicon and define them as follows:[45]

(69) a. $the_{sg} \in (\text{N} \cap \text{SG} \cap \text{NOM} \cap \text{ACC})/(\text{COMMON} \cap \text{SG})$
b. $the_{pl} \in (\text{N} \cap \text{PL} \cap \text{NOM} \cap \text{ACC})/(\text{COMMON} \cap \text{PL})$

(70) a. If $\text{USE}(\alpha, the_{sg})$ then $Lex(\alpha, u, w) = \lambda f \in \mathbf{D}_{\langle \text{r,e} \rangle}$. There exists some x such that for all r, if $Ref_{u,\alpha} \preccurlyeq r$ and $f(r)$ is defined, then $x = f(r)$. $f(r')$, where r' is any arbitrary extension of $Ref_{u,\alpha}$ such that $f(r')$ is defined.

b. If $\text{USE}(\alpha, the_{pl})$ then $Lex(\alpha, u, w) = \lambda f \in \mathbf{D}_{\langle \text{r,e} \rangle}$. There exists some x such that for all r, if $Ref_{u,\alpha} \preccurlyeq r$ and $f(r)$ is defined, then x includes $f(r)$. $f(r')$, where r' is any arbitrary extension of $Ref_{u,\alpha}$ such that $f(r')$ is defined.

3.4. Tense

Almost all analyses of tense treat it as indicating that some time involved in interpreting the sentence is placed relative to the time of utterance[46] — before, concurrently, or after, depending on the tense, with some languages also distinguishing near vs. distant past, etc. There are two separate ways in which indexicality sometimes enters the analysis, both of which are controversial to some extent.

First and most commonly, tense is often treated as making indexical reference to the time of utterance: much as the content of an indexical pronoun like *we* depends on the speaker of the context, the content of a tense marker is analyzed as depending on the time of the context. Second, the time which the tense places before, concurrently or after the time of utterance is also sometimes treated as picked out indexically — typically as dependent on the referential intentions of the speaker. Let us consider these two appeals to indexicality in the analysis of tense separately, beginning with the claim that tense markers involve indexical reference to the time of utterance.

If tense markers make indexical reference to the time of utterance, then given all our other assumptions, the *content* of a tensed sentence will vary with the time of utterance. For example, when uttered at 2:46 p.m., July 31, 2011, the sentence *John sees Mary* will express the content that John sees Mary at 2:46 p.m., July 31, 2011; but when uttered at 1:00 a.m. August 1, 2011, it will express the content that John sees Mary at 1:00 a.m., August 1, 2011.

The main alternative to this idea is that sentence contents are "time neutral"; that is, that the content of a sentence like *John sees Mary* is the same at all times — just the content that John sees Mary. Of course the truth value of this sentence seems to vary with the time of utterance, so if we take the content to be the same at all times, the truth value of the content will have to vary as well. This implies that time is a parameter relative to which sentence contents have truth values. We would therefore have to revise our truth definitions:

[45] In a more detailed analysis, it would be desirable to give a single definition for *the*, covering both singular and plural uses. We distinguish singular and plural *the* here in order to allow our oversimple account of subject-verb agreement to work properly.

[46] Or some proxy for the time of utterance, as in special uses like the "historical present."

(71) Φ is true relative to *w*, *t* iff $[\![\Phi]\!]^{w,t}$ = **truth**
Φ is false relative to *w*, *t* iff $[\![\Phi]\!]^{w,t}$ = **falsity**

Of course all our other semantic rules would also be revised to assign denotations relative to times — but such revisions would not be hard to carry out. Indeed, this is essentially the technique underlying the standard semantics for tense logic, and implemented in Montague (1968; 1973), Lewis (1972), Kaplan (1989) and many other places.

However, a number of objections can be (and have been) given to this approach. For our purposes, the most important of these is that it greatly complicates the definition of contradiction. We do not, as yet, have a formal definition of contradiction in place, but the natural one if sentences are *not* time-neutral would be one like this:

(72) Φ contradicts Ψ iff there is no world *w* such that Φ is true relative to *w* and Ψ is true relative to *w*

Contradiction is defined here at the level of sentence contents. The definition reflects very directly the intuition that two contents contradict each other iff it is impossible for them to both be true.

If we were to adapt such a definition to a semantic theory in which sentence contents are time-neutral, we might try simply replacing "*w*" in the definition with "*w*, *t*":

(73) Φ contradicts Ψ iff there is no *w*, *t* such that Φ is true relative to *w*, *t* and Ψ is true relative to *w*, *t*

But this does not capture our intuitions about contradiction at all. For example, a sentence like *John does not see Mary* presumably expresses a content which is true relative to a given word *w* and time *t* iff the content expressed by *John sees Mary* is false relative to *w*, *t* — but then, the contents expressed by these two sentences will contradict each other. This is the result we want if the two sentences are uttered at the same time, but if one person says "John sees Mary" at 1:00, and another says "John does not see Mary" at 2:00, there is no contradiction. Nonetheless the definition in (73) is satisfied, because in this analysis the contents of these sentences do not change over time.

We might try reformulating the definition to correspond better to our intuitions, but it is not at all straightforward to come up with a formulation which defines contradiction as a relation between sentence contents and treats such contents as time-neutral, as readers may confirm for themselves.

I will therefore assume that tensed sentences are not to be treated as having time-neutral contents, and that the sensitivity they exhibit to the time of utterance is due to indexical reference to that time, rather than a time parameter relative to which sentence contents are assigned truth values. (Arguments against this view will be addressed in Section 7.5. below.)

The second indexical element which is sometimes — more controversially — claimed to be involved in the semantics of tense is the time which is placed before, concurrently, or after the speech time. In this kind of analysis, a past tense sentence like *John saw a dog* is taken to be about one particular past time, fixed by the referential intentions of the speaker; the sentence is true if John saw a dog at that particular time, false if he did not (even if there were other past times when he saw

a dog). This kind of approach is developed by Partee (1973), Enç (1981; 1986) and others. The alternative is to take tensed sentences as involving existential quantification over times, so that this sentence is true if there is at least one past time when John saw a dog, and false if there are no such times. This approach is exemplified in the standard semantics for Tense Logic, and in analyses which build on it such as Montague (1973) or Kaplan (1989).

The choice between these alternatives involves a complex set of issues and cannot be addressed in full detail here. However, I wish to give at least a brief defense of the latter, quantificational approach, and show how to implement it in our current system.

Two arguments that are frequently cited in favor of the indexical approach, both due essentially to Partee, are (1) that it more accurately captures the truth conditions of sentences in which tense appears with negation, and (2) it more easily accounts for the fact that tenses are bindable.

The issue with negation is easily seen in sentences like *I didn't turn off the stove*. Partee asks us to imagine someone uttering this sentence while driving down the highway after suddenly remembering that he or she left the house with the stove still burning. If the interpretation of the past tense involves existential quantification, there would seem to be two possibilities for its scope relative to negation — either the negation takes wide scope, or the quantifier does:

(74) a. $\sim\exists t$ [*t* is prior to speech time and I turn off the stove at *t*]
b. $\exists t$ [*t* is prior to speech time and ~ I turn off the stove at *t*]

But neither of these seems to be the correct truth conditions. (44)a. seems better paraphrased as "I have never turned off the stove," which is too strong, and (44)b. is true if there is any past time at all when the speaker did not turn off the stove, which is much too weak. If, however, we take the sentence to express a content which is true iff the speaker did not turn off the stove at the one particular time he or she is referring to, we seem to obtain the right interpretation.

However, this is only a weak argument, because it presumes the quantification is unrestricted. Quantification in natural language is virtually always over some pragmatically limited domain, left inexplicit in the sentence itself. If, at the beginning of a meeting, I say "Everyone is here; let's begin," my use of *everyone* quantifies over everyone who is expected to be at the meeting, not over everyone in the world. By the same token, we might take *I didn't turn off the stove* to quantify not over all previous times, but only those just prior to the speaker leaving the house. Letting negation take wide scope with respect to the quantifier then gives perfectly appropriate truth conditions for this sentence.

A defender of the indexical analysis might point out that quantification with *never* is also presumably restricted to a pragmatically established domain, but that *I didn't turn off the stove* seems weaker than *I never turned off the stove*. But nothing in the appeal to pragmatically restricted domains of quantification requires that they must be the same in these two sentences. It is quite conceivable that different quantifiers serve as signals of different degrees of pragmatic "domain narrowing" — an idea that has been argued for independently in the analysis of negative polarity items (Kadmon and Landman 1993).[47]

[47] It is perhaps worth mentioning here the possibility of analyzing *never* as *not+ever*, where *ever*, as a negative polarity item, would be independently expected to signal a widened domain of quantification in a Kadmon-and-Landman-style analysis.

The argument based on bindability is also quite weak. Here, the concern is with sentences like *Whenever Susan arrived, John left*, which seems to involve universal quantification over past times when Susan arrived. If the past tense in *Susan arrived* expressed existential quantification over past times when Susan arrived, there would seem to be no appropriate variable left free for universal binding by *whenever*; but if we treat the past tense as similar to a pronoun — available for binding as a variable, but referring indexically if no quantifier binds it — we obtain more reasonable results. However, this argument shows only that the tense marker itself does not bind off the variable and make it unavailable for further quantification; it does not show that the free-standing sentence *Susan arrived* does not involve existential quantification over past times.

An analogy may be drawn here to indefinite descriptions. There can be little doubt that a sentence like *An elephant arrived* involves existential quantification over elephants. Yet indefinites may also be bound by quantificational adverbs: *An elephant always has 56 chromosomes* expresses universal quantification over elephants. We can obtain the correct truth conditions for both sentences very straightforwardly using well-known techniques: either we attribute the existential force in *An elephant arrived* to a general operation of existential closure rather than the internal semantics of *an elephant*, as developed by Kamp (1981) and Heim (1982) and adopted in Sec. 3.3. above, or we move to a "dynamic" semantics in which quantified variables are still available for further quantification, as in Groenendijk and Stokhof (1989; 1991), Dekker (1993). In either approach, the existential force of *An elephant arrived* presents no obstacle to treating *An elephant always has 56 chromosomes* as involving universal quantification over elephants. In the same way, we may analyze *Susan arrived* as expressing existential quantification over past times, either because of a general operation of existential closure or by treating the tense as a dynamic quantifier, without threatening our ability to analyze *Whenever Susan arrived, John left* as expressing universal quantification over times.

Although these arguments against a quantificational analysis of tense are easily overcome, it still remains to be shown that there is any advantage to such an analysis over one in which a past or future tense makes indexical reference to a specific past or future time, based on the referential intentions of the speaker. I believe there is an advantage to the quantificational analysis, which may be seen in the ability of tense to serve as the antecedent to a *Wh*-phrase in "sluicing" and "pre-sluicing" constructions.[48]

Sluicing is the grammatical construction illustrated in (75):

(75) a. John is dating someone, but I don't know who.
b. A book is in this box, and John will guess which one.
c. Something smells bad. I wonder what?
d. Mary left her wallet someplace, but can't remember where.

Sluicing sentences are usually analyzed as derived via syntactic ellipsis from corresponding *pre-sluicing* constructions as in (76):

[48] A similar argument is given briefly in Ogihara (1995:667). The general idea of appealing to sluicing and pre-sluicing constructions as a means of distinguishing specific reference from existential quantification owes a great deal to Fillmore (1986), though Fillmore applies this idea to implicit arguments rather than tense.

(76) a. John is dating someone, but I don't know who he is dating.
b. A book is in this box, and John will guess which one is in this box.
c. Something smells bad. I wonder what smells bad?
d. Mary left her wallet someplace, but can't remember where she left her wallet.

In the pre-sluicing constructions a subordinate clause appears which is structurally parallel to another clause earlier in the discourse. Let us this earlier clause the *antecedent clause*. The subordinate clause contains a *Wh*-phrase (that is, an interrogative pronoun or phrase headed by an interrogative pronoun), which typically corresponds directly to an indefinite description in the antecedent clause. For example in (76)a., the interrogative pronoun *who* in *who he is dating* corresponds to *someone* in *John is dating someone*; in (76)b. the *Wh*-phrase *which one* in *which one is in the box* corresponds to *a book* in *A book is in the box*; etc. In the sluicing constructions, most of the subordinate clause is deleted, leaving just the *Wh*-phrase.

In some cases, it is possible to construct a sluicing or pre-sluicing sentence in which there is no overt indefinite description in the antecedent clause. In such cases, the antecedent clause may nonetheless be analyzed as involving covert existential quantification, perhaps because of a syntactic "sprouting" operation as in Chung, Ladusaw, and McCloskey (1995). This existential quantification corresponds to the *Wh*-phrase in much the same way as existential quantification which is overtly expressed by an indefinite description. In (77)a., the clause *John ate dinner* may be analyzed as implying that there is someone with whom he ate; *with whom* corresponds to this implicit existential quantification in much the same way as it corresponds to the overt quantification expressed by *someone* in (77)b. If one overtly denies the existential quantification, as in (77)c., the result is semantically incoherent:

(77) a. John ate dinner, but I don't know with whom.
b. John ate dinner with someone, but I don't know with whom.
c. *??*John didn't eat dinner with anyone, but I don't know with whom.

Sluicing or pre-sluicing sentences may also be constructed in which the phrase corresponding to the *Wh*-item is definite rather than indefinite (Chung et al. 1995; Dayal and Schwarzschild 2010):

(78) He announced that he had eaten the asparagus. We didn't know which asparagus.

However, in many cases one finds that a sluicing or pre-sluicing sentence in which the antecedent clause contains an indefinite description becomes anomalous if a definite description, demonstrative, personal pronoun, or proper name is substituted for the indefinite:

(79) a. John is dating someone, but I don't know who.
b. *??*John is dating the woman in the blue dress, but I don't know which one.
c. *??*John is dating that woman, but I don't know which one.
d. *??*John is dating her, but I don't know who.
e. *??*John is dating Mary, but I don't know who.

As argued by Dayal and Schwarzschild (2010), the anomaly in these cases is due to a conflict between the assertion made in the second clause — namely, that the speaker does not know the

answer to the question who (or which one) John is dating — and a familiarity presupposition carried by the definite description, demonstrative, pronoun, or name in the first clause.[49] One may reasonably assert that John is dating someone without knowing who he is dating, but if one is in a position to assert that John is dating the woman in the blue dress, then one has an answer to the question "Which woman is John dating?" — namely, the woman in the blue dress — so (79)b. conflicts with itself pragmatically (and similarly for (79)c.–e.) Example (78) does not show the same anomaly because the verb *announce* acts as a plug to presupposition projection (Karttunen (1973)), allowing felicitous use by speakers who are unfamiliar with the particular asparagus mentioned in the announcement.

We may use the pattern illustrated in (79) to "test" whether tensed sentences express existential quantification over times or instead involve indexical reference to a specific time. If they involve existential quantification, we expect that a tensed sentence may be continued by the phrase *but I don't know when* without anomaly. But if tensed sentences involve reference to a particular time fixed by the referential intentions of the speaker, we expect the opposite, at least if we assume that a speaker who intends reference to a particular time must know which time he or she intends to refer to. The actual pattern we observe is the former one; it is perfectly natural to follow a tensed sentence with *but I don't know when*:

(80) John turned off the stove, but I don't know when.

This suggests that at least sometimes, tensed sentences express existential quantification over times rather than indexical reference to a particular time intended by the speaker.

Of course, this does not mean that the tense marker itself is an existential quantifier; as discussed on p. 49 above, it may be that tenses, like indefinite descriptions, contribute a free variable to the semantic derivation, with the quantificational force contributed by a general operation of existential closure rather than by the internal semantics of the tense morpheme itself. This is, in fact, what I will assume.

To summarize: I have argued that tensed sentences *do* make indexical reference to the time of utterance, so that they express different contents at different times — as opposed to expressing time-invariant contents which receive different truth values at different times — but that they do *not* (always) make indexical reference to a specific time fixed by the referential intentions of the speaker, and located before, concurrent with, or after the time of utterance, according to whether the tense is past, present or future. Rather, such sentences may express existential quantification over past, present or future times. But I assume this quantificational force is due to a general operation of existential closure, and that the tense morphemes themselves simply contribute free variables ranging over past, present or future times to the semantic derivation.

To implement this technically, let us treat verbs as derived morphologically from their infinitival "base" forms, through the addition of inflectional affixes or other markers. We need not

[49] Dayal and Schwarzschild's analysis is actually a little more complicated than this, and imposes an additional constrain that the *Wh*-phrase must match its antecedent in the presence or absence of a contentful head noun. See their article for details concerning this constraint, and for discussion of examples involving affirmative sentences rather than negative sentences in the second clause, verbs other than *know*, and non-first-person subjects.

concern ourselves here with the details of morphological theory, but simply assume that for any verb α, its base form BASE(α) may be identified, with the result that:

(81) a. BASE($sleep_{1sg}$) = BASE($sleep_{2sg}$) = BASE(*slept*) = BASE($sleep_{pl}$) = BASE(*sleeps*) = $sleep_{inf}$

b. BASE(see_{1sg}) = BASE(see_{2sg}) = BASE(*saw*) = BASE(see_{pl}) = BASE(*sees*) = see_{inf}

We regard the BASE relation as holding at the level of uses, not just expressions:

(82) $\exists\beta[\beta = \text{BASE}(\alpha)]$ iff $[\text{USE}(\alpha,\boldsymbol{\alpha}) \;\&\; \text{USE}(\beta, \boldsymbol{\beta}) \;\&\; \boldsymbol{\beta} = \text{BASE}(\boldsymbol{\alpha})]$

It will be important later that infinitival verbs do *not* include a free variable over times, but instead simply have an open argument place for times. We therefore define $sleep_{inf}$ and see_{inf} as follows:

(83) a. If USE(α, $sleep_{inf}$) then $Lex(\alpha, u, w) = \lambda x : x \in \mathbf{D}_e \,.\, \lambda t \in \mathbf{D}_i \,.\, x$ sleeps at t in w

b. If USE(α, see_{inf}) then $Lex(\alpha, u, w) = \lambda x : x \in \mathbf{D}_e \,.\, \lambda y : y \in \mathbf{D}_e \,.\, \lambda t \in \mathbf{D}_i \,.\, y$ sees x at t in w

We may now define tensed verbs as follows, where '$t < t'$' means that t precedes t':[50]

(84) For any use α of a verb $\boldsymbol{\alpha}_i$, types $\tau_1,\ldots,\tau_n$, and worlds u, w, if α is of type $\langle\tau_1,\langle\ldots,\langle\tau_n,\langle i,t\rangle\rangle\rangle\rangle$, then:

a. If $\boldsymbol{\alpha}_i$ is of category PAST, then $Lex(\alpha, u, w) = \lambda r : Ref_{u,\alpha} \preccurlyeq r$ and $r(i) < time_{u,\alpha} \,.\, \lambda x_1 \in \mathbf{D}_{\tau_1} \ldots \lambda x_n \in \mathbf{D}_{\tau_n} \,.\, Lex(\text{BASE}(\alpha), u, w)(x_1)\ldots(x_n)(r(i))$

b. If α is of category PRES, then $Lex(\alpha, u, w) = \lambda r : Ref_{u,\alpha} \preccurlyeq r$ and $r(i) = time_{u,\alpha} \,.\, \lambda x_1 \in \mathbf{D}_{\tau_1} \ldots \lambda x_n \in \mathbf{D}_{\tau_n} \,.\, Lex(\text{BASE}(\alpha), u, w)(x_1)\ldots(x_n)(r(i))$

This gives the result that if φ is a use of $John_3$ $slept_2$, for example, then $[\![\varphi]\!]^{u,w}$ is $[\lambda r : Ref_{u,\varphi} \preccurlyeq r$ and $r(2)$ precedes $time_{u,\varphi}$ and $Ref_{u,\varphi}(3)$ is named *John* in $u \,.\, Ref_{u,\varphi}(3)$ sleeps at $r(2)$ in $w]$.

Nothing in our rules requires the temporal index on a tense to be assigned a value by $Ref_{u,\varphi}$, and nothing prohibits it from being assigned a value. If it is not assigned a value by $Ref_{u,\varphi}$, it will be existentially quantified by our general existential closure operation — the interpretation suggested by the sluicing patterns discussed above. If it is assigned a value by $Ref_{u,\varphi}$, it will be interpreted as making indexical reference to a particular time according to the referential intentions of the speaker — essentially the interpretation originally advocated by Partee. Although I think it is clear that the former interpretation must be available, I see no harm in also making the second interpretation

[50] These definitions should not be construed as applying to $don't_{1sg}$, $don't_{2sg}$, *doesn't*, $don't_{pl}$ and *didn't*, since (in the "*do* support" usage modeled here) these verbs do not have an infinitival base form.

available as well, as long as it is not the only option. We therefore impose no condition on whether the temporal index must be assigned a value by $Ref_{u,\varphi}$ or not.[51]

3.5. Spatial deixis

In this section we consider locative indexicals such as *here* and *there*, as well as expressions such as *in front of* which sometimes vary in content according to a relevant perspective point. As we will also see, there is a kind of indexicality built into all locative expressions, because locations themselves are differentiated in a context-sensitive way.

Most work on indexicality in logical semantics has had little to say on the subject of spatial location — it often seems to be assumed that some theory of metaphysics or common-sense ontology will simply provide us with a set of locations, which can then serve as the referents for locative expressions in straightforward manner, without the semantic theory itself having much to say about how these locations are structured or differentiated. This is, I think, a reasonable practice for many purposes, but in the present context it is important to recognize a complication which is often ignored: What counts as "the same location" is dependent on context in a way which is a reminiscent of other contextual effects on interpretation.

Of course modern theories of physics do entirely without a notion of absolute location; only relative to an inertial frame of reference can it be said whether an object changes location or not. It is quite debatable whether this has relevance to the issue of how to formulate the semantic component of a natural language grammar, since the facts which motivate this conception of location are far removed from the everyday communicative practices and cognitive apparatus by which the structure of our language has presumably been shaped. But it would be a mistake to conclude therefore that the language portrays location as absolute. We don't actually need to appeal to the exotica of Special Relativity to see the effects of something like frames of reference (though not exactly *inertial* frames of reference) on our judgments about location and change of location.

Suppose, for example, that John is riding the train. He gets up to use the restroom, leaving his cell phone on his seat. When he returns, he notices that the phone is no longer on his seat, and exclaims, pointing: "I left it right here!" Of course the train has continued its journey while John was in the restroom, so relative to the earth, he *didn't* actually leave it in the same location as he points to as he makes his utterance; but nonetheless he intuitively says something true. The reason, of course, is that he is differentiating locations based on their position relative to the train, not relative to the earth, much less absolutely.

Locations are differentiated based on their position relative to the train in this scenario because to someone riding in the train, it serves naturally as a background against which the foci of attention stand out; in other words it serves as the ground in figure/ground perception.[52] I will not

[51] The analysis of tense given here is still very simplified, and leaves a number of complications unaccounted for, including generic uses of the past and present tenses, binding of the present tense by temporal adverbials, futurate uses of the present, etc. A full-scale analysis of tense in English is well beyond the scope of this study.

[52] The fundamental point that locations are differentiated in natural language on the basis of figure-ground relations goes back at least to Talmy (2000), though it is somewhat obscured by Talmy's choice of terminology. Talmy adopts a tripartite distinction among a "Figure," a "Ground" (also called a "reference object") and a "background" (also called a "frame of reference"). He makes clear that the background should be identified with the ground "in the psychological sense," even though it is distinct from the Ground (i.e. reference object). It is the background, not the Ground, which is regarded as stationary for the purposes of distinguishing locations, determining whether something has moved, etc.

attempt here to formalize the representation of figure/ground relations (but see, e.g., Bittner (1997), Bittner (1998)) or discuss the cognitive factors which determine the identification of a figure and ground, but will simply assume that for each individual *x*, and each interval *I* and world *w* at which *x* has the appropriate cognitive capacities, there is a set $space_{x,I,w}$, construed by *x* throughout *I* in *w* as stationary spatial points. Let also assume (for all such *x, I, w*) that certain subsets of $space_{x,I,w}$ are distinguished as lines, that each line segment is assigned a definite length and each pair of intersecting lines a definite angle, and that all these concepts relate to each other as in common-sense (presumably Euclidean) three-dimensional geometry.

Of course the various spaces set up in this way are not totally independent of one another. At any given time, a point in one space may have a corresponding point in any other space; we simply allow the correspondence to "shift" over time. More formally, for every instant *t* in an interval *I*, let us require a reflexive, symmetric, transitive relation $\equiv_t$ on $\cup_{x,I,w} space_{x,I,w}$, with the further requirement that if $p \equiv_t p'$ and $p, p' \in space_{x,I,w}$, then $p = p'$. (That is, the only point equivalent to *p* in its own space is *p* itself.)

By a *location* (relative to *x,I,w*), let us mean any subset of $space_{x,I,w}$; and let us further assume that (for each *x, I, w*) every element *y* of some distinguished set of objects is assigned a location in $space_{x,I,w}$ relative to certain times and worlds (namely the times and worlds where it exists and is located in $space_{x,I,w}$).[53] We write '$loc_{x,I,w}(y,t,w')$' to mean "the location of *y* at *t′* in *w′*, in the space given by *x*'s figure-ground articulation over *I* in *w*.[54] Where *l, l′* are locations in $space_{x,I,w}$, $space_{x',I',w'}$, respectively, we say that $l \equiv_t l'$ iff $l' = \{p' \in space_{x',I',w'} : \exists p \in l\ [p \equiv_t p']\}$. We require that for all *x, x′, y, t, I, I′, w, w′, l, l′*: If $loc_{x,I,w}(y, t, w'') = l$ and $loc_{x',I',w'}(y, t, w'') = l'$, then $l \equiv_t l'$. That is, object locations are preserved across the $\equiv_t$ relation.[55]

To integrate locations into our type system, we expand it as follows (cf. (55), (56)):

(85) a. l is a type.
b. $\mathbf{D}_l = \{l \mid \exists x \in \mathbf{D}_e\ \exists I \in \mathbf{D}_i\ \exists w[w \text{ is a possible world}] \,.\, l \subseteq space_{x,I,w}\}$

If as suggested here, locations are linguistically differentiated on the basis of figure-ground relations, an interesting question arises: For any given (use of) a locative expression, who is the cognitive agent whose perceptions determine the relevant figure-ground relations (and at what time and world)? A natural first guess, which works well in the simple cases — for example the case of John saying "I left it right here," is that the relevant agent is the speaker, at the time and world of the context of use. As we shall see, things are really quite a bit more complicated than this, but let us

[53] Perhaps it should be clarified that even though locations are differentiated on a cognitive basis, the intention here is that the location assigned to an object relative to the actual world should be its actual location, relative to the space defined by the figure-ground articulation of the relevant agent — not the location which that agent perceives, imagines, or believes the object to have. Appealing to cognitive factors in the differentiation of space does not imply that the notion of actual location must be conflated with that of perceived or imagined location.

[54] For typical concrete individuals, we should presumably also impose a requirement that their locations be contiguous (that is, that for any two points in the individual's location, there must be a path which connects them without passing through any points not in that location). But there are numerous concrete objects for which such a condition would be inappropriate — archipelagos, coin collections, etc.

[55] We normally interpret language with an assumption that not just location, but also distance and volume are preserved across this relation, but it is not clear to me that this assumption is built into the grammatical structure of the language in any way.

delay consideration of the complications just a little bit.

Let us turn now to *here* and *there*. The analysis in Kaplan (1989) treated each context as having a "position," and *here* as denoting, relative to a given context *c*, the position of *c*; we followed this approach in our initial treatment of indexicality in (18). Kaplan noted (following Michael Bennett) that *here* also has a "demonstrative" use, in which it denotes some other location indicated by the speaker, as when the speaker points at a map and says "In two weeks, I will be here." The formal treatment of *here* in Kaplan's paper — which does not cover this demonstrative use — is therefore at best only a reasonable simplification for the purposes of making the theoretical points Kaplan had in mind. A more detailed analysis of *here* for its own sake would have to account for the demonstrative use and for a number of other complications not mentioned by Kaplan.

Even setting aside examples in which the speaker points at a map or similar representation, *here* is used not just to denote the location of utterance but to denote any location which is sufficiently close to the location of the speaker. For example, I can instruct someone where on my desk to place a book by saying "Put it here, not here," pointing at two locations on my desk as I utter the two tokens of *here*. Neither spot on the desk is the location where the utterance takes place; they are both simply nearby enough to warrant the use of *here*. As linguists have long put it, *here* is a "proximal" locative expression, in contrast with *there*, which is "distal."[56] It seems appropriate in incorporating *here* and *there* into our grammar, therefore, to subscript each occurrence of either of these words with a positive integer, to indicate the intended location. In addition, we assume each use α of *here* is associated (in each world *u* where it occurs) with a location $proximity_{u,\alpha}$ in $space_{speaker_{u,\alpha},time_{u,\alpha},u}$, such that $loc_{speaker_{u,\alpha},time_{u,\alpha},u}(speaker_{u,\alpha}) \subseteq proximity_{u,\alpha}$. We might then stipulate that if USE(α, $here_i$), then $[\![\alpha]\!]^{u,w} = Ref_{u,\alpha}(i)$, provided $Ref_{u,\alpha}(i)$ is part of $proximity_{u,\alpha}$ (undefined otherwise); and that USE(α, $there_i$), then $[\![\alpha]\!]^{u,w,p} = Ref_{u,\alpha}(i)$, provided $Ref_{u,\alpha}(i)$ does not overlap with $proximity_{u,\alpha}$ (undefined otherwise). Note that the semantics of each of these items is sensitive to two separate contextual parameters: the particular location picked out by the referential index *i*, and the area specified as $proximity_{u,\alpha}$.

Returning now to the map example, we may see that there are at least three location parameters involved in the semantics of *here* and not just two. When the speaker points at the map and says "I will be here next week," the location where the speaker says he or she will be next week is determined contextually. But even though the speaker uses *here* rather than *there*, this location need not be proximal at all. The use of *here* is licensed not because this location is proximal, but because the portion of map to which he or she points is proximal. This portion of the map represents a location, and it is this represented location at which the speaker must be located the following week in order for the assertion to be true; but it is the location of this portion of the map which governs the choice between *here* and *there*. More generally, the semantics of *here* is sensitive to the specific location demonstrated by the speaker, to the region around the speaker which counts as proximal, and to the location which the demonstrated location represents; all of these locations are determined

[56] The proximal/distal contrast also governs the choice between *this* and *that*. Some languages (such as Japanese) show a three-way contrast, roughly corresponding to "near the speaker," "near the addressee," and "far from both the speaker and the addressee," in effect tying locative reference to person-marking. Other languages — including archaic and dialectal English — draw finer distinctions by marking evidentiality together with distance, as in *yonder* "distal but (literally or metaphorically) within view."

contextually. This suggests we should revise our analysis of *here* and *there* to assign them each two syntactic indices, and give definitions as follows:[57]

(86) a. If USE(α, $here_{i,j}$) then $Lex(\alpha, u, w) = Ref_{u,\alpha}(j)$, provided $Ref_{u,\alpha}(i) \subseteq proximity_{u,\alpha}$ and $Ref_{u,\alpha}(i)$ represents $Ref_{u,\alpha}(j)$ (undefined otherwise)

b. If USE(α, $there_{i,j}$) then $Lex(\alpha, u, w) = Ref_{u,\alpha}(j)$, provided $Ref_{u,\alpha}(i)$ does not overlap with $proximity_{u,\alpha}$ and $Ref_{u,\alpha}(i)$ represents $Ref_{u,\alpha}(j)$ (undefined otherwise)

It may be worth seeing one of these definitions deployed in an actual derivation. To do this, we will need to expand our syntax a little. Let us expand our set of major categories to include a third category, which I will notate 'P'. We let *here* and *there* belong to this category. We also add the various forms of the verb *leave*:[58]

(87)
- $leave_{inf}$: (((ACC\V)/P)/ACC) ∩ INF
- $leave_{1sg}$: ((((NOM,1P,SG)\V)/P)/ACC) ∩ PRES
- $leave_{2sg}$: ((((NOM,2P,SG)\V)/P)/ACC) ∩ PRES
- *leaves*: ((((NOM,3P,SG)\V)/P)/ACC) ∩ PRES
- $leave_{pl}$: ((((NOM,PL)\V)/P)/ACC) ∩ PRES
- $left_{vb}$: (((NOM\V)/P)/ACC) ∩ PAST

Our grammar will now generate sentences such as $John_1$ $left_{vb2}$ it_3 $here_{4,5}$.

To leave something at a place is roughly to move in such a way that one is no longer near its location. With this idea in mind, we may define *leave* as follows:[59]

(88) If USE(α, $leave_{inf}$) then $Lex(\alpha, u, w) = \lambda x : x \in \mathbf{D}_e . \lambda l : l \in \mathbf{D}_l . \lambda z : z \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i .$ $loc_{speaker_{u,\alpha},time_{u,\alpha},u}(x, t, w) \subseteq l$ and there exists some $t' < t$ such that for all t'', if $t' < t'' < t$, then $loc_{speaker_{u,\alpha},time_{u,\alpha},u}(z, t'', w)$ is near l, and $loc_{speaker_{u,\alpha},time_{u,\alpha},u}(z, t, w)$ is not near l

[57] To handle cases where nothing like a map is used, we consider every location to represent itself, in addition to whatever other locations it may represent. These definitions are still too simple to handle special uses of *here* and *there* such as occur in free indirect discourse; see Banfield (1982), Doron (1991), Schlenker (2004).

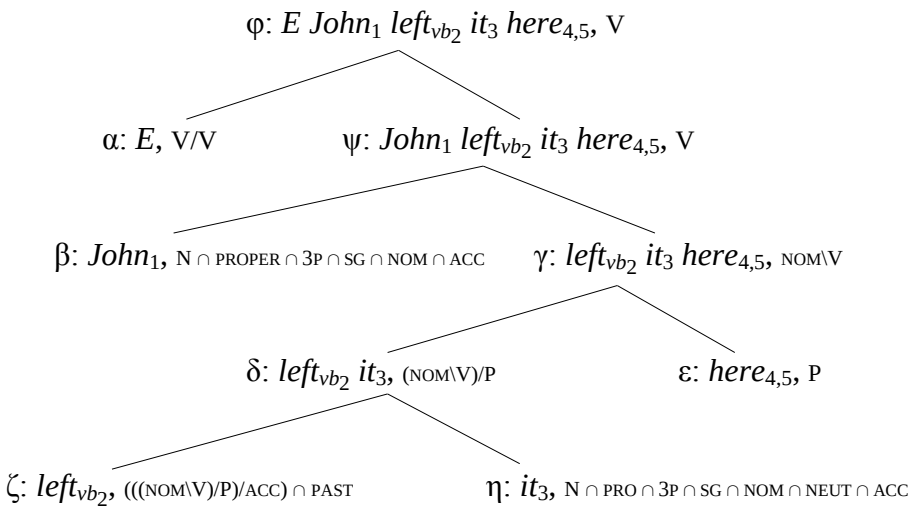
[58] The past tense form is notated $left_{vb}$ to distinguish it from the homophonous noun, antonym of *right*.

[59] One might quibble with this definition. It requires the left object to be at a location at the moment the subject stops being near that location, but does not require the left object to be at that location before or after that one moment. It is debatable whether this is correct. Certainly it would be unusual to say that A left B at L, in the case where B was not at L during the time when A was near L or for any time after the moment when A stopped being near L. (For example, it would be odd to describe two people who pass each other on the street as leaving each other at the place where they pass each other.) But even though it would be *odd to say* that A left B at L in such a case, I am much less sure that it would be literally *false*. It could be that the inference that B was at L longer than that one moment is some sort of implicature, rather than part of the truth conditions.

The definition is also stated using the vague word *near*, yet requires there to be a particular instant at which a moving object stops being near a location; but correcting this problem would require a general theory of vagueness, which cannot be undertaken here.

Of course we assume BASE(*leave*$_{1sg}$) = BASE(*leave*$_{2sg}$) = BASE(*leaves*) = BASE(*leave*$_{pl}$) = BASE(*left*$_{vb}$) = *leave*$_{inf}$. By our rules now, a use φ of *E John*$_1$ *left*$_{vb2}$ *it*$_3$ *here*$_{4,5}$, will have the syntactic structure in (89) and the content in (90):

(89)



(90) $\llparenthesis \varphi \rrparenthesis^u = \lambda w$. $\exists r$ [$Ref_{u,\varphi} \preccurlyeq r$ and $Ref_{u,\varphi}(1)$ is named John and $r(2) < time_{u,\zeta}$ and $Ref_{u,\varphi}(3)$ is sexless in u and $Ref_{u,\varphi}(4)$ is part of $proximity_{u,\varepsilon}$ and $Ref_{u,\varphi}(4)$ represents $Ref_{u,\varphi}(5)$. $loc_{speaker_{u,\zeta},time_{u,\zeta},u}(Ref_{u,\varphi}(3), r(2), w) \subseteq Ref_{u,\varphi}(5)$ and there exists some $t' < r(2)$ such that for all t'', if $t' < t'' < r(2)$, then $loc_{speaker_{u,\delta},time_{u,\zeta},u}(Ref_c(1), t'', w)$ is near $Ref_{u,\varphi}(5)$, and $loc_{speaker_{u,\zeta},time_{u,\zeta},u}(Ref_c(1), r(2), w)$ is not near $Ref_{u,\varphi}(5)$

A somewhat more complex pattern of spatial indexicality is exhibited by orientation expressions such as *in front of*, *behind*, *left* and *right*. As classically pointed out by Fillmore (1997), a semantic analysis of these expressions requires us to distinguish objects which have intrinsic fronts and tops from those which don't. *John is in front of the house* normally means that he is outside the house but near it, and nearer to the intrinsic front of the house than to the sides or back; but because trees do not normally have intrinsic fronts, *John is in front of the tree* means that he is between the tree and the relevant perspective point (which for now we may take to be that of the speaker). *In front of* thus introduces a kind of indexicality when its complement denotes an object without an intrinsic front, but does not seem to do so when its complement denotes an object which does have an intrinsic front.

The characteristics of an object which lead it to have an intrinsic front are complex and cannot be dealt with in detail here. To summarize some of the main considerations: The front of a person, or of an animal with a face, is the side with the face. If an object with no face is moving, the side toward the direction of movement is the front.[60] If a movable object is stationary, but has a side

[60] It is rather odd to call the front "intrinsic" in this case, since the object may have no front until it moves, and stop having a front as soon as it stops moving. In the case of a rolling object, which side counts as the front is constantly

which is *normally* oriented toward the direction of movement when the object does move, that side is the front. If there is a side which is normally toward a person who views, approaches or interacts with the object, that side is the front.[61]

It is presumably this last condition which provides the conceptual link between indexical and non-indexical uses of *in front of*. Just as certain objects have no intrinsic front when stationary, but acquire one when they move, certain objects may have no intrinsic front when no one views, approaches, or interacts with them, but acquire a front when someone does. If only one person at a time could view, approach or interact with such an object, this would not motivate an indexical semantics for the word *front*; we could simply say that the front of the object (at a given time *t*) was the side facing the unique relevant individual (at *t*). But because different individuals may simultaneously view, approach or interact with an object from different sides, each describing some other object as being in front of the first (or not), based on his or her own location, we must regard the truth values of such statements as depending on the contexts in which they are made.

It is not just the truth values of such statements which depend on the context, but also the contents. If John and Mary approach the tree from the west, and John tells Mary "The rock is in front of the tree," while Bill and Sue approach the tree from the east, and Bill tells Sue "The rock is not in front of the tree," we do not take John and Bill to be contradicting each other. Even though the form of Bill's sentence is a negated version of John's sentence, the fact that there is no contradiction shows that Bill is not denying the same content as John is asserting, any more than if John had said "My name is *John*" and Bill had said "My name is not *John*."

As a rough first attempt at a semantics for *in front of*, then, we might take *in front of* to be of syntactic category (P/ACC) and define it as follows:[62]

(91) If USE(α, *in front of*) then $Lex(\alpha, u, w) = \lambda x \in \mathbf{D}_e$. $\{l \in space_{speaker_{u,\alpha},time_{u,\alpha},u} \mid l$ is near $loc_{speaker_{u,\alpha},time_{u,\alpha},u}(x, time_{u,\alpha}, w)$; and if x has an intrinsic front at $time_{u,\alpha}$ in w, then l is nearer to the intrinsic front of x at $time_{u,\alpha}$ in w than to any other side of x, else l is between $loc_{speaker_{u,\alpha},time_{u,\alpha},u}(x, time_{u,\alpha}, w)$ and $loc_{speaker_{u,\alpha},time_{u,\alpha},u}(speaker_{u,\alpha}, time_{u,\alpha}, u)$ at $time_{u,\alpha}$ in w.$\}$

More informally,[63] a use of *in front of X* denotes a location; specifically, a location *l* such that all of *l* is near the location of *X* at the time of use; and if *X* has an intrinsic front at that time, then all of *l* is nearer to the intrinsic front of *X* than to any other side of *X*; but if *X* does not have an intrinsic front at that time, then all of *l* is between *X* and the speaker's location at that time.

This is still unsatisfactory, however, for well-known reasons.[64] The relevant perspective for determining whether something is in front of an object with no intrinsic front is not always that of the speaker; nor is the relevant time always the time of use or the world the world of use. It is quite

---

changing. But I will continue to label such examples involving an "intrinsic" front, to distinguish them from cases where the assignment of a front depends on the perspective of the speaker or other relevant agent.

[61] In a few cases, these considerations conflict with one another. If my house is being carried away by a landslide, do I count as in front of it if I am in its path, or if I am on the side of the main door, which had been facing the street? It seems to me we may say I am in front of the house in either case.

[62] I assume that *in front of* has been fossilized into a kind of complex preposition, rather than treating phrases of the form *in front of X* as composed of the preposition *in* taking as its object a noun phrase of the form *front of X*.

[63] and mixing object language and metalanguage in the interest of readability

[64] See especially Fillmore (1997). Fillmore devotes more attention to *come* and *go* than to *in front of*, but the problems that come up are directly analogous.

common for a speaker to "adopt the perspective" of the addressee in using expressions like *in front of*. For example, if I am conveying instructions to you over the phone about where to find the key to my house, and I know that you are standing in my yard, I can say "Turn and face the big oak tree. In front of the tree is a rock. The key is under that rock." Especially in narrative discourse, it is common for the relevant perspective to be that of a third person, neither the speaker nor the addressee: "Bill entered the yard and saw a large tree. In front of the tree was a rock." And as this example also shows, it might not be the relevant individual's perspective at the time of utterance which is relevant, but rather his or her perspective at some other time. In some cases, the relevant perspective varies with the values of variable:

(92) The contestants all approached the tree from different angles. As they came closer, each contestant noticed an odd rock in front of the tree.

The second sentence in (92) admits an interpretation in which each contestant notices a different rock, and for each contestant, the rock he or she notices is located between that contestant and the tree. We will not obtain that reading if we treat the relevant perspective for the interpretation of *in front of* as fixed globally by the context of use.

Therefore, let us annotate each occurrence of *in front of* with a referential index *i* and a temporal index *j*, then revise the definition for *in front of* as follows:

(93) If USE(α, *in front of*$_{i,j}$) then $Lex(\alpha, u, w) = \lambda r : Ref_{u,\alpha} \preccurlyeq r \,.\, \lambda x \in \mathbf{D}_e \,.\, \{l \in space_{r(i),r(j),w} \mid l$ is near $loc_{r(i),r(j),w}(x, r(j), w)$; and if x has an intrinsic front at $r(j)$ in w, then l is nearer to the intrinsic front of x at $r(j)$ in w than to any other side of x; else l is between $loc_{r(i),r(j),w}(x, r(j), w)$ and $loc_{r(i),r(j),w}(r(i), r(j), u)$ at $r(j)$ in $w\}$

That is, we simply replace references to $speaker_{u,\alpha}$ with references to the value of i, and references to $time_{u,\alpha}$ with references to the value of j. In effect, this approach treats *in front of* as containing two free variables — one over individuals and the other over times. These variables would then be available for binding by quantifiers, in much the same ways as pronouns may be bound by quantifiers.[65]

Having recognized that the relevant perspective point for the interpretation of *in front of* is not always that of the speaker, we should reconsider the semantics for *leave*. It too may be interpreted in a way which depends on a differentiation of locations by someone other than the speaker. For example, one may say *John left his phone on the seat*, treating the seat as a stationary location because it is stationary relative to the train on which John was riding, even if the speaker is not on the train too. Therefore, let us index *leave* with a pair of indices *i, j* and revise (88) as follows:

[65] Some readers may have the intuition — which I also initially had — that all occurrences of *in front of* (and similar expressions) in a single clause must be indexed in the same way. There is something odd about saying *The rock was in front of the tree in front of the telephone pole* to mean that the rock was between the tree and John, and the tree was between the telephone pole and Mary (who was standing apart from John), for example. With enough context, however, I think precisely this kind of perspective shift within a single clause is possible. Suppose you and I are standing in the yard and can see various trees, as well as a telephone pole. I am telling you about what happened to Bill earlier that day: "He saw a rock in front of that tree in front of the telephone pole." Here, the interpretation is that the rock was between Bill and the tree, but the tree is between us and the telephone pole.

(94) If USE(α, $leave_{infl\,i,j}$) then $Lex(\alpha, u, w) = \lambda r : Ref_{u,\alpha} \preccurlyeq r \,.\, \lambda x : x \in \mathbf{D}_e \,.\, \lambda l : l \in \mathbf{D}_l \,.\, \lambda z : z \in \mathbf{D}_e \,.\, \lambda t \in \mathbf{D}_i \,.\, loc_{r(i),r(j),w}(x, t, w) \subseteq l$ and there exists some $t' < t$ such that for all t'', if $t' < t'' < t$, then $loc_{r(i),r(j),w}(z, t'', w)$ is near l, and $loc_{r(i),r(j),w}(z, t, w)$ is not near l

What about *here* and *there*? Should they also be redefined to allow binding of their syntactic indices? At first glance the answer might seem to be no: *here* and *there* are tied to the speaker much more tightly than *in front of, come, go*, etc. It is not appropriate to say *It was right here, in front of the tree*, for example, pointing to a position which is between the addressee and the tree but distant from the speaker. This suggests that the denotation of *here* is not "shiftable" away from the speaker the way the relevant perspective for the interpretation of *in front of* is, which might lead us to expect that it will likewise not be bindable.

However, it becomes clear that there is a bindable parameter in the interpretation of *here* as soon as we remember that it is only the *representing* location which is required to be proximal to the speaker, while the *represented* location need not be. Suppose each man rides in a different train car of the same model, and leaves his phone on the rightmost seat in the third row. A speaker can describe this by pointing at a diagram showing the layout of that model of train car, indicating the part of the diagram representing the relevant seat, and saying *Every man left his phone here*. For each man *x*, the differentiation of locations is different — treating the car in which *x* was riding as the stationary background — and the location where *x* left his phone is different too (since there was no pile of phones in the end). We can obtain this effect adding a third index *k* to *here* and *there*, bindable by a nominal quantifier. This requires revising our definitions so that *here* and *there* are of type ⟨r, l⟩, and so that index *i* (for the representing location) takes its value from $Ref_{u,\alpha}$ (keeping it strictly indexical and not bindable), but indices *j* and *k* can take their values from extensions of $Ref_{u,\alpha}$:

(95) a. If USE(α, $here_{i,j,k}$) then $Lex(\alpha, u, w) = \lambda r$: $Ref_{u,\alpha} \preccurlyeq r$ & $Ref_{u,\alpha}(i) \subseteq proximity_{u,\alpha}$ & $\exists t\, r(j) \subseteq space_{r(k),t,w}$ & $Ref_{u,\alpha}(i)$ represents $r(j)$. $Ref_{u,\alpha}(j)$

b. If USE(α, $there_{i,j,k}$) then $Lex(\alpha, u, w) = \lambda r$: $Ref_{u,\alpha} \preccurlyeq r$ & $Ref_{u,\alpha}(i)$ does not overlap with $proximity_{u,\alpha}$ & $\exists t\, r(j) \subseteq space_{r(k),t,w}$ & $Ref_{u,\alpha}(i)$ represents $r(j)$. $Ref_{u,\alpha}(j)$

3.6. Negation and contradiction

Since much of our concern in giving a semantics for personal taste sentences will be with accurately capturing our intuitions about when two assertions contradict each other, it may be useful to explore briefly how contradiction works in the current system, which has indexical expressions but no predicates of personal taste. This is most directly accomplished if we have a semantics for negation.

Negation is expressed in the present grammar via the auxiliary verbs *don't*$_{1sg}$, *don't*$_{2sg}$, *doesn't*, *don't*$_{pl}$ and *didn't*. Of course these are composed morphologically of *do*, tense and negation, but we simplify here and define them directly as wholes:[66]

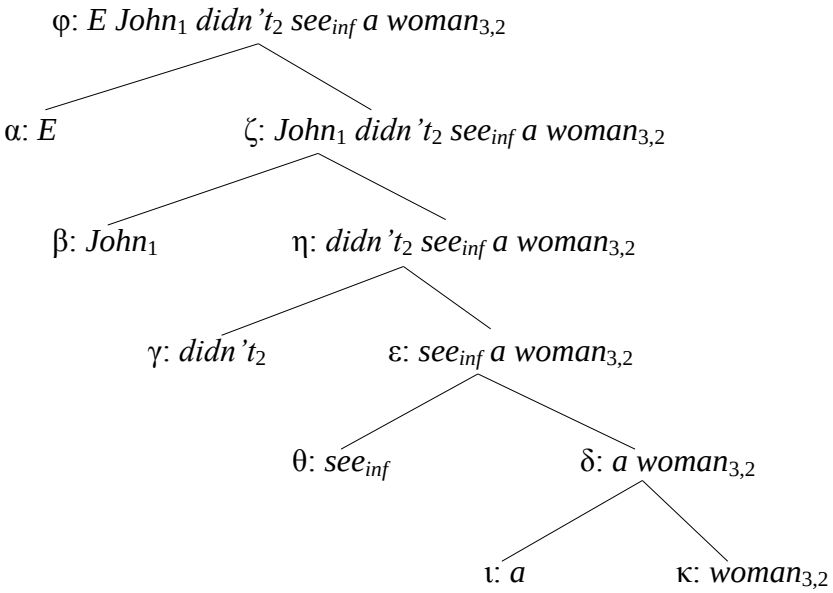
[66] It may be noticed that our semantics does not allow negation to scope over the subject of its clause. My suspicion is that when negation which appears syntactically on a verb appears to scope over the subject of the verb, it is actually metalinguistic negation in the sense of Horn (1985) and not wide-scope logical negation in the usual sense. Readers who

(96) a. If USE(α, *don't*$_{1sg_i}$) or USE(α, *don't*$_{2sg_i}$) or USE(α, *doesn't*$_i$) or USE(α, *don't*$_{pl_i}$), then *Lex*(α, *u*, *w*) = [$\lambda r : Ref_{u,\alpha} \preccurlyeq r \,.\, \lambda f : f \in \mathbf{D}_{\langle r,\langle i,\langle e,t\rangle\rangle\rangle} \,.\, \lambda x : x \in \mathbf{D}_e$. There is no r' such that $Ref_{u,\alpha} \preccurlyeq r'$ and $r'(i) = time_{u,\alpha}$ and $f(r')(x)(r'(i))$ = **truth**]

b. If USE(α, *didn't*$_i$) then *Lex*(α, *w*, *w'*) = [$\lambda r : Ref_{u,\alpha} \preccurlyeq r \,.\, \lambda f : f \in \mathbf{D}_{\langle r,\langle i,\langle e,t\rangle\rangle\rangle} \,.\, \lambda x : x \in \mathbf{D}_e$. There is no r' such that $Ref_{u,\alpha} \preccurlyeq r'$ and $r'(i)$ is a time and $r'(i) < time_{u,\alpha}$ and $f(r')(x)(r'(i))$ = **truth**]

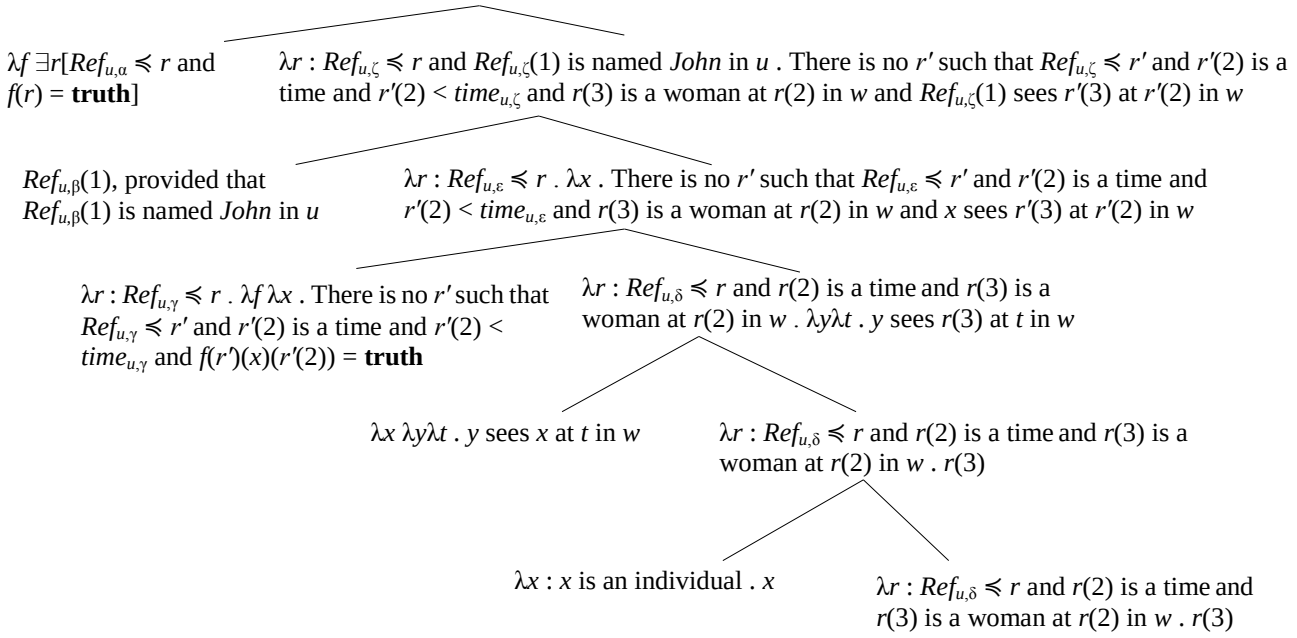
As an example, consider the sentence *E John*$_1$ *didn't*$_2$ *see*$_{inf}$ *a woman*$_{3,2}$, interpreted relative to worlds *u*, *w*:

(97)



are skeptical about this may recast the analysis using their favorite technique for allowing negation to scope over subjects; nothing in the rest of the analysis hinges on this point.

(98) $\exists r$[$Ref_{u,\varphi} \preccurlyeq r$ and $Ref_{u,\varphi}(1)$ is named *John* in u . There is no r' such that $Ref_{u,\varphi} \preccurlyeq r'$ and $r'(2)$ is a time and $r'(2) < time_{u,\varphi}$ and $r(3)$ is a woman at $r(2)$ in w and $Ref_{u,\varphi}(1)$ sees $r'(3)$ at $r'(2)$ in w]



Putting it more informally: *John didn't see a woman* presupposes that the individual referred to by *John* in context is named *John*; and is true if there is no woman whom that individual sees at any relevant time preceding the time of the context of the verb; and false if there is such a woman.[67]

Comparing this to the affirmative sentence, readers should be able to confirm that *E John*$_1$ *saw*$_2$ *a woman*$_{3,2}$ presupposes that the individual referred to by *John* in its context is named *John*; and is true if there is a woman whom that individual sees at some relevant time preceding the time of the context of the verb, and false if there is no such woman.

Notice that if φ and ψ are uses of *E John*$_1$ *saw*$_2$ *a woman*$_{3,2}$ and *E John*$_1$ *didn't*$_2$ *see*$_{inf}$ *a woman*$_{3,2}$, respectively, where the name *John* is used to refer to the same individual and the times of the tensed verbs are the same, there will be no world w such that $[\![\varphi]\!]^{u,w}$ = **truth** and $[\![\psi]\!]^{u,w}$ = **truth** (for any u). By our notational conventions, this is just a variant of saying that there is no world w such that $[\![\langle\varphi\rangle^u]\!]^w$ = **truth** and $[\![\langle\psi\rangle^u]\!]^w$ = **truth**. That is, there is no world w such that $\langle\varphi\rangle^u$ is true relative to w and $\langle\psi\rangle^u$ is true relative to w. This matches our definition of contradiction, which said that Φ contradicts Ψ iff there is no world w such that Φ is true relative to w and Ψ is true relative to w. So we can say that as long as the references of the indexical items are kept fixed, *John saw a woman* contradicts *John didn't see a woman*.

However, it is entirely possible that there might be uses of these sentences φ, ψ and worlds u, w, such that $[\![\varphi]\!]^{u,w}$ = **truth** and $[\![\psi]\!]^{u,w}$ = **truth**. This would happen, for example, if the referent of *John* sees a woman prior to $time_{u,\varphi}$, but not prior to $time_{u,\psi}$. In this case $\langle\varphi\rangle^u$ and $\langle\psi\rangle^u$ are both true relative to w, and there is no contradiction. Thus we see that because different uses of sentence may

[67] Note that the existential quantification imposed by *E* at the last step is vacuous, and serves mainly to convert the type from ⟨r,t⟩ to t.

express different contents, it can happen that some uses of a pair of sentences express mutually contradictory contents but others uses of the same sentences do not. In this example, the effect is due to the indexical interpretation of tense, but it is easily seen that the indexicality of personal pronouns will produce an analogous result.

3.7. Intensionality

Now we must deal with cases where an expression appears in an intensional context. In this case, we want it to denote what would otherwise be its content.

In many analyses of intensionality, it is assumed that the identification of a context as intensional or extensional is made on a purely linguistic basis: if an expression is in the scope of a modal operator, attitude predicate, or something similar, its context is intensional. That is, the relevant notion of context is that of *linguistic* context, not *pragmatic* context.

In our current grammar, we have been assigning contents to "uses" of expressions, rather than assigning contents to expressions relative to contexts. But each use of an expression, in any world where it occurs, is associated with a context. Contexts are assumed to provide a speaker, addressee, time, etc.; so they are clearly pragmatic contexts in some sense. But because each individual use of an expression occurs in its own context — different contexts for different constituents of the same (use of a) sentence — we may assume that the context of a given use of an expression will also provide information about its position in the larger syntactic structure. In other words, our notion of context is that of combined pragmatic-and-linguistic context.

Let us delay for a moment considering the issue of what makes one context intensional and another extensional. Assuming we have some basis for drawing the distinction, we may label a given use of an expression α as "intensional in u" or "extensional in u," according as its context in u is intensional or extensional. We may now stipulate intensional uses of lexical items denote what would customarily be their contents (cf. (42)):

(99) For any lexical item **α**, if USE(α, **α**), then:
  a. If α is extensional in u, then $[\![\alpha]\!]^{u,w} = Lex(\alpha, u, w)$;
  b. If α is intensional in u, then $[\![\alpha]\!]^{u,w} = \lambda w' . Lex(\alpha, u, w')$.

Next, we need to generalize our compositional rules (in (51)), to allow for the fact that expressions will now sometimes denote functions from possible worlds. We do this first by inductively defining an operation APPLY:[68]

(100) Case A: If $f \in \mathbf{D}_{\langle\tau,\upsilon\rangle}$ and $x \in \mathbf{D}_{\tau}$, then APPLY(f, x) = $f(x)$.
  Case B: (i) If $f \in \mathbf{D}_{\langle s,\tau\rangle}$ and there is no υ such that $x \in \mathbf{D}_{\langle s,\upsilon\rangle}$, then APPLY($f$, x) = [λw : w is a possible world . APPLY($f(w)$, x)]
    (ii) If $f \in \mathbf{D}_{\langle r,\tau\rangle}$ and there is no υ such that $x \in \mathbf{D}_{\langle r,\upsilon\rangle}$, then APPLY($f$, x) = [λr : r is a partial function from positive integers to $\mathbf{D}_e \cup \mathbf{D}_i$. APPLY($f(r)$, x)]

[68] In (100) τ, υ, and ω are variables over types, hence cannot be set to r or s (which appear in type labels, but are not types themselves).

Case C: (i) If $f \in \mathbf{D}_{\langle \tau,\upsilon \rangle}$ and $x \in \mathbf{D}_{\langle s,\omega \rangle}$, then APPLY($f$, x) = [λw : w is a possible world . APPLY(f, $x(w)$).

(ii) If $f \in \mathbf{D}_{\langle \tau,\upsilon \rangle}$ and $x \in \mathbf{D}_{\langle r,\omega \rangle}$, then APPLY($f$, x) = [λr : r is a partial function from positive integers to $\mathbf{D}_e \cup \mathbf{D}_i$. APPLY(f, $x(r)$)]

Case D: (i) If $f \in \mathbf{D}_{\langle s,\tau \rangle}$ and $x \in \mathbf{D}_{\langle s,\omega \rangle}$, then APPLY($f$, x) = [λw : w is a possible world . APPLY($f(w)$, $x(w)$)].

(ii) If $f \in \mathbf{D}_{\langle r,\tau \rangle}$ and $x \in \mathbf{D}_{\langle r,\omega \rangle}$, then APPLY($f$, x) = [λr : r is a partial function from positive integers to $\mathbf{D}_e \cup \mathbf{D}_i$. APPLY($f(r)$, $x(r)$)].

Intuitively, APPLY "skips over" any argument places for possible worlds or assignments of values to variables, filling *f*'s "outermost" argument place of the appropriate type with *x*.

Now we interpret complex expressions via the rule in (101):

(101) Where γ is a use of a complex expression consisting of α and β:
$[\![\gamma]\!]^{u,w}$ = APPLY($[\![\alpha]\!]^{u,w}$, $[\![\beta]\!]^{u,w}$) or APPLY($[\![\beta]\!]^{u,w}$, $[\![\alpha]\!]^{u,w}$), whichever is defined.

It may be useful to work through some examples. Assume that φ is a use of *E John*$_1$ *sleeps*$_2$ consisting of a use α of *E* and a use β of *John*$_1$ *sleeps*$_2$, which in turn consists of a use γ of *John*$_1$ and a use δ *sleeps*$_2$. Suppose in addition that γ and δ are intensional in *u* while α is extensional in *u*. Then:

(102) $[\![\gamma]\!]^{u,w}$ = $\lambda w'$. *Lex*(γ, *u*, *w*′)
= $\lambda w'$. $Ref_{u,\gamma}(1)$, provided that $Ref_{u,\gamma}(1)$ is named *John* in *u*.

Similarly:

(103) $[\![\delta]\!]^{u,w}$ = $\lambda w'$. *Lex*(δ, *u*, *w*′)
= $\lambda w'$. λr : $Ref_{u,\delta} \preccurlyeq r$ and $r(2) = time_{u,\delta}$. λx . *Lex*(BASE(δ), *u*, *w*′)(*x*)(*r*(2))
= $\lambda w'$. λr : $Ref_{u,\delta} \preccurlyeq r$ and $r(2) = time_{u,\delta}$. λx . *x* sleeps at *r*(2) in *w*′.

Because α is extensional in *u*:

(104) $[\![\alpha]\!]^{u,w}$ = *Lex*(α, *u*, *w*) = $\lambda f \in \mathbf{D}_{\langle r,t \rangle}$. $\exists r[Ref_{u,\alpha} \preccurlyeq r$ and $f(r) =$ **truth**]

By (101):

(105) $[\![\beta]\!]^{u,w}$ = APPLY($[\![\delta]\!]^{u,w}$, $[\![\gamma]\!]^{u,w}$)
= [$\lambda w'$. APPLY($[\![\delta]\!]^{u,w}(w')$, $[\![\gamma]\!]^{u,w}(w')$)]
= [$\lambda w'$. [λr : *r* is a partial function from positive integers to $\mathbf{D}_e \cup \mathbf{D}_i$.
APPLY($[\![\delta]\!]^{u,w}(w')(r)$, $[\![\gamma]\!]^{u,w}(w')$)]]
= [$\lambda w'$. [λr : $Ref_{u,\beta} \preccurlyeq r$ and $r(2) = time_{u,\delta}$ and $Ref_{u,\beta}(1)$ is named *John* in *u* . [λx . *x* sleeps at *r*(2) in *w*′]($Ref_{u,\beta}(1)$)]]
= [$\lambda w'$. [λr : $Ref_{u,\beta} \preccurlyeq r$ and $r(2) = time_{u,\delta}$ and $Ref_{u,\beta}(1)$ is named *John* in *u* . $Ref_{u,\beta}(1)$ sleeps at *r*(2) in *w*′]]

Similarly:

(106) $[\![\varphi]\!]^{u,w}$ = APPLY($[\![\alpha]\!]^{u,w}$, $[\![\beta]\!]^{u,w}$)
= $[\lambda w'$. APPLY($[\![\alpha]\!]^{u,w}$, $[\![\beta]\!]^{u,w}(w')$)]
= $[\lambda w'$. $\exists r[Ref_{u,\varphi} \preccurlyeq r$ and $[\![\beta]\!]^{u,w}(w')(r)$ = **truth**]]
= $[\lambda w'$. $\exists r[Ref_{u,\varphi} \preccurlyeq r$ and $Ref_{u,\varphi}(1)$ is named *John* in u and $r(2) = time_{u,\delta}$ and $Ref_{u,\varphi}(1)$ sleeps at $r(2)$ in w']]

In other words, in a context as given, this sentence denotes the set of worlds in which the individual named John intended by the speaker sleeps at the time of the context of the verb *sleeps*. Note that this is the *denotation* of the sentence in such a context, not its content.

Now let us consider how such a denotation may figure in the semantic composition of a larger sentence. A relevant case would be where the sentence appears as the scope of the adverb *necessarily*. We analyze *necessarily* in the usual way as universally quantifying over possible worlds:[69]

(107) If USE(α, *necessarily*), then *Lex*(α, u, w) = $\lambda p \in \mathbf{D}_{\langle s,t\rangle}$. For all possible worlds w', $p(w')$ = **truth**

It should be clear from this definition that if ψ is a use of the sentence *Necessarily E John*$_1$ *sleeps*$_2$ with φ as a constituent, then ψ expresses the content in world w with the truth condition $\forall w'\exists r[Ref_{u,\psi} \preccurlyeq r$ and $Ref_{u,\psi}(1)$ is named *John* in u and $r(2) = time_{u,\delta}$ and $Ref_{u,\psi}(1)$ sleeps at $r(2)$ in w'].

We obtained this reading on the assumption that the contexts of *John*$_1$ and *sleeps*$_2$ were intensional. What if their contexts are extensional? In that case, *E John*$_1$ *sleeps*$_2$ denotes **truth** relative to worlds u, w iff the individual intended by the speaker is named *John* in u and there is a time t prior to the time of the verb *sleeps*$_2$ such that that individual sleeps at t in w. (It denotes **falsity** if that individual does not sleep at t in w, and has no denotation if that individual is not named *John* in u.) But if its denotation is **truth** or **falsity**, it cannot combine sensibly with *necessarily*, which denotes a function applying to objects of type $\langle s, t\rangle$, not type t. No denotation will be assigned to such a use of the whole sentence *Necessarily E John*$_1$ *sleeps*$_2$. We may assume that some sort of denotation must be assigned to a sentence use in order for that use to be felicitous (and in particular that it must be assigned a denotation of type t for its content to be asserted), so it is clear that *Necessarily E John*$_1$ *sleeps*$_2$ will be anomalous if used in such a way that all its constituents are in extensional contexts.

In other words, the combinatorics of our type system already guarantee that in order to obtain a coherent reading, the scope of *necessarily* must be interpreted at least partly in an intensional context. We don't need to posit any special parameters like $speaker_{u,\alpha}$, $Ref_{u,\alpha}$, etc. to represent

[69] This interpretation of *necessarily*, in which it expresses unrestricted universal quantification over worlds, really only occurs in specialized philosophical discourse. A more adequate analysis would account for other interpretations as well, presumably treating them as particular cases of a more general pattern; but to develop such an analysis here would take us too far afield.

syntactic aspects of context. We simply assume that rational speakers will normally intend referents which result in coherent readings, and let the status of a context as intensional or extensional vary with the referential intentions of the speaker.

We should next consider the interaction of indexicality with modality. The crucial point, familiar from Kaplan (1989), is that indexical reference varies with context of use, but not world of evaluation. This pattern is preserved in the current system: Even if the pronoun *I*, for example, appears in an intensional context, a use φ of the sentence *Necessarily, E* I_1 $sleep_2$ is true (relative to *u*, *w*) iff for every possible world *w′*, the speaker of φ in *u* sleeps in *w′*. The derivation parallels that just given for *Necessarily, E John*$_1$ *sleeps*$_2$ exactly: Assume USE(γ, I_1), USE(δ, $sleep_2$), and γ and δ are both intensional in *u*. Then:

(108) $[\![\gamma]\!]^{u,w}$ $= \lambda w' . Lex(\gamma, u, w')$
$= \lambda w' . Ref_{u,\gamma}(1)$, provided that $Ref_{u,\gamma}(1) = speaker_{u,\gamma}$

(109) $[\![\delta]\!]^{u,w}$ $= \lambda w' . Lex(\delta, u, w')$
$= \lambda w' . \lambda r : Ref_{u,\delta} \preccurlyeq r$ and $r(2) = time_{u,\delta} . \lambda x . x$ sleeps at $r(2)$ in w'.

Now supposing β is the use of I_1 $sleep_2$ which consists of γ and δ:

(110) $[\![\beta]\!]^{u,w}$ = $[\lambda w' . [\lambda r : Ref_{u,\beta} \preccurlyeq r$ and $Ref_{u,\beta}(1) = speaker_{u,\beta}$ and $r(2) = time_{u,\delta} . Ref_{u,\beta}(1)$ sleeps at $r(2)$ in $w']]$

As before letting α be a use of *E* and φ consist of α and β:

(111) $[\![\varphi]\!]^{u,w} = [\lambda w' . \exists r[Ref_{u,\varphi} \preccurlyeq r$ and $Ref_{u,\varphi}(1) = speaker_{u,\varphi}$ and $r(2) = time_{u,\delta}$ and $Ref_{u,\varphi}(1)$ sleeps at $r(2)$ in $w']]$

And finally letting ψ consist of a use of *necessarily* and φ:

(112) $[\![\psi]\!]^{u,w}$ = **truth** if for all possible worlds w': $\exists r[Ref_{u,\psi} \preccurlyeq r$ and $Ref_{u,\psi}(1) = speaker_{u,\psi}$ and $r(2) = time_{u,\delta}$ and $Ref_{u,\psi}(1)$ sleeps at $r(2)$ in $w']$

Note that the function denoted by the pronoun *I* (relative to *u*, *w*) is constant — returning the same individual for every world (even though different people may be speaking in different worlds). The function denoted by *E* I_1 $sleep_2$ may yield different truth values in different worlds; but for any given world *w′*, that truth value depends on whether the speaker of this sentence in *u* sleeps in *w′*, not on whether the speaker in *w′* sleeps in *w′*.

Chapter 4 : Basic Relativist Semantics

In this chapter, we consider how to revise and extend the grammar presented in Chapter 3 to treat personal taste sentences as expressing contents whose truth value can vary from person to person. This initial presentation of a relativist semantic theory will deal only with simple cases — taste predicates appearing in main clauses, used in the ordinary way. More complex cases, including taste predicates in attitude contexts, or used in pragmatically unusual ways, will be considered in later chapters.

4.1. The judge parameter

In Lasersohn (2005), I presented an analysis of predicates of personal taste in which sentence contents were assigned truth values relative to an *individual* index, alongside the possible world index. The intuitive idea was that if a sentence was about a matter of taste rather than a matter of fact, its truth value could vary from person to person according to their tastes, and not just from world to world. For example, the sentence *This chili is tasty* might express a content (in a given context *c*) which is true relative to ⟨John, w_1⟩ but false relative to ⟨Mary, w_1⟩, meaning that in w_1, the chili is tasty according to John's tastes, but not according to Mary's. I will build on the same basic idea in this chapter, but we will have to alter the details of the formalization a little bit.

The system in Lasersohn (2005) was modeled closely on that of Kaplan (1989), with minimal adjustments to allow variations in truth value from individual to individual (in the case of sentences about matters of taste). An interesting feature of Kaplan's system, which often seems forgotten in subsequent discussion, is that it appeals to context in two distinct ways in assigning truth values to sentences. First, context plays a role in the assignment of contents to linguistic expressions. A given expression may have different contents in different contexts, so we may think of context as what has to be supplied (together with character) to determine what content is being expressed. Second, in Kaplan's system, context plays a role in the assignment of truth values to sentence contents — or more generally, in the assignment of denotations to contents of all types. Modeling the contents of sentences as functions from world-time pairs to truth values, Kaplan defines a notion of truth-in-context such that a given sentence **φ** is true in *c* iff the content of **φ** maps the pair ⟨$time_c$, $world_c$⟩ onto **truth** (Kaplan (1989) p. 547).[70] The context supplies a particular time and world, which are used not in determining what content is expressed, but what truth value is assigned to that content. In Lasersohn (2005) I suggested that we could regard each context as supplying an individual, called the *judge* of the context, and consider a sentence **φ** to be true relative to a context *c* iff the content of *c* mapped the triple ⟨$judge_c$, $time_c$, $world_c$⟩ onto **truth**.[71] The judge parameter of contexts played no role in the determination of content, but only in the assignment of denotations to contents; so the content of a sentence about a matter of taste was treated as "judge neutral" — the same for everyone — even though the truth value could vary from person to person.

---

[70] More technically, and using Kaplan's notation, φ *is true in the context c* (in the structure $\mathfrak{A}$) iff for every assignment *f*, $\{\varphi\}^{\mathfrak{A}}_{c,f}(c_T, c_W)$ = Truth. Here, the curly braces {} play a role similar to our round double brackets ⦇⦈.

[71] In the article I followed Kaplan and other authors in assuming sentence contents were true or false relative to times, in order to account for tense. Given the concerns outlined in Section 3.4. above, this use of a time index no longer seems appropriate. However, we will still require a time index to account for the fact that people's tastes can change over time, as discussed below.

Relativizing the truth values of sentence contents to individuals in this way is not, by itself, enough to get the desired semantic effect. If we assume that for each use of an expression, there is a matter of fact about what the context of that use is, and that each context supplies a judge, then for each use there will be a matter of fact about who the judge is. In such a system, a sentence as used in context *c* could be regarded as true simpliciter iff it is true relative to the judge (and time and world) of *c*. But the point of introducing the judge parameter was to draw a semantic distinction between sentences whose truth values depend purely on matters of fact and sentences whose truth values depend on matters of taste; only the former should have truth values simpliciter, while the latter should have truth values only relative to the judge parameter.

Therefore, I suggested in Lasersohn (2005), we should give up the assumption that there is a matter of fact about what the context is, in which a given use of an expression occurs. Rather, for any concrete situation in which a sentence is used, there are as many different "contexts" as there are individuals who might judge whether or not that use was truthful. The formal notion of context as represented in the grammatical theory was thus distinguished from the intuitive idea of the practical environment in which an expression is used; the practical environment underdetermined the formal context. Contextual parameters other than the judge were assumed to be fixed by matters of fact, of course; so the connection between the context and matters of fact about the practical environment was not entirely severed. Sentences that were purely about matters of fact could be distinguished from sentences about matters of taste in that their contents did not vary in truth value among contexts which differed only in the value of the judge parameter.

This technique served its purpose adequately, but it cannot be incorporated into our current system without some changes. The reason is that we have been assigning contents to "uses" of linguistic expressions, and thinking of each use as having a unique context in each world where it occurs.[72] If each context supplies a judge, then each use will have (in any world where it occurs) a particular judge specified as the judge of its context in that world, and we will have returned to a system in which each sentence content can be assigned a definite truth value, regardless of whether it is about matters of fact or matters of taste. If the context of φ in $w_@$ supplies *j* as the judge and *t* as the time, for example, we could easily say that φ is true (simpliciter) iff $\text{⟨}\varphi\text{⟩}^{w_@}(j, t, w_@)$ = **truth**; but then even sentences whose truth values vary with the judge will be assigned truth values simpliciter — that is, "absolutely" and not just relatively.

We could maintain this strategy but still recover a relativistic effect by distinguishing between "uses" in the technical sense of pairs of an expression and a context and the uses in the intuitive sense of the actual employment of an expression by a particular speaker for a particular purpose on a particular occasion, much as we distinguished contexts in the technical sense of a specification of a time, world, speaker, addressee, judge, etc., from contexts in the intuitive sense of the practical environment in which an utterance is made; but as more and more of our technical concepts become separated from the intuitive ideas they were introduced to model, one has the feeling that something has gone awry.

I think a better response would be to recognize that "judging" whether a use of a sentence is true is a separate matter from using the sentence; and that such judgment may be based in part on contexts which are different from the contexts in which the sentence is used. That is, we may articulate each formal context in the sense of Lasersohn (2005) into two parts, corresponding to the situation in which an expression is used, and the situation in which a truth value is "judged." In

[72] See the discussion on p. 13.

sentences about matter of taste, the truth value may vary with both parts.

This idea of relativizing truth assignment to two contexts had already been argued at the time of my 2005 paper by MacFarlane (2003), and since then has been further developed in considerable detail by MacFarlane and others. In MacFarlane's terminology, the two contexts are called the *context of use* and *context of assessment*; let us adopt this terminology here. Intuitively, the judge parameter of Lasersohn (2005) is a feature of the context of assessment, not the context of use; this is the reason why a series of parameter values including the judge cannot be regarded as determined by the concrete situation in which an utterance is made.

With this idea in mind, we may continue to regard each use of an expression as having a unique context of use in each world where it occurs, but need not distinguish contexts of use any more finely than facts about the situation in which that use takes place. We continue to regard each context of use as supplying a speaker, addressee, time, sets of familiar and relevant objects, reference assignment, etc.

Contexts of assessment, on the other hand, must supply at least an individual (the judge), a time, and a world; there is no need regard them as supplying a speaker, addressee, etc. Of course someone assessing a sentence may need to know who the speaker, addressee, etc. were in order to assess it properly; but this just means that the assessor must have some knowledge of the context of use, not that the context of assessment provides its own values for these parameters, independently of the context of use.

The terms "judge" and "assessment" carry some risk of misunderstanding. It would be natural to get the impression from these terms that a sentence content Φ counts as true relative to a context of assessment *c* iff the judge of *c* considers Φ in *c* and comes to the conclusion that Φ is true. This is not the intended idea, however.

As discussed in Section 2.3. above, we each operate from a position of epistemic privilege with respect to our own tastes; but it is still possible to be mistaken in judgment about matters of taste. If, for example, John misremembers the flavor of elderberries, he might consider the claim that elderberries are tasty and come to the conclusion that this claim is false, even though elderberries are actually delicious according to his standard of taste. In this situation, we should regard the sentence *Elderberries are tasty* as expressing a content which is *true* relative to a context of assessment in which John is the judge, even though his conclusion is that this content is *false*. The truth value of a sentence content relative to an individual is not necessarily the truth value that individual believes it to have; it is the truth value which is congruent with that individual's standard of taste. To ask whether a sentence content is true relative to a context of assessment *c* is *not* to ask what the result is of the judge of *c* performing some psychological act of assessment in *c*. Contrary to what the terminology suggests, there is no requirement that the judge even perform an act of assessment in *c*.

To put the matter differently, the strategy we are developing to account for personal taste sentences is to say that their truth values depend on the value of a parameter which is left indeterminate by the context in which the sentence is used; it is this indeterminacy which gives the effect that such sentences are true or false only subjectively. If we resolve the indeterminacy and fix a particular value for this parameter, a truth value may be assigned on a completely objective basis. To say that *Elderberries are tasty* is true relative to a context of assessment with John as the judge is to say that it is objectively true that elderberries are tasty for John; this is no more to be identified with John's believing that elderberries are tasty than any other objective claim Φ should be identified with John's believing Φ.

4.2. Revising the grammar

We now turn to the task of revising the grammar sketched in Chapter 3 so that we can add predicates of personal taste to it. The primary changes to be made are: (1) denotations will be assigned relative not just to a possible world index, but to a world index and a "perspective" index, where each perspective index itself is identified with an ordered triple of an individual, a time and a world; (2) contents accordingly will be identified not with functions mapping worlds onto denotations, but with functions mapping world-perspective pairs (or equivalently, world-individual-time-world quadruples) onto denotations; (3) contexts of assessment will be distinguished from contexts of use, and will supply perspectives to serve as arguments to contents in order to derive denotations, including truth values.

To facilitate these revisions, it will be useful to introduce some notation to distinguish contexts of use from contexts of assessment. Let us use *c* as a metalanguage variable over contexts of assessment. We write $judge_c$, $time_c$ and $world_c$ for the parameter values supplied by context of assessment *c*.[73]

The next step is to update our type theory. Previously, the denotation domains for types beginning with "s" were sets of functions from the set of possible worlds; we now revise this to the set of world-perspective pairs (cf. (56), (85)):

(113) a. $\mathbf{D}_e = \{x \mid x \text{ is a group or individual}\}$;
b. $\mathbf{D}_t = \{\mathbf{truth}, \mathbf{falsity}\}$;
c. $\mathbf{D}_i = \{x \mid x \text{ is a time}\}$;
d. $\mathbf{D}_l = \{l \mid \exists x \in \mathbf{D}_e\ \exists I \in \mathbf{D}_i\ \exists w[w \text{ is a possible world}] .\ l \subseteq space_{x,I,w}\}$
e. If τ is a type, then $\mathbf{D}_{\langle s,\tau\rangle} = \mathbf{D}_\tau^{\{\langle w, \langle x, t, w'\rangle\rangle \mid x \text{ is an individual, } t \text{ is a time, and } w, w' \text{ are possible worlds}\}}$;
f. If τ is a type, then $\mathbf{D}_{\langle r,\tau\rangle} = \mathbf{D}_\tau^{\{r \mid r \text{ is a partial function from positive integers to } \mathbf{D}_e \cup \mathbf{D}_i\}}$;
g. If τ is a type and υ is a type, then $\mathbf{D}_{\langle \tau,\upsilon\rangle} = (\mathbf{D}_\upsilon)^{\mathbf{D}\tau}$.

With these changes we will also need to update our notation. Where before we wrote '$Lex(\alpha, u, w)$' we will now write '$Lex(\alpha, u, w, p)$'; where before we wrote '$[\![\alpha]\!]^{u,w}$' we will now write '$[\![\alpha]\!]^{u,w,p}$'. We understand *p* to range over individual-time-world triples, so we will sometimes use notation like '$[\![\alpha]\!]^{u,w,\langle x,t,a\rangle}$' (where *x* is an individual, *t* is a time, and *a* is a world — the "world of assessment"). Of course we regard '$[\![\alpha]\!]^{u,w,p}$' as an abbreviation for '$[\![\ulcorner\alpha\urcorner^u]\!]^{w,p}$', adapting our earlier convention to the new indexation system.

Now we must revise our existing lexical entries. The necessary revisions are simple notational changes; for these words the new perspective index acts as a "third wheel" — present, but with no significant effect.

(114) a. If USE(α, $John_i$) then $Lex(\alpha, u, w, p) = Ref_{u,\alpha}(i)$, provided that $Ref_{u,\alpha}(i)$ is named *John* in *u*.

[73] Although contents are indexed to *two* possible world parameters — the ordinary one we have been using all along, and the world parameter which serves to distinguish perspectives — we need only regard the context of assessment as specifying a single world. We will need to distinguish the two world parameters to give an adequate account of modal operators, not because of contextual effects on truth values.

b. If USE(α, $Mary_i$) then $Lex(\alpha, u, w, p) = Ref_{u,\alpha}(i)$, provided that $Ref_{u,\alpha}(i)$ is named *Mary* in u.
c. If USE(α, $The\ Giant\ Dipper_i$) then $Lex(\alpha, u, w, p) = Ref_{u,\alpha}(i)$, provided that $Ref_{u,\alpha}(i)$ is named *The Giant Dipper* in u.
(Undefined otherwise)

(115) If USE(α, $\boldsymbol{\alpha_i}$) and $\boldsymbol{\alpha_i}$ is a pronoun:

a. If $Ref_{u,\alpha}(i)$ is defined and $Ref_{u,\alpha}$ is a feature-respecting assignment with respect to u, α, then $Lex(\alpha, u, w, p) = Ref_{u,\alpha}(i)$.

b. If $Ref_{u,\alpha}(i)$ is not defined, then $Lex(\alpha, u, w, p) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r$ and r is a feature-respecting assignment with respect to u, α . $r(i)]$

(116) If USE(α, $puppy_{i,j}$) then $Lex(\alpha, u, w, p) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r$ and $r(j)$ is a time and $r(i)$ is a puppy at $r(j)$ in w . $r(i)]$. (Likewise for other common nouns)

(117) If USE(α, a) then $Lex(\alpha, u, w, p) = [\lambda x : x$ is an individual . $x]$

(118) If USE(α, E) then $Lex(\alpha, u, w, p) = [\lambda f \in \mathbf{D}_{\langle r,t\rangle} . \exists r[Ref_{u,\alpha} \preccurlyeq r$ and $f(r) = \mathbf{truth}]]$

(119) If USE(α, e_i) then $Lex(\alpha, u, w, p) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r. r(i)]$

(120) If USE(α, *every*) then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle r,e\rangle} . \lambda g : g \in \mathbf{D}_{\langle r,t\rangle}$. For all r, if $Ref_{u,\alpha} \preccurlyeq r$ and $f(r)$ is defined, there is at least one r' such that $r \preccurlyeq r'$ and $g(r') = \mathbf{truth}]$

(121) a. If USE(α, the_{sg}) then $Lex(\alpha, u, w, p) = [\lambda f \in \mathbf{D}_{\langle r,e\rangle}$. There exists some x such that for all r, if $Ref_{u,\alpha} \preccurlyeq r$ and $f(r)$ is defined, then $x = f(r) . f(r')]$, where r' is any arbitrary extension of $Ref_{u,\alpha}$ such that $f(r')$ is defined.

b. If USE(α, the_{pl}) then $Lex(\alpha, u, w, p) = [\lambda f \in \mathbf{D}_{\langle r,e\rangle}$. There exists some x such that for all r, if $Ref_{u,\alpha} \preccurlyeq r$ and $f(r)$ is defined, then x includes $f(r) . f(r')]$, where r' is any arbitrary extension of $Ref_{u,\alpha}$ such that $f(r')$ is defined.

(122) a. If USE(α, $sleep_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda y : y \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . y$ sleeps at t in $w]$

b. If USE(α, see_{inf}) then $Lex(\alpha, u, w, p) = [\lambda y : y \in \mathbf{D}_e . \lambda z : z \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . z$ sees y at t in $w]$

(123) If USE(α, $\boldsymbol{\alpha_i}$) and $\boldsymbol{\alpha_i}$ is a verb of type $\langle\tau_1,\langle\ldots,\langle\tau_n,\langle i,t\rangle\rangle\rangle\rangle$, then:

a. If $\boldsymbol{\alpha_i}$ is of category PAST, then $Lex(\alpha, u, w, p) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r$ and $r(i) < time_{u,\alpha} . \lambda y_1 \in \mathbf{D}_{\tau_1} \ldots \lambda y_n \in \mathbf{D}_{\tau_n} . Lex(\text{BASE}(\alpha), u, w, p)(y_1)\ldots(y_n)(r(i))]$

c. If $\boldsymbol{\alpha_i}$ is of category PRES, then $Lex(\alpha, u, w, p) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r$ and $r(i) = time_{u,\alpha} . \lambda y_1 \in \mathbf{D}_{\tau_1} \ldots \lambda y_n \in \mathbf{D}_{\tau_n} . Lex(\text{BASE}(\alpha), u, w, p)(y_1)\ldots(y_n)(r(i))]$

(124) a. If USE(α, $here_{i,j,k}$) then $Lex(\alpha, u, w, p) = \lambda r$: $Ref_{u,\alpha} \preccurlyeq r$ & $Ref_{u,\alpha}(i) \subseteq proximity_{u,\alpha}$ & $\exists t\, r(j) \subseteq space_{r(k),t,w}$ & $Ref_{u,\alpha}(i)$ represents $r(j)$. $Ref_{u,\alpha}(j)$

b. If USE(α, $there_{i,j,k}$) then $Lex(\alpha, u, w, p) = \lambda r$: $Ref_{u,\alpha} \preccurlyeq r$ & $Ref_{u,\alpha}(i)$ does not overlap with $proximity_{u,\alpha}$ & $\exists t\, r(j) \subseteq space_{r(k),t,w}$ & $Ref_{u,\alpha}(i)$ represents $r(j)$. $Ref_{u,\alpha}(j)$

(125) If USE(α, *in front of*$_{i,j}$) then $Lex(\alpha, u, w, p) = \lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda x \in \mathbf{D}_e . \{l \in space_{r(i),r(j),w} \mid l$ is near $loc_{r(i),r(j),w}(x, r(j), w)$; and if x has an intrinsic front at $r(j)$ in w, then l is nearer to the intrinsic front of x at $r(j)$ in w than to any other side of x; else l is between $loc_{r(i),r(j),w}(x, r(j), w)$ and $loc_{r(i),r(j),w}(r(i), r(j), u)$ at $r(j)$ in $w\}$

(126) If USE(α, $leave_{infl_{i,j}}$) then $Lex(\alpha, u, w) = \lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda x : x \in \mathbf{D}_e . \lambda l : l \in \mathbf{D}_l . \lambda z : z \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . loc_{r(i),r(j),w}(x, t, w) \subseteq l$ and there exists some $t' < t$ such that for all t'', if $t' < t'' < t$, then $loc_{r(i),r(j),w}(z, t'', w)$ is near l, and $loc_{r(i),r(j),w}(z, t, w)$ is not near l

(127) a. If USE(α, $don't_{1sg_i}$) or USE(α, $don't_{2sg_i}$) or USE(α, $doesn't_i$) or USE(α, $don't_{pl_i}$) then $Lex(\alpha, u, w, p) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda f : f \in \mathbf{D}_{\langle i,\langle e,t\rangle\rangle} . \lambda y : y \in \mathbf{D}_e$. There is no r' such that $Ref_{u,\alpha} \preccurlyeq r'$ and $r'(i) = time_{u,\alpha}$ and $f(y)(r'(i)) =$ **truth**]

b. If USE(α, $didn't_i$) then $Lex(\alpha, u, w, p) = [\lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda f : f \in \mathbf{D}_{\langle i,\langle e,t\rangle\rangle} . \lambda y : y \in \mathbf{D}_e$. There is no r' such that $Ref_{u,\alpha} \preccurlyeq r'$ and $r'(i)$ is a time and $r'(i) < time_{u,\alpha}$ and $f(y)(r'(i)) =$ **truth**]

The assignment of denotations based on the status of a context as intensional or extensional must also be revised in accordance with our new notation and indices (cf. (99)):

(128) For any lexical item $\boldsymbol{\alpha}$, if USE(α, $\boldsymbol{\alpha}$), then:
a. If α is extensional in u, then $[\![\alpha]\!]^{u,w,p} = Lex(\alpha, u, w, p)$;
b. If α is intensional in u, then $[\![\alpha]\!]^{u,w,p} = \lambda\langle w', p'\rangle . Lex(\alpha, u, w', p')$.

Of course we now understand the content of an expression to be a function which maps any world-perspective pair to its denotation relative to that world and perspective:

(129) For any use α: $\text{⟨}\alpha\text{⟩}^u = \lambda\langle w, p\rangle . [\![\alpha]\!]^{u,w,p}$

Our definition of APPLY and compositional rule must also be revised to reflect our new assumptions about indices and contexts (cf. (100), (101)):

(130) Case A: If $f \in \mathbf{D}_{\langle\tau,\upsilon\rangle}$ and $x \in \mathbf{D}_\tau$, then APPLY(f, x) = $f(x)$.
Case B: (i) If $f \in \mathbf{D}_{\langle s,\tau\rangle}$ and there is no υ such that $x \in \mathbf{D}_{\langle s,\upsilon\rangle}$, then APPLY($f$, x) = $[\lambda\langle w, p\rangle : w$ is a possible world and p is a perspective . APPLY($f(w,p)$, x)]

(ii) If $f \in \mathbf{D}_{\langle r,\tau\rangle}$ and there is no υ such that $x \in \mathbf{D}_{\langle r,\upsilon\rangle}$, then APPLY($f$, x) = [λr : r is a partial function from positive integers to $\mathbf{D}_e \cup \mathbf{D}_i$. APPLY($f(r)$, x)]

Case C: (i) If $f \in \mathbf{D}_{\langle \tau,\upsilon\rangle}$ and $x \in \mathbf{D}_{\langle s,\omega\rangle}$, then APPLY($f$, x) = [$\lambda\langle w, p\rangle$: w is a possible world and p is a perspective . APPLY(f, $x(w, p)$).

(ii) If $f \in \mathbf{D}_{\langle \tau,\upsilon\rangle}$ and $x \in \mathbf{D}_{\langle r,\omega\rangle}$, then APPLY($f$, x) = [λr : r is a partial function from positive integers to $\mathbf{D}_e \cup \mathbf{D}_i$. APPLY(f, $x(r)$)]

Case D: (i) If $f \in \mathbf{D}_{\langle s,\tau\rangle}$ and $x \in \mathbf{D}_{\langle s,\omega\rangle}$, then APPLY($f$, x) = [$\lambda\langle w, p\rangle$: w is a possible world and p is a perspective . APPLY($f(w, p)$, $x(w, p)$)].

(ii) If $f \in \mathbf{D}_{\langle r,\tau\rangle}$ and $x \in \mathbf{D}_{\langle r,\omega\rangle}$, then APPLY($f$, x) = [λr : r is a partial function from positive integers to $\mathbf{D}_e \cup \mathbf{D}_i$. APPLY($f(r)$, $x(r)$)].

(131) Where γ is a use of a complex expression consisting of α and β:
$[\![\gamma]\!]^{u,w,p}$ = APPLY($[\![\alpha]\!]^{u,w,p}$, $[\![\beta]\!]^{u,w,p}$) or APPLY($[\![\beta]\!]^{u,w,p}$, $[\![\alpha]\!]^{u,w,p}$), whichever is defined.

4.3. Extension to taste expressions

Now that we have made the necessary background revisions, we can begin extending the grammar to include expressions of personal taste, specifically, the verb *stink* and the adjectives *fun* and *tasty*.[74] The first step is to enlarge our set of syntactic categories to include adjectives. At this point we will only include adjectives in predicate (postcopular) position, and not adjectives in attributive position. We also set aside for the moment the fact that (most) adjectives are scalar.[75] We add A to our set of basic syntactic categories and stipulate that *tasty* ∈ A and *fun* ∈ A.

(132) a. If USE(α, *tasty*) then *Lex*(α, u, w, $\langle x, t, a\rangle$) = [$\lambda y : y \in \mathbf{D}_e$. $\lambda t' \in \mathbf{D}_i$. y is tasty at t' in w by the standards of x at t in a]

b. If USE(α, *fun*) then *Lex*(α, u, w, $\langle x, t, a\rangle$) = [$\lambda y : y \in \mathbf{D}_e$. $\lambda t' \in \mathbf{D}_i$. y is fun at t' in w by the standards of x at t in a]

The various forms of the verb *stink* belong to the expected categories, and are defined analogously to *tasty* and *fun*:

(133) a. $stink_{inf} \in$ (ACC\(V∩INF))

b. $stink_{1sg_i} \in$ ((NOM∩1P∩SG)\V) ∩ PRES

c. $stink_{2sg_i} \in$ ((NOM∩2P∩SG)\V) ∩ PRES

d. $stinks_i \in$ ((NOM∩3P∩SG)\V) ∩ PRES

e. $stink_{pl_i} \in$ ((NOM∩PL)\V) ∩ PRES

f. $stank_i \in$ (NOM\V) ∩ PAST

[74] *Fun* also has a use as a noun, and in the idiolects of conservative speakers, is always a noun and never an adjective. (Such speakers find it ungrammatical to say *This is as fun as that*, for example, preferring instead *This is as much fun as that*.) I take for granted here that the adjectival use of *fun*, in those varieties of English where it occurs, should be analyzed as a normal, grammatical usage, and not as an error.

[75] We will return to scalarity briefly in Section 9.4.

(134) If USE(α, $stink_{inf}$) then $Lex(\alpha, u, w, \langle x, t, a\rangle)$ = [$\lambda y : y \in \mathbf{D}_e . \lambda t' \in \mathbf{D}_i . y$ stinks at t' in w by the standards of x at t in a]

We need to add various forms of the copula to our grammar:

(135) a. $am_i \in (((\text{NOM}\cap 1\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PRES}$
b. $are_{2sg_i} \in (((\text{NOM}\cap 2\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PRES}$
c. $is_i \in (((\text{NOM}\cap 3\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PRES}$
d. $are_{pl_i} \in (((\text{NOM}\cap\text{PL})\backslash\text{V})/\text{A}) \cap \text{PRES}$
e. $was_{1sg_i} \in (((\text{NOM}\cap 1\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PAST}$
f. $were_{2sg_i} \in (((\text{NOM}\cap 2\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PAST}$
g. $was_{3sg_i} \in (((\text{NOM}\cap 3\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PAST}$
h. $were_{pl_i} \in (((\text{NOM}\cap\text{PL})\backslash\text{V})/\text{A}) \cap \text{PAST}$
i. $am\ not_i \in (((\text{NOM}\cap 1\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PRES}$
j. $aren't_{2sg_i} \in (((\text{NOM}\cap 2\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PRES}$
k. $isn't_i \in (((\text{NOM}\cap 3\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PRES}$
l. $aren't_{pl_i} \in (((\text{NOM}\cap\text{PL})\backslash\text{V})/\text{A}) \cap \text{PRES}$
m. $wasn't_{1sg_i} \in (((\text{NOM}\cap 1\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PAST}$
n. $weren't_{2sg_i} \in (((\text{NOM}\cap 2\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PAST}$
o. $wasn't_{3sg_i} \in (((\text{NOM}\cap 3\text{P}\cap\text{SG})\backslash\text{V})/\text{A}) \cap \text{PAST}$
p. $weren't_{pl_i} \in (((\text{NOM}\cap\text{PL})\backslash\text{V})/\text{A}) \cap \text{PAST}$

(136) a. If USE(α, am_i) or USE(α, are_{2sg_i}) or USE(α, is_i) or USE(α, are_{pl_i}) then $Lex(\alpha, u, w, p)$ = [$\lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda f : f \in \mathbf{D}_{\langle e,\langle i,t\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . r(i) = time_{u,\alpha}$ and $f(x)(r(i))$ = **truth**]

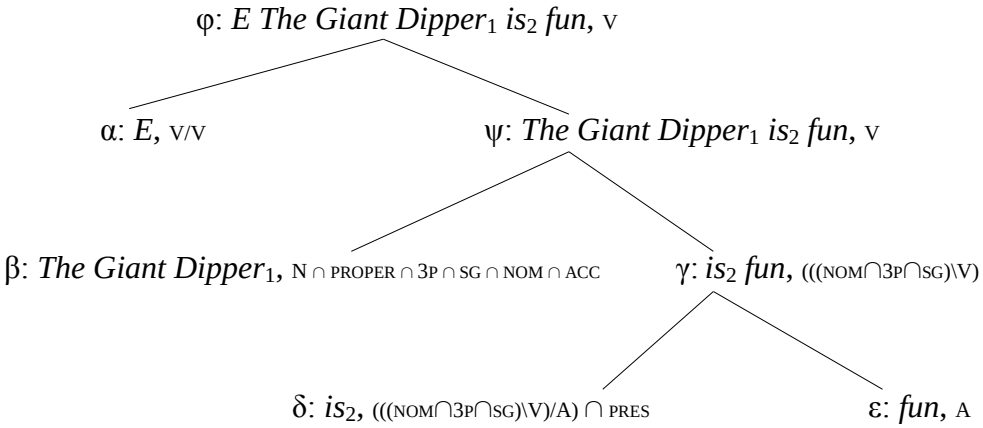
b. If USE(α, was_{1sg_i}) or USE(α, $were_{2sg_i}$) or USE(α, was_{3sg_i}) or USE(α, $were_{pl_i}$) then $Lex(\alpha, u, w, p)$ = [$\lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda f : f \in \mathbf{D}_{\langle e,\langle i,t\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . r(i) < time_{u,\alpha}$ and $f(x)(r(i))$ = **truth**]

c. If USE(α, $am\ not_i$) or USE(α, $aren't_{2sg_i}$) or USE(α, $isn't_i$) or USE(α, $aren't_{pl_i}$) then $Lex(\alpha, u, w, p)$ = [$\lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda f : f \in \mathbf{D}_{\langle r,\langle e,\langle i,t\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e$. There is no r' such that $Ref_{u,\alpha} \preccurlyeq r'$ and $r'(i) = time_{u,\alpha}$ and $f(r')(x)(r'(i))$ = **truth**]

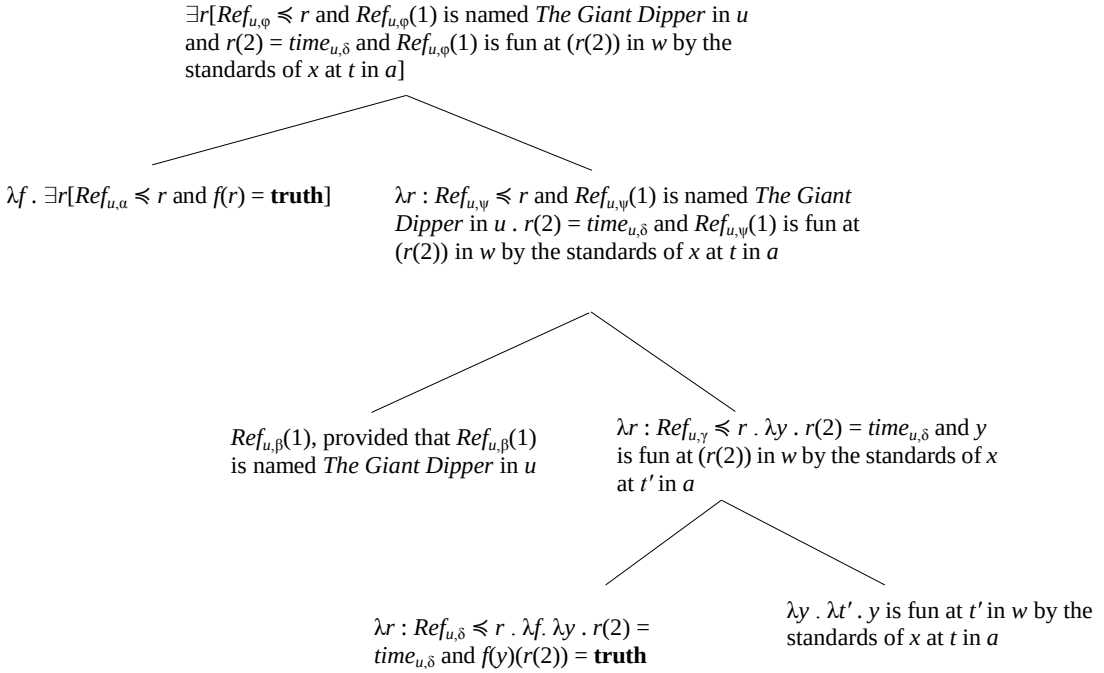
d. If USE(α, $wasn't_{1sg_i}$) or USE(α, $weren't_{2sg_i}$) or USE(α, $wasn't_{3sg_i}$) or USE(α, $weren't_{pl_i}$) then $Lex(\alpha, u, w, p)$ = [$\lambda r : Ref_{u,\alpha} \preccurlyeq r . \lambda f : f \in \mathbf{D}_{\langle r,\langle e,\langle i,t\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e$. There is no r' such that $Ref_{u,\alpha} \preccurlyeq r'$ and $r'(i) < time_{u,\alpha}$ and $f(r')(x)(r'(i))$ = **truth**]

We can now derive sentences like *John is fun* as in (137), with denotations assigned relative to worlds u, w and perspective $p = \langle x, t, a\rangle$ as shown in (138):

(137)



(138)



4.4. Truth in a relativist semantics

We still need to relate the assignment of denotations relative to an arbitrary world and perspective to the notion of truth relative to a context of assessment. For contents, this is trivial: we let a sentence content be true relative to a world *w* and context of assessment *c* iff it is true relative to *w* and the perspective consisting of the judge of *c*, time of *c* and world of *c*:

(139) a. Φ is true relative to a world *w* and context of assessment *c* iff $\Phi(w, \langle judge_c, time_c, world_c \rangle)$ = **truth**;
b. Φ is false relative to *w* and *c* iff $\Phi(w, \langle judge_c, time_c, world_c \rangle)$ = **falsity**.

We can define truth relative to contexts of assessment for *uses* of sentences relative to triples of a world of use, a world of modal evaluation, and a context of assessment, and define truth for *sentences* relative to quadruples of a use, a world of use, a world of modal evaluation, and a context of assessment:

(140) a. φ is true in world *u* relative to world *w* and context of assessment *c* iff $[\![\varphi]\!]^{u,w,\langle judge_c,\, time_c,\, world_c\rangle}$ = **truth**;
b. φ is false in *u* relative to *w* and *c* iff $[\![\varphi]\!]^{u,w,\langle judge_c,\, time_c,\, world_c\rangle}$ = **falsity**.

(141) a. **φ** is true relative to a use φ, in world *u*, relative to world *w* and context of assessment *c* iff USE(φ, **φ**) and φ is true in *u* relative to *w* and *c*;

b. **φ** is false relative to φ, in *u*, relative to *w* and *c* iff USE(φ, **φ**) and φ is false in *u* relative to *w* and *c*.

We can also define monadic (or "absolute") truth for sentence contents by universally quantifying on the perspective indices, and fixing the modal evaluation world index to the actual world $w_@$:

(142) a. Φ is true iff for all individuals *x*, times *t*, and worlds *a*: $\Phi(w_@, \langle x, t, a\rangle)$ = **truth**;
b. Φ is false iff for all individuals *x*, times *t*, and worlds *a*: $\Phi(w_@, \langle x, t, a\rangle)$ = **falsity**.

Modadic truth can also be defined for sentence uses:

(143) a. φ is true iff for all individuals *x*, times *t*, and worlds *a*: $[\![\varphi]\!]^{w_@,w_@,\langle x,\, t,\, a\rangle}$ = **truth**;
b. φ is false iff for all individuals *x*, times *t*, and worlds *a*: $[\![\varphi]\!]^{w_@,w_@,\langle x,\, t,\, a\rangle}$ = **falsity**.

We can define truth for sentences relative to uses (with no relativization to contexts of assessment) by appealing to monadic truth for uses:

(144) a. **φ** is true relative to a use φ iff USE(**φ**, φ) and φ is true;
b. **φ** is false relative to a use φ iff USE(**φ**, φ) and φ is false.

It should be noticed that the notions of truth and falsity defined in (142), (143) and (144) only apply in certain cases. If a sentence contains *fun* or *tasty*, its content may be true relative to some contexts of assessment but false relative to others. But then it will not be true or false *tout court*. In contrast, a sentence like *John sleeps* has no constituents whose denotations vary with the perspective. For any worlds *u*, *w* and perspectives *p*, *p'*, $[\![John_i\ sleeps_j]\!]^{u,w,p} = [\![John_i\ sleeps_j]\!]^{u,w,p'}$; therefore this sentence expresses a content which is (absolutely) true or (absolutely) false.

Our grammar thus allows us multiple notions of truth (at the various levels such as contents, uses, and sentences), including both relativistic notions of truth which apply to all examples, and absolute notions of truth that apply only in examples which are strictly about matters of fact, not matters of taste.

4.5. Derelativization in the object language

Although the presence of expressions like *fun* or *tasty* may allow a sentence to be true relative to some contexts of assessment and false relative to others (hence to lack an absolute truth value), some sentences will have absolute truth values even though they contain such expressions. For example, if a sentence contains *fun* or *tasty* with a prepositional phrase headed by *for* or *to*, it loses its subjective "flavor": even though *The Giant Dipper is fun* might be true relative to John but false relative to Mary, *The Giant Dipper is fun for John* should have the same truth value relative to everyone.

We can obtain this effect by defining (this use of) *for* in such a way that any use of a sentence *x is P for y* is true relative to $u, w, \langle z, t, a\rangle$ iff *x is P* is true relative to $u, w, \langle y, t, a\rangle$.[76] In other words, we treat *for* as a kind of modal operator, but rather than quantifying on the ordinary possible worlds index, it quantifies on the individual part of the perspective index.[77]

Implementing this idea is a straightforward adaptation of whatever formalism is used for intensional operators; for example in Lasersohn (2005), I effectively[78] just stipulated that $[\![\alpha \textit{ for } \beta]\!]^{u,w,\langle x,t,a\rangle} = [\![\alpha]\!]^{u,w,\langle [\![\beta]\!]^{u,w,\langle x,t,a\rangle},t,a\rangle}$; that is, that α *for* β denotes, relative to *x*, whatever α denotes relative to the denotation of β. In our present, system, however, we must give a slightly different formulation in order to respect our assumptions about compositionality, which invoke a Fregean reference shift rather than letting the modal operator directly "see" the denotations its scope has relative to a variety of indices. We may define *for* as follows:

(145) a. $\textit{for} \in (A\backslash A)/N$

b. $\textit{Lex}(\textit{for}, u, w, \langle x, t, a\rangle) = \lambda z \in \mathbf{D}_{u,e}\, \lambda f \in \mathbf{D}_{u,\langle s,\langle e,\langle i,t\rangle\rangle\rangle}\, .\, f(w, \langle z, t, a\rangle)$

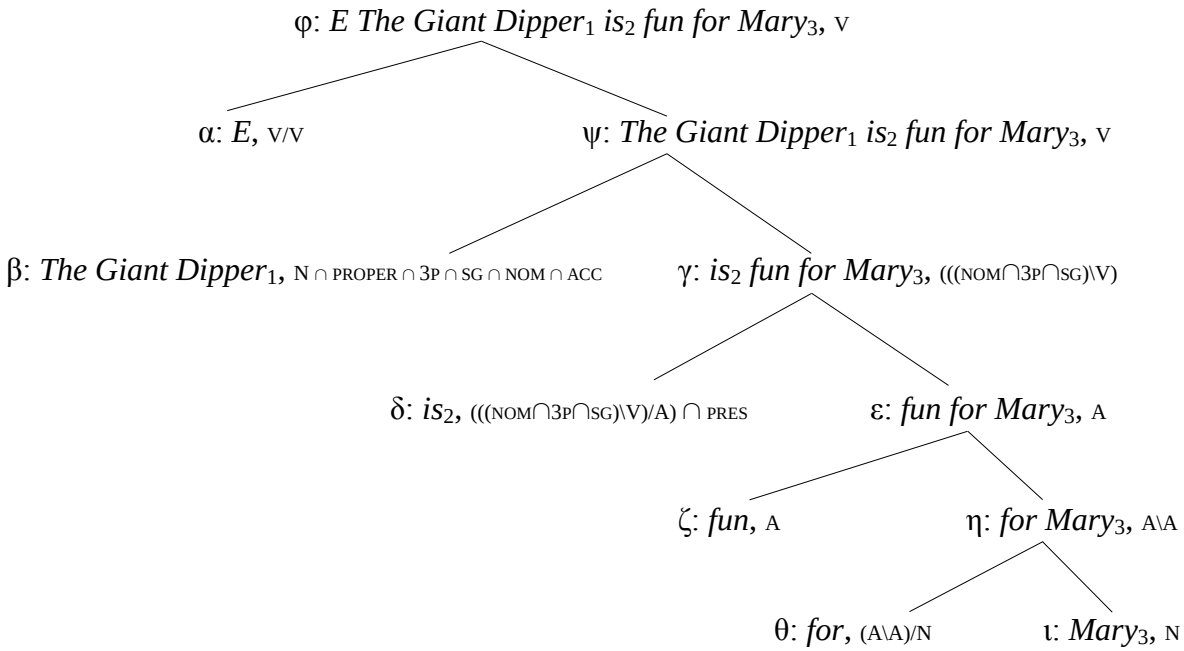
As an example, consider a use of the sentence *The Giant Dipper is fun for Mary*, again with denotations assigned relative to an arbitrary worlds *u*, *w* and perspective $p = \langle x, t, a\rangle$:

[76] This way of putting it mixes metalanguage and object language, but is, I hope, sufficiently clear. A more precise formulation will be given below.

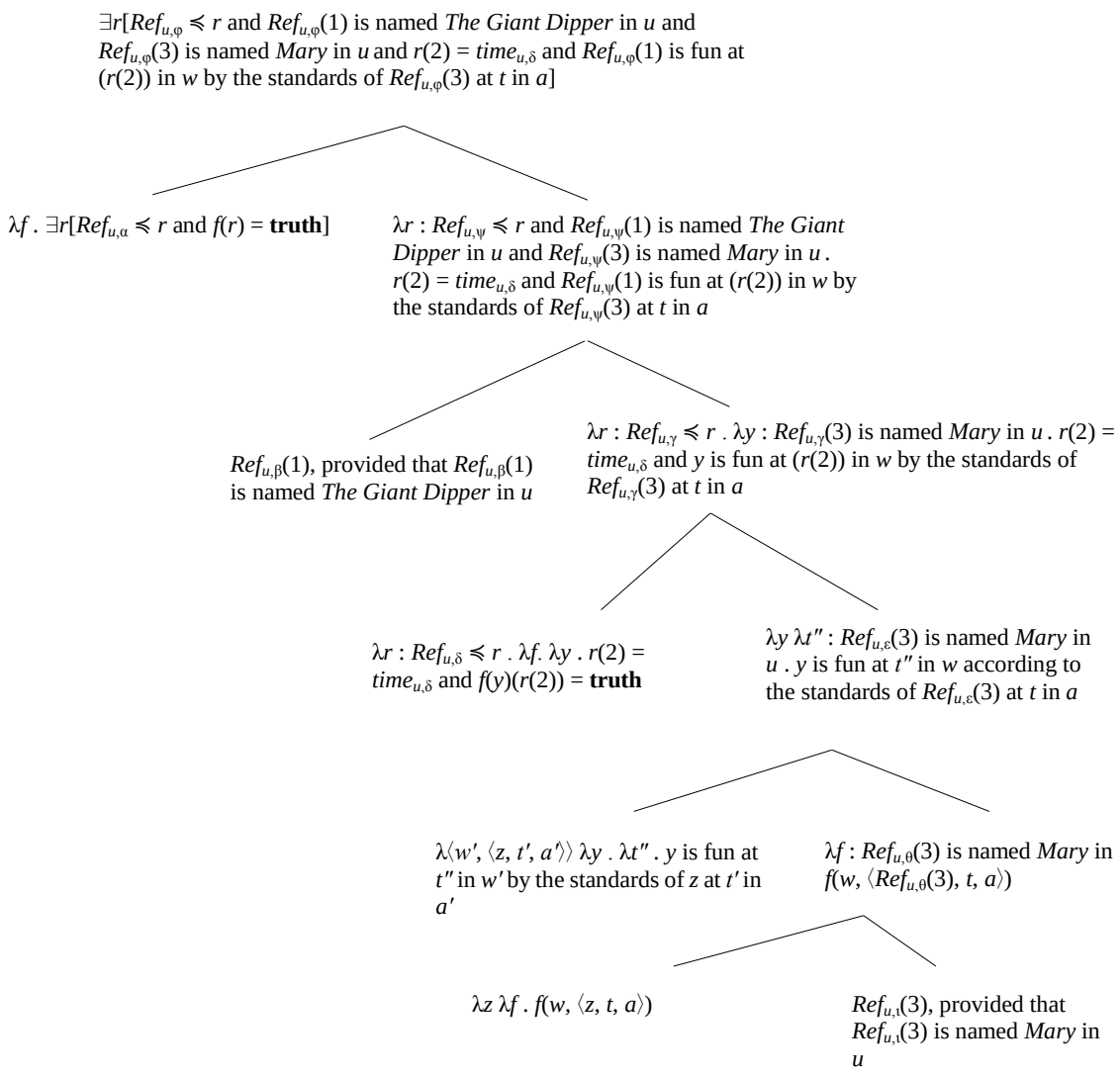
[77] Several authors have criticized this approach to *for*-phrases by pointing to evidence that they are arguments of the taste predicate, rather than adjuncts; see Stephenson (2007b), Glanzberg (2007), Schaffer (2009), Snyder (2013), Bylinina (2013). However, Collins (2013) provides extended arguments to the contrary; see also Lasersohn (2008). I regard this as a relatively minor issue; the question of whether *for*-phrases are arguments or adjuncts is completely independent of the more fundamental question whether predicates of personal taste *without* overt *for*-phrases should be analyzed in a relativist or a contextualist manner.

[78] I have made some adjustments here for differences in notation, number of indices, etc.

(146)


φ: E The Giant Dipper1 is2 fun for Mary3, V
α: E, V/V
ψ: The Giant Dipper1 is2 fun for Mary3, V
β: The Giant Dipper1, N ∩ PROPER ∩ 3P ∩ SG ∩ NOM ∩ ACC
γ: is2 fun for Mary3, (((NOM∩3P∩SG)\V)
δ: is2, (((NOM∩3P∩SG)\V)/A) ∩ PRES
ε: fun for Mary3, A
ζ: fun, A
η: for Mary3, A\A
θ: for, (A\A)/N
ι: Mary3, N

(147)



Notice that the final truth condition states that the sentence is true relative to an arbitrary perspective p iff The Giant Dipper is fun by Mary's standards (at the time and world of p). The truth value does not vary from individual to individual, but is assigned "objectively," at least as far as dependence on choice of individuals is concerned. (The sentence may still vary in truth value based on the other elements of the context of assessment, *viz*. the time and world.)

The analysis extends straightforwardly to quantificational examples, such as *The Giant Dipper is fun for every man*, as readers may easily confirm for themselves.[79]

The analysis just given is far from comprehensive. Even for the sentences it deals with, it fails to account for the full range of naturally available interpretations; and of course it will be

[79] Contrary to what Snyder (2013) seems to imply, dealing with such examples in a relativist framework does not involve a use-mention confusion, or an appeal to "Lewis Carroll" interpretations in which quantifiers are treated as names, or any other such technical problems.

necessary to extend it to a wider range of sentences as well (such as attitude sentences). But it should have provided a detailed enough initial formulation that we can now begin to consider alternatives.

Chapter 5 : "Hidden" and "Disguised" Elements

The indexical expressions we have discussed so far — English pronouns and tense markings — are "overt" indexicals, in the sense that they are manifested by stretches of sound in speech (and strings of letters in writing), and have semantic contents whose denotations vary in a relatively direct and obvious way with features of the pragmatic context. Sentences like *The roller coaster is fun* or *The licorice is tasty* do not contain such elements, beyond the usual ones which they share also with sentences about matters of fact. This is not enough to rule out the possibility of an indexical analysis, however, because the defender of such an analysis would presumably claim that the indexicality in these sentences is in some way "hidden" or "disguised" — that is, that it either cannot be traced to any particular stretch of sound in these sentences, or else must be traced to an expression whose status as indexical is for some reason not very obvious. To assess the plausibility of such an analysis as an alternative to the relativist analysis developed in Chapter 4, we should consider what kinds of hidden or disguised elements can contribute to semantic interpretation, and whether expressions of personal taste can be analyzed in the same way.

Any appeal to hidden elements in linguistic analysis is to some extent controversial. But there are a variety of ways an element which contributes to semantic interpretation might be said to be "hidden," each with its own set of theoretical issues, some of which may be more problematic than others. It will be useful to consider separately the various levels at which an element can be "hidden" — phonology, syntax, etc. — and the kinds of interpretation a hidden element may have — indexical, indefinite, etc.

5.1. Phonological reduction

Perhaps the least controversial kind of hidden element would be expressions which are phonologically reduced to the point of not being pronounced. A fairly clear example of this in English is the omission of unstressed material in utterance-initial position. For example, in casual speech styles, the struck through portions of examples (148)a.–c. need not be pronounced:

(148) a. ~~There is a~~ big storm coming.
b. ~~It's~~ time to go.
c. ~~I~~ saw a funny thing yesterday.

Zwicky and Pullum (1983) argue that this kind of omission should be analyzed as a (morpho-) phonological phenomenon, not a syntactic one.[80] Note that the portion of the sentence which is omitted need not form a syntactic constituent; few theories of syntax would treat *There is a* in (148)a. as coherent functional unit which could be targeted by syntactic rules.[81] In fact, as Zwicky and Pullum point out, even proper parts of words may be omitted in this way:

[80] See Weir (2012) for a more recent analysis.
[81] The exception would be those versions of Categorial Grammar (such as the Lambek Calculus or Combinatory Categorial Grammar) which allow such a flexible division into constituents that the whole idea of constituency loses most of its force.

(149) a. ~~I~~s he going?
b. ~~Di~~d he go?

Moreover, the portion of a sentence which is *not* omitted in this way often shows phonological effects of the omitted material. For example, in English, consonants are devoiced if they immediately follow a voiceless sound in the same syllable; accordingly *it's* is pronounced with a voiceless [s] but *there's* is pronounced with a final [z]. This effect persists even if *it* or *there* is deleted:

(150) a. ~~It~~'s[s] really cold in here.
b. ~~There~~'s[z] really no hope.

Listeners who encounter utterances of sentences with the initial portion omitted in this way obviously must use the pragmatic context (along with their knowledge of the grammar and phonology of the language) to interpret them properly; and in that sense, such sentences are contextually sensitive. But this seems like a very different phenomenon from indexicality. The omitted elements themselves are not required to be indexical — or to have any specific kind of interpretation, as evidenced by the examples just discussed, in which the omitted portion does not form a syntactic constituent, whole word, or even a meaningful subpart of word. This is a rule of language that targets units other than those dealt with by a compositional truth definition. Like unstressed material generally, they are interpreted as backgrounded and "given," but this is a separate matter.

Accordingly, we may assume that at the level of representation on which a compositional truth definition is stated ("logical form"), such sentences are represented with all the phonologically omitted material in place. In other words, the examples in (148) – (150) may be treated for the purposes of truth-theoretic semantic theory as though the struck-through portions were present. If someone says "Saw a funny thing yesterday," or "'D he go?" we count it as an utterance of the sentence *I saw a funny thing yesterday* or *Did he go?* The pronunciation is rather different than in slow, highly careful speech, of course; but any theory must recognize that the pronunciation of any expression varies widely from occasion to occasion, and the presence or absence of a short unstressed portion of the sentence seems well within the range of possible variation.

I know of no one who has appealed to this particular kind of "hiding" in the analysis of personal taste sentences in order to account for the intuition that their truth values may vary from person to person; nor does it seem at all likely that an adequate analysis could be developed along these lines. If we claim, for example, that the sentence *Licorice is tasty* is just a phonologically reduced version of a sentence of the form *Licorice is tasty for X*, we would need to give some account of the phonological factors which motivate this reduction; but no such account seems forthcoming. Nor do there seem to be any phonological effects of the omitted material, of the kind illustrated in (150). Finally, the analysis would leave it a mystery why the omitted phrase has the particular semantic interpretation it does, since this kind of reduction is motivated by phonological considerations, and no systematic truth-conditional effect is expected.

5.2. Null elements in syntax

In addition to cases like those just discussed, in which linguistic material is hidden via some phonological process, there may also be phenomena of a more "syntactic" nature which result in expressions being hidden. A classic example is *pro-drop* — the systematic omission of pronouns, except as required for emphasis, contrast or similar pragmatic effects. This may be seen in examples (151)a. and b., from Spanish and Italian, respectively.[82] The omitted pronouns are parenthesized in the English translations, but not present in the overt forms of the original sentences.

(151) a. Cantaron.
"(They) sang."

b. Ha comprato quella casa.
"(He/she/it) has bought that house."

Pro-drop is unavailable in English but common cross-linguistically. A rather heterogeneous set of cases have been analyzed under this heading, and a completely unified analysis is probably neither possible nor desirable; but the classical cases represented by Spanish and Italian show a few similarities with, but also significant differences from, the case of English omission of initial unstressed material just discussed.

As in English, classical examples of pro-drop are limited to cases where the omitted item would not bear phonological stress. But this is to be expected for any example of an unpronounced linguistic expression; an item must be pronounced to bear stress, so when stress is obligatory, so is overt pronunciation.

In other respects, pro-drop is rather different from English initial unstressed deletion. It is not limited to casual or colloquial style. It is not limited to sentence-initial position. It *is* limited to pronouns, rather than unstressed material more generally[83]; for this reason we do *not* find examples, as we do in English, where the omitted material does not form a syntactic unit. At least in Spanish and Italian, pro-drop is limited to subjects; and this sensitivity to grammatical role suggests an analysis which treats pro-drop at least in part as a syntactically conditioned phenomenon, and not a purely phonological one.

The syntactic nature of pro-drop renders it more questionable than in the English case whether we can analyze an utterance of a pro-drop sentence simply as an alternative pronunciation of the very same sentence as its non-pro-drop counterpart. But even if we do not claim that the sentences are the same, we may still analyze pro-drop sentences as having a syntactic structure with a pronoun — albeit unpronounced — in the position of the missing subject. This is, in fact, by far the most popular syntactic analysis, and we have no reason to call it into question here. Among the advantages for syntactic theory in assuming the subject is syntactically present are that it simplifies and regularizes the analyses of subject-predicate agreement and the distribution of reflexive

[82] Examples are from Zagona (2002) p. 25, Decker (2010) p. 119.

[83] This is true for Spanish and Italian, but probably not for Chinese and Japanese, where examples frequently discussed as involving "pro-drop" actually receive interpretations which are unavailable for overt pronouns. See Takahashi (2008). This and other differences have convinced many linguists that "pro-drop" is a misnomer in such languages, which really involve a different phenomenon — perhaps syntactic ellipsis.

pronouns. These considerations will be familiar and elementary to readers well versed in modern syntactic theory, but are worth reviewing here because they provide illustrations of the kinds of evidence one can cite to support the claim that a hidden element is syntactically present — an issue that will be of concern to us in evaluating analyses of personal taste sentences which appeal to such elements.

In many languages, predicates are marked morphologically to agree with their subjects (and in some languages, with other arguments). In the familiar European languages, verbs are typically marked to agree with their subjects in person and number, while adjectives are often marked to agree in number and gender, as illustrated in the following examples from Spanish:[84]

(152) a. El chico es simpático.
the.MASC boy is nice.MASC
"The boy is nice."

b. La chica es simpática.
the.FEM girl is nice.FEM
"The girl is nice."

The exact conditions governing agreement vary from language to language and in some cases are sensitive to a complex mixture of syntactic and semantic factors which cannot be reviewed in detail here. The relevant point for our purposes is that some examples of agreement are problematic to account for in purely semantic terms, and that some sort of syntactic mechanism therefore appears to be a necessary part of the analysis. In many languages, for example, grammatical gender is assigned on a largely arbitrary basis, rather than on the basis of sex as in English. In the following examples, the adjective is marked to agree with the gender of the subject nouns, both of which refer to inanimate, sexless objects:

(153) a. El papel es blanco
the.MASC paper is white.MASC
"The (piece of) paper is white."

b. La hoja es blanca.
the.FEM sheet is white.FEM
"The sheet (of paper) is white."

Masculine or feminine morphology on an adjective in Spanish therefore cannot be analyzed as restricting the adjective to hold only of male or female individuals respectively. One *could* analyze masculine-marked adjectives as predicates that hold only of those individuals which are in the extensions of grammatically masculine nouns, and feminine-marked adjectives as holding only of individuals in the extensions of feminine nouns; but this will not account for agreement, since a given individual can be in the extensions of many different nouns of different genders — one even finds synonyms of opposite gender — yet predicative adjectives still must agree in gender with their subjects. This is easy to account for if agreement is imposed by a syntactic rule forcing the adjective

[84] Thanks to Jon MacDonald for help with Spanish examples.

to be marked morphologically with the same grammatical gender as the subject phrase, rather than by semantic restrictions on what kinds of things the adjective can truthfully be predicated of.

But if we say this, and then claim that pro-drop sentences do *not* have syntactically represented subjects, the appearance of gender marking on predicative adjectives in pro-drop sentences becomes a mystery:

(154) a. Es blanco
is white.MASC
"He/it is white."

b. Es blanca
is white.FEM
"She/it is white."

Gender agreement thus supports the view that pro-drop sentences do have syntactically represented subjects, even though those subjects are not realized in phonetic or written form.

The distribution of reflexive pronouns provides a second kind of evidence for syntactically represented subjects in pro-drop sentences. In many languages, a reflexive pronoun may be used only if it has a local antecedent — roughly, an antecedent in the same minimal clause.[85] For example, (155)a. is ungrammatical in Spanish because the reflexive has no antecedent. Yet (155)b. is grammatical, suggesting that an antecedent is structurally present, even though not overtly present in speech or writing:[86]

(155) a. *El problema fue resuelto por sí mismo
The problem was solved by himself

b. Hablaba consigo mismo
spoke with himself
"(He) talked to himself."

To summarize, positing a hidden subject pronoun in pro-drop examples allows us to make sense of syntactic patterns which otherwise would appear all to treat such examples as irregular and exceptional. In this sense we may say that there is *syntactic evidence* for the hidden pronouns — we posit them not because doing so is necessary to treat pro-drop sentences as expressing "complete propositions" or any other such semantic concerns, but because doing so facilitates the smooth operation and explanatory power of syntactic theory.

The same might be said for elements which are hidden due to *syntactic ellipsis*. Some caution about terminology is called for here, because the term *ellipsis* has often been used in the

[85] It is not unusual for languages to conform mostly to this pattern, but allow a specific, limited class of exceptions. This is the case in English, for example, where the antecedence conditions on reflexives in coordinate structures and in certain adjuncts are different from the conditions which hold on reflexives more generally. Such complications do not undermine the point that the distribution of reflexives can be used to argue for syntactically represented hidden elements.
[86] Examples are from Zagona (2002), p. 125.

philosophical literature in a much less specific way than it is usually used in the modern linguistic literature, as emphasized by Neale (2004). Often, philosophers will claim that one sentence is elliptical for another, meaning simply that the shorter sentence in some sense goes proxy for the longer. The exact nature of this proxy relation may be incidental to the main concerns of the analysis and deliberately left unclear — even to the point where no claim is made whether it is to be characterized in syntactic, semantic, or pragmatic terms. In the linguistic literature, however, the term *ellipsis* is usually reserved for a specific kind of syntactic pattern, licensed in particular structural configurations and governed by general syntactic principles.

An example is English verb phrase ellipsis. Examples like (156) contain a verb phrase "gap," indicated here by an underscore, interpreted like some earlier verb phrase in the discourse, modulo tense and agreement marking:

(156) John built a table. Mary will ____ too.
(= John built a table. Mary will *build a table* too.)

The gap is not simply filled in pragmatically, but must have a linguistic antecedent (Hankamer and Sag (1976)). If the speaker simply points to John building the table rather than saying "John built a table," it would be quite odd to continue with "Mary will too." This sensitivity to linguistic structure is suggestive of a syntactic analysis.

Moreover, the acceptability of verb phrase ellipsis depends on whether the missing material would violate *island constraints* (that is, syntactic principles which prohibit movement from certain types of phrase). In English, an interrogative pronoun cannot be extracted from an adverbial clause, as illustrated in (157)a. Here, the interrogative pronoun *what* appears at the beginning of the sentence, but functions as the object of *playing*. The example is ungrammatical because in order to move *what* from canonical object position (indicated by *e*), it must escape the temporal adverbial clause enclosed in brackets, which is a syntactic island. Now compare (157)b., due to Chung, Ladusaw, and McCloskey (1995). The second sentence of this example is also ungrammatical, presumably because if the verb phrase gap (indicated by the underscore) is filled in with the missing material, the result is a sentence which could only result from an island violation (namely, extraction of *what* from the adverbial clause *before they did start playing*):

(157) a. *What did you leave [before they started playing *e*]?
b. We left before they started playing party games. *What did you leave [before they did ___]?
(=…*What did you leave [before they did start playing *e*]?)

This sensitivity of grammaticality patterns to the structural configuration of the missing material provides convincing evidence that this material is present at some level of syntactic representation, even if not overtly realized.

A similar argument for hidden syntactic material can be made on the basis of *crossover* violations. Normally, an interrogative pronoun cannot "cross over" a coindexed pronoun, thereby binding it, even though the position from which the interrogative pronoun originates does not c-command the pronoun it binds. Example (158) is ungrammatical under the indexation shown; the sentence cannot be assigned the structure which would result in the interpretation "I know which *x* is such that the man who bought *x* threw away *x*."

(158) *I know what$_i$ the man who bought it$_i$ threw away e_i

We can use this pattern to argue for the presence of hidden pronoun-like expressions in certain syntactic positions. For example, consider uses of the verb *forget* in which it does not have an overt direct object or clausal complement, as in *John remembered to wash the dishes, but Mary forgot.* If we place such a use of *forgot* in a position where it would be "crossed over" by an interrogative pronoun, we find the same effect as we did with overt pronouns. Example (159) resists the interpretation "I know for which *x*, the man who forgot *x* later attempted *x*:

(159) *I know what$_i$ the man who forgot ___ later attempted e_i.

This is easy to explain if we analyze the sentence syntactically as having a pronoun-like element at the position of the underscore, which would be subject to the same constraints on indexation which rule out (158). Here again, we find that a linguistic pattern is easily and straightforwardly explained in terms of independently motivated syntactic principles, if we analyze sentences as containing expressions which are not overtly realized in speech or writing. In these sorts of cases, where positing such expressions leads to a simpler or more explanatory syntactic theory, we may say that there is syntactic evidence for them.

The idea that personal taste sentences contain a pronoun-like expression which is hidden for syntactic reasons has much greater initial plausibility than the idea (discussed in the last section) that they contain an expression which is reduced by phonological processes to the point of omission. This is the position advocated by Glanzberg (2007), Schaffer (2009), and Snyder (2013), for example. However, this idea too faces significant problems.

First is the basic semantic problem pointed out in Chapter 2, which motivated our exploration of a relativist alternative in the first place: If we think of the hidden element as having an interpretation like an indexical pronoun, then the content of any sentence in which it appears will vary with the contextually supplied referent of the hidden expression. For example, *Roller coasters are fun* will express a different semantic content when the hidden element refers to John than it does when it refers to Mary. In that case, *Roller coasters are fun* need not contradict *Roller coasters are not fun*. If we are using this analysis to try to account for the intuition that John and Mary may use these respective sentences to express their personal opinions of roller coasters, each sincerely and without making an error of fact, then we fail to account for the intuition that they contradict each other in doing so.

But in addition, such an analysis faces significant problems for syntactic analysis. If *fun* combined with a syntactically present but hidden pronoun, we would expect it to produce crossover effects, much like the hidden argument of *forget*.[87] However, it does not. To see this, first we must recognize that some speakers allow bound readings for the parameter relative to which something is judged to be fun or tasty.[88] Such speakers allow a reading of (160) which means that the speaker knows for which *x* it is the case that *x* went on a ride which was fun for *x*:

[87] I repeat here with some modification an argument given in Lasersohn (2008).
[88] I have found that a significant number of speakers do not allow such readings, but many do.

(161) I know who$_i$ e_i went on a fun ride.

In (161) the interrogative pronoun *who* does not "cross over" *fun*, so crossover constraints should not rule out a bound reading. But now consider (162):

(162) I know who$_i$ a ride that is fun ___ would appeal to e_i.

If *fun* has a syntactically represented pronoun-like argument, it should not be bindable by *who* in (162), so the sentence should resist a reading paraphrased as "I know for which *x* a ride that is fun for *x* would appeal to *x*." In fact, however, those speakers who accept a bound reading for (161) seem uniformly to accept it for (162) as well, indicating that there is *not* a syntactically represented hidden pronoun argument for *fun*.

The lack of cross-over effects in sentences like (162) is evidence that *fun* does not take a syntactically represented hidden argument, but we still must account for the bound reading of sentences like (161), for those speakers who allow it. Extending our grammar to include interrogative clauses (like the subordinate clause in (161)) would take us too far afield, but such speakers generally also allow a bound reading in examples like (163), in which the binder is an ordinary quantifier rather than an interrogative pronoun:

(163) Every customer received a tasty dish.

The relevant reading here is that every customer *x* received a dish that was tasty for *x* (though not necessarily tasty according to the standards of whoever assesses the sentence for truth or falsity). Stanley (2000) has argued that anything interpreted as a bound variable must be syntactically represented, and one might argue more particularly that examples like (163) provide evidence for a syntactically represented hidden argument to *tasty*.[89]

However, the bound reading is easy enough to account for without claiming that *fun* or *tasty* combine with a syntactically represented hidden argument, if we simply allow quantification on the perspective index, in effect allowing ordinary quantifiers to bind this index as they would a variable. One technique for achieving this effect, developed in Lasersohn (2008), appeals to a special "μ-operator" which ties the value of the perspective index to that of a free variable; I repeat this analysis here with some adaptation.

Like an analysis which appeals to syntactically represented variable serving as a hidden argument to *fun*, this analysis will involve a hidden element in syntax. However, it will be a variable-binding operator, rather than a variable or indexical, so the analysis will not give us any reason to expect cross-over effects, unlike a hidden-variable analysis.

For any positive integer *i*, we let $\mu_i \in$ V/V, and define it as follows:

[89] In giving oral presentations on relativist semantics over many years, I have found that audiences raise examples like (163) as a potential problem with some regularity. Schaffer (2009) offers a similar objection, but is careful to present it only as an argument that *tasty* combines with a variable at "logical form," remaining neutral as to whether the variable is projected syntactically.

(164) If USE(α, μ_i) then $Lex(\alpha, u, w, p) = [\lambda q: q \in \mathbf{D}_{\langle r,\langle s,t\rangle\rangle} . [\lambda r : Ref_{u,\alpha} \preccurlyeq r . \exists t\, q(w, \langle r(i), t, w\rangle) =$ **truth**$]]$

To see how this works, consider a clause containing a trace and a predicate of personal taste, such as [e_1 *received*$_2$ *a tasty dish*$_{3,2}$]. Our grammar as currently formulated will not actually generate this clause, since it makes no provision for adjectives in attributive position, and does not include the verb *received*; but straightforward extensions to our rules would assign the following denotation to a use α of this expression in an intensional context:

(165) $[\lambda r : Ref_{u,\alpha} \preccurlyeq r$ & $r(2) < time_{\alpha,u} . \lambda\langle w, \langle x, t, a\rangle\rangle . r(1)$ receives $r(3)$ at $r(2)$ in w and $r(3)$ is tasty at $r(2)$ in w by the standards of x at t in a and $r(3)$ is a dish at $r(2)$ in $w]$

This is of the right type to substitute for q in (164), yielding the following denotation for a corresponding use of μ_1[e_1 *received*$_2$ *a tasty dish*$_{3,2}$]:

(166) $[\lambda r : Ref_{u,\alpha} \preccurlyeq r$ & $r(2) < time_{\alpha,u} . \lambda\langle w, \langle x, t, a\rangle\rangle . \exists t'\, r(1)$ receives $r(3)$ at $r(2)$ in w and $r(3)$ is tasty at $r(2)$ in w by the standards of $r(1)$ at t' in w and $r(3)$ is a dish at $r(2)$ in $w]$

This may now combine with *every customer*$_{1,2}$ to give the bound reading in a straightforward fashion. No hidden pronoun-like argument to *tasty* is needed to derive this reading.

A different argument for the syntactic presence of a hidden argument for *fun* is offered by Schaffer (2009). Schaffer suggests that only a syntactically represented expression can serve as the "controller" of the PRO subject of an infinitival clause. For example, in (167)a., *John* is the controller of PRO, so we understand the sentence to mean that John sank the ship so that he, John, could collect the insurance. In (167)b., there is no overt controller, but — it is often assumed — there is a syntactically represented but phonologically unpronounced phrase interpreted something like *by someone*, so that the sentence means that the ship was sunk by someone so that that person, the one who sank the ship, could collect the insurance. But in (167)c., where there is no hidden *by*-phrase, there is no controller for PRO, so the sentence is ungrammatical:

(167) a. John sank the ship [PRO to collect the insurance].
b. The ship was sunk [PRO to collect the insurance].
c. The ship sank [PRO to collect the insurance].

As Shaffer points out, infinitival clauses can serve as arguments to *fun*:

(168) It is fun [PRO to dance].

The natural interpretation here one in which the person relative to whom dancing is judged to be fun is the person whose dancing is judged to be fun. That is, the subject of *dance* is the same as the person relative to whom dancing is judged to be fun. Schaffer suggests that this provides evidence that *fun* has a hidden "experiencer" argument which serves as controller for PRO.

However, it is far from clear that we need to appeal to control of PRO to account for this semantic pattern. This becomes clear as soon as we examine cases where the overt argument of *fun* is

an ordinary noun phrase instead of an infinitival or gerundive phrase:

(169) Board games are fun.

The natural reading for (169) is that the person relative to whom board games are judged to be fun is the same as the one whose playing of the board games is judged to be fun. But in such examples, we presumably do not want to analyze the noun phrase *board games* as containing a hidden PRO argument, denoting the person who plays the games. Rather, we simply predicate *fun* of the games themselves, and count a game as fun relative to a world w and perspective $p = \langle x, t, w' \rangle$ iff events of interacting with that game in the normal (or other contextually salient) way are fun relative to w and p. As discussed in more detail in Section 6.2. below, it makes sense to assess whether an event e is fun relative to $\langle x, t, w' \rangle$ only if x participates in e, giving the effect that (169) is interpreted so that the person relative to whom board games are judged to be fun is the person whose playing of the board games is judged to be fun — with no need for PRO or a syntactically represented controller at all.

Once we recognize the availability of this approach for examples like (169), there is no reason we should not use it in examples like (168) as well. The semantics of PRO will be a topic for Section 7.4. Simplifying somewhat, the analysis presented there treats an infinitival clause like [PRO *to dance*] as denoting the relation which an individual x stands in to a world w iff x dances in w.[90] Just as with board games, we may count such a relation as fun relative to a world w and perspective p iff the events by virtue of which an individual stands in that relation to w are fun relative to w, p. Given that an event e is fun relative to $\langle x, t, w' \rangle$ only if x participates in e, we obtain the result that the person relative to whom dancing is judged to be fun is the person whose dancing is judged to be fun — still with no need for a syntactically represented controller.

5.3. Unarticulated constituents

Phonological reduction, pro-drop and syntactic ellipsis all involve linguistic expressions which are present in syntactic representation but unpronounced. We should also consider the possibility that some elements of semantic interpretation are not represented linguistically even at the level of syntax — that they play a role in semantic content, but are provided on an entirely non-linguistic basis, rather than as the contents of linguistic expressions. Such elements have usually been termed *unarticulated constituents*, following Perry (1986), because they may be seen as constituents of propositional content which are not articulated in syntactic representation.[91] This terminology makes most sense intuitively if a proposition is taken to be a structured object with identifiable parts, at least some of which are the contents of syntactic constituents of a sentence expressing the proposition. In the present framework, where the content of a sentence use is not such a structured object, but a function from indices to truth values, the terminology is less than ideal; but it is by now

[90] The actual denotation is more complex: [PRO$_i$ *to dance*] denotes the function $[\lambda r \lambda \langle w, p \rangle \lambda x \lambda t\, \lambda t' : r(i)$ includes x & $t \leq t' . r(i)$ dances at t' in $w]$. The complications are added to deal with "partial" control and with the temporal interpretation of infinitival *to*. See Section 7.4. for details.

[91] Many linguists may find this an odd or confusing choice of terminology, since in the linguistic literature "unarticulated" is usually understood as implying lack of *phonetic* articulation and "constituent" as referring to units of *syntactic* representation. But the idea behind unarticulated constituents is not that of syntactically represented but phonetically unrealized elements, but rather of aspects of propositional content which are not contributed by linguistic expressions at all.

so well entrenched that it cannot be avoided. Standard examples of sentences for which an analysis using unarticulated constituents has been offered include *It is raining*, used to express the proposition that it is raining in some particular location,[92] but which arguably does not have any element of its syntactic structure which denotes that location, and quantificational sentences like *Every bottle is green*, used to express the proposition that every bottle in some limited, pragmatically salient set of bottles is green, but which does not seem to have any element of its syntactic structure expressing the limitation to this set.

For our purposes, we need give only a very informal characterization of when (an analysis of) a sentence involves an unarticulated constituent. Unarticulated constituents are provided through extragrammatical pragmatics — they are used in deriving the contents of complex expressions, without being encoded in the contents of the syntactic parts of those expressions or supplied via grammatical rule. Even in an analysis which makes some use of unarticulated constituents, the grammar must specify a set of semantic operations, by which the contents of complex expressions are *related* to the contents of their parts. (For example, in our present grammar, these operations are defined via (130).) If these operations and the contents of the parts are insufficient to *determine* the content of the complex expression, there must be some extragrammatical factor involved in assigning it. In other words, if the content of a complex expression cannot be derived from the contents of its parts via the semantic operations specified by the grammar, then there must be an unarticulated constituent involved in deriving its content.

The debate over whether unarticulated constituents can be legitimately appealed to in semantic analysis is too extensive to be reviewed or evaluated here. But it should be noted that no such appeal is possible under the assumptions already adopted in this book. In particular, any analysis which uses unarticulated constituents directly violates the assumption that content is assigned homomorphically, as given in (13). Any analysis which appeals to unarticulated constituents ought therefore to address the intuition underlying this assumption — which, I hasten to repeat, is based on the idea that "the truth of what you say depends on what you are talking about," as discussed in Section 1.5. above, and not on considerations of language learnability or people's ability to understand novel sentences.

5.4. Constructional indexicality

It is perhaps worth noting that a relatively slight revision of our assumptions about compositionality would allow for certain cases where a complex expression varies in content with the context, but where this variation cannot be traced to a hidden or overt indexical expression which it has as a syntactic part. All that is necessary is to allow the grammar of a language to invoke different semantic operations in different contexts.

As an illustration, consider an analysis in which contextual restrictions on domains of quantification are built into the rule which derives the denotation of a complex noun phrase from the denotation of its determiner and the denotation of the noun or noun phrase with which the determiner combines. (I am *not* suggesting that this is the correct account of such restrictions; the point is simply to illustrate the theoretical possibility of this kind of analysis.) We might adopt a rule for combining quantificational determiners with their nouns to the effect of (170), for example[93]:

---

[92] I think this is not in fact the semantic content of this sentence, as argued on p. 99 below.

[93] Cf. (43) in Stanley and Szabó (2000), (15) in Lasersohn (2012). For the purposes of illustration I revert here to a

(170) Where α is a use of a complex expression consisting of δ and ν, if $[\![\delta]\!]^{u,w} \in \mathbf{D}_{\langle\langle r,e\rangle,\tau\rangle}$ and $[\![\nu]\!]^{u,w} \in \mathbf{D}_{\langle r,e\rangle}$, then $[\![\alpha]\!]^{u,w} = [\![\delta]\!]^{u,w}(\lambda r : \mathit{Ref}_{u,\nu} \preccurlyeq r \text{ and } [\![\nu]\!]^{u,w}(r) \in \mathit{Relevant}_{u,\nu} . [\![\nu]\!]^{u,w}(r))$].

This rule applies the denotation of the determiner δ to the result of limiting the range of the denotation of the noun ν to $\mathit{Relevant}_{u,\nu}$, the set of things which are pragmatically relevant in the context of ν (in world of use *u*). This appeal to $\mathit{Relevant}_{u,\nu}$ in the formulation of the rule allows for cases where to uses α, β of a noun phrase can have different contents, even if every constituent of α is identical in content to the corresponding constituent of β. This is allowed because the contents of α and β depend not just on the contents of their constituents, but also on what is relevant in their respective contexts.

Some semanticists (Stanley and Szabó (2000), Stanley (2000)) are explicit in maintaining that any appeal to contextual variation in the semantic operations which are used in interpreting a phrase amounts to a violation of compositionality. Whether or not this is correct obviously depends on exactly what we mean by "compositionality," but if we understand compositionality as a constraint that interpretation be assigned homomorphically, and allow the semantic algebra with respect to which the homomorphic interpretation assignment is defined to vary by context, the claim is not correct. Allowing contextual variation in semantic operations does *not* undermine the idea that interpretation is assigned by the same grammar in different contexts, as discussed in more detail in Lasersohn (2012). A grammar must provide some context-invariant way of saying which semantic operations are to be used, but this does not mean that the operations themselves must be the same in all contexts. For example, the rule in (170) may be construed as part of the grammar, and need not be reformulated from one context to the next. The grammar is a set of rules which determine which operations are used; it is not to be identified with the operations themselves.

Contextual variation in content which is produced in this way may be termed *constructional indexicality*, since the variation is traced to the construction of complex expressions, rather than to the elementary parts from which they are constructed. Whether this technique is actually needed in the analysis of natural language is a separate question, which cannot be resolved here.[94] It is clear in any case that the rule in (170) is inadequate for a serious treatment of domain restriction in English; we return to this topic at the end of Section 5.5. below.

5.5. Sublexical and compound indexicality

Another way an expression might exhibit contextual variation in content without being obviously indexical is if its indexicality is *disguised* — that is, traceable to some particular lexical item whose status as indexical is not immediately apparent.

Of course the extent to which indexicality is "apparent" depends not just on its own properties but on the theoretical assumptions we make in considering it. If, for example, we take one of the lessons of Kaplan (1989) to be that contexts can be represented as sequences of a time, a location, a speaker, etc., and that each indexical item directly and rigidly denotes one of these

---

non-relativist semantics.

[94] Certain syntactic theories, especially Construction Grammar (Fillmore, Kay, and O'Connor (1988), Goldberg (1995)), would seem in principle to be theoretically committed to such an approach, but the assignment of truth conditions is a largely unexplored topic in Construction Grammar.

elements, then any item which has a non-rigid denotation, or whose denotation is not some simple contextual parameter, will seem *prima facie* not to be indexical. (Never mind that Kaplan himself did not actually represent contexts in this way, or claim that each indexical must denote one of the simple parameters he stipulates each context to provide.)

Arguably, there are many examples of expressions whose denotations vary with some aspect of the context but which do not denote that element. Rather, the denotation may be defined partly in terms of the contextual parameter, but partly in non-contextual terms. The denotation itself may therefore be modally non-rigid, despite the indexicality.

Natural kind terms analyzed in something like the style of Putnam (1975) provide an easy illustration. Suppose we take the noun *lemon* to denote (relative to a given world *w*) anything which is sufficiently similar to some sample of lemons which we, or experts in our community to whom we defer, identify in the actual world $w_@$. The comparison is always to this actual sample, not to different samples in different worlds, so there is a modally rigid aspect to the meaning. But the number of things which are sufficiently similar to this sample can surely vary from world to world, so the denotation of the noun must vary as well; the denotation as a whole is not rigid.

At the same time, the choice of similarity relations depends on the context, and particularly the interests and purpose of the speaker. In a scientific context, the requisite similarity is one of underlying structure — genetic or other biological properties — but in the context of a photographer setting up a shot, a similarity of outward appearance may be all that is required, so that even a plastic lemon would qualify. This contextual variation in the choice of similarity relation effectively assigns different contents to the noun in different contexts. Under this analysis, then, we have variation both in the content of *lemon* from context to context, and in the denotation of each content from world to world; the two types of variation are not mutually exclusive.

In this kind of case, there is indexicality involved, but the value of the contextual parameter does not serve as the denotation. The indexicality is *sublexical*; one can imagine analyses in which the word is decomposed into smaller expressions, some but not all of which are indexical.

A full evaluation of the extent to which content varies with context would take us far beyond the scope of this book. I personally have rather strong doubts whether a plastic lemon can ever be truthfully described as a lemon; but the fundamental point that content may be determined on the basis of a context-dependent choice of similarity relations does not seem at all implausible. Therefore, we will make no assumption here that indexicality is limited to a small, closed class of lexical items like pronouns, *here* and *now*, nor any assumption that each indexical expression directly and rigidly denotes some element of the pragmatic context.

By adopting such a liberal position on the extent of contextual variation we also strengthen any argument we may give against an indexical analysis of expressions of personal taste or other matters of opinion: if we were to assume in advance a highly restrictive theory of indexicality, and then show that it was inadequate to the analysis of expressions of personal taste, the possibility might remain that the inadequacy was due to our having adopted an unrealistically narrow view of what indexicality is like. But if even a relatively permissive theory of indexicality fails in accounting for such expressions, a stronger case may be made that something other than indexicality is needed to explain them.

Adopting such a permissive view of the extent of contextual variation in content will not permit us to limit our statement of truth conditions to the classic disquotational style of Tarski (1944). When a semanticist *S* formulates a T-sentence like *'Every lemon is yellow' is true iff every lemon is yellow*, this T-sentence will assign a truth value to the sentence *Every lemon is yellow* based

on the content which *lemon* has on the right-hand side of the *iff* — and if the content of *lemon* varies with its context of use, this truth value will depend on the context in which semanticist *S* formulates the T-sentence, not the context in which the speaker asserts *Every lemon is yellow*, which might assign a quite different content to *lemon*. This is an obviously wrong result. A theorist who is committed to the use of disquotation in the statement of truth conditions will therefore naturally be attracted to the view that contextual variation in content is highly limited and constrained, so that the disruption it produces in the use of disquotation may also be seen as limited and constrained. Conversely, if we want to leave open the possibility of more varied or pervasive contextual variation in content (as I do here), then we cannot commit ourselves too strictly to a disquotational format for truth conditions.

Nor can we commit ourselves too strictly to lexical denotation assignments in the format exemplified in (171)a., since this uses the word *lemon* disquotationally. Various other strategies suggest themselves: For example, we might elaborate Putnam's strategy and regard each use α of word like *lemon* associated (in each world *u* where it occurs) with a contextually-provided similarity relation $Sim_{u,\alpha}$, then assign denotations more-or-less on the model of (171)b., or in some other way:

(171) a. If USE(α, $lemon_{i,j}$), then $[\![\alpha]\!]^{u,w,p}$ = [λ*r* : $Ref_{u,\alpha} \preccurlyeq r$ and *r*(*i*) is a lemon in *w* at *r*(*j*) . *r*(*i*)]

b. If USE(α, $lemon_{i,j}$), and *l* is a sample which $speaker_{u,\alpha}$ (or experts to whom $speaker_{u,\alpha}$ defers) assigns to the word *lemon* in $world_c$, $[\![\alpha]\!]^{u,w,p}$ = [λ*r* : $Ref_{u,\alpha} \preccurlyeq r$ and $Sim_{u,\alpha}$(*r*(*i*), *l*, *w*, *r*(*j*)) . *r*(*i*)]

(Here "$Sim_{u,\alpha}$(*r*(*i*), *l*, *w*, *r*(*j*))" means that in world *w* at time *r*(*j*), *r*(*i*) is similar to *l* in the way relevant to α in *u*.)

It would take us too far afield to seriously address here the advisability of such an analysis, let alone the issue of which vocabulary items should be defined in this way, or the details of how such definitions should be formulated. The point is merely that for the purposes of our argument here, we shall not assume that contextual variation in content is limited to a few obviously indexical items, nor that expressions which display such variation are always modally rigid, nor that truth conditions are to be stated disquotationally. It may be that any or all of these assumptions are correct, but my aim is to show that an indexical analysis of sentences expressing matters of opinion is problematic even if they are not.

It should be recognized that sublexical indexicality is not limited to words like *lemon* whose status as indexical is non-obvious or debatable. As we have seen already in Section 3.5. even as clear an example of indexicality as the word *here* turns out on close inspection to involve multiple contextual parameters, so that it cannot be treated as simply denoting "the" location parameter of the context. Since multiple parameters are involved, we may say that *here* provides an example of *compound* indexicality.

Another interesting case of sublexical indexicality involves *domain restrictions* on quantification. An illustrative analysis of domain restriction was given in Section 5.4. but this account was not adequate as serious model of domain restriction in natural language. The primary reason is that it will not account for the fact that the domain restriction on a quantifier may vary with the values of a variable, as pointed out by von Fintel (1994), Stanley (2000), Stanley and Szabó (2000), among others. For example, suppose we are discussing a contest in which each

contestant was assigned a series of hidden coins to try to find — different coins for each contestant. In this context, one can say:

(172) No contestant found every coin.

The relevant reading is one which can be paraphrased as "No contestant found every coin he or she was assigned"; that is, the quantification is universal with respect to a different domain for each contestant. This reading cannot be obtained by simply restricting quantification to some set $Relevant_{u,\alpha}$ provided globally by the context in which the quantifier is used; if the quantification were universal with respect to this set, the sentence would be unambiguously true even if every contestant found all the coins he or she was assigned, provided none of them also found all the coins assigned to the other contestants.

We can allow the domain of quantification to vary with the value of a variable by assuming that each use of a determiner δ is associated (in each world *u* where it occurs) with a function $Rel_{u,\delta}$, mapping each extension of $Ref_{u,\delta}$ onto a subset of $Relevant_{u,\delta}$. That is, each assignment of values to variables is assigned a set of relevant objects. Now we simply revise our lexical entries for determiners to mention $Rel_{u,\delta}$. For example:

(173) If USE(δ, *every*) then $[\![\delta]\!]^{u,w,p} = [\lambda r : Ref_{u,\delta} \preccurlyeq r .\ \lambda f : f \in \mathbf{D}_{\langle r,e\rangle} .\ \lambda g : g \in \mathbf{D}_{\langle r,t\rangle}$. For all r', if $Ref_{u,\delta} \preccurlyeq r'$ and $f(r')$ is defined and $f(r') \in Rel_{u,\delta}(r')$, there is at least one r'' such that $r' \preccurlyeq r''$ and $g(r'') =$ **truth**]

This analysis treats determiners as having a kind of sublexical or disguised indexicality; they will receive different contents in different contexts. It matters relatively little for our present purposes whether domain restrictions are treated as affecting the content of the determiner as suggested here, or instead the content of the noun with which the determiner combines;[95] but it may nonetheless be interesting to address one set of well-known arguments that the indexicality belongs to the "noun side" of the phrase rather than the "determiner side." Stanley and Szabó (2000) argue that domain restrictions should not be treated through the use of a hidden variable on the determiner node, claiming that such an analysis faces problems in accounting for patterns of anaphora. They point out that in a context which establishes the set of things in a particular village as the domain of quantification, the pronoun *they* in (174) may be interpreted in either of two ways:

(174) Most people regularly scream. They are crazy.

On one reading, *they* refers to the people in the village. On the other, it refers to the people in the village who regularly scream. If the domain restriction is treated as a hidden indexical on the determiner, the noun *people* will presumably denote the set of all people (not just people in the village), so there will be no node in the syntactic tree for the first sentence which denotes the people in the village, hence (Stanley and Szabó suggest), no appropriate antecedent for *they* which will yield the first reading. If, however, the denotation of the noun node varies with the domain of

[95] Or, as suggested by Stanley and Szabó (2000), the contents of the node "cohabited" by the noun and a separate hidden variable — though I have considerable difficulty making syntactic sense of this notion.

quantification, this node will denote the set of people in the village and will serve as an appropriate antecedent.

This is only a weak argument, because in many examples, precisely the opposite kind of reading is available, in which the pronoun denotes the group corresponding to the unrestricted noun. If we are discussing a particular university course in which I was the instructor, for example, I might say:

(175) I identified three plagiarized papers. They are a lot easier to spot these days, thanks to anti-plagiarism software.

The relevant reading here is one where I identified three plagiarized papers in that particular class, but it is plagiarized papers in general — not just those in that one class — which are now easier to spot because of the software. In an analysis like Stanley and Szabó's, there is no node in the syntactic tree for the first sentence which denotes the set of plagiarized papers in general, so the anaphora should be unacceptable, on their assumption that cross-sentential anaphora requires the antecedent to be a node in a preceding logical form.

Stanley and Szabó also suggest that the second reading of (174) is problematic for an analysis which locates domain restrictions on the determiner node instead of the noun node. Following Neale (1990) (who in turn was building on Evans (1977)), they suggest that when a pronoun is anaphoric on, but not c-commanded by, a (non-maximal[96]) quantifier '[*Dx*:*Fx*]' that occurs in a clause '[*Dx*:*Fx*](*Gx*)', the pronoun is interpreted as '[the *x*: *Fx* & *Gx*]'. For example, the pronoun *they* in the second sentence in (174), if anaphoric to *most people*, is interpreted as '[the *x*: *person*(*x*) & *regularly-scream*(*x*)]'. Note that the quantifier *most*, though present in the antecedent, is not used in interpreting the pronoun. If the domain restriction in the first sentence is encoded in the content of *most*, therefore, it should not be preserved in the interpretation of the pronoun; but since the pronoun is interpreted as referring to the people *in the village* who regularly scream, the restriction is clearly preserved.

This argument is only as strong as the assumption that the pronoun is interpreted via the particular syntactic algorithm Neale suggests. Stanley and Szabó claim that the argument can be extended to alternative analyses as well, but in point of fact it is straightforward to derive the intended interpretation in any framework in which quantificational noun phrases have the "live on" property of Barwise and Cooper (1981) — as they do in virtually all modern theories. Assuming for the moment that each quantificational noun phrase denotes a set Q of sets of individuals as in standard Generalized Quantifier Theory, we say that Q *lives on* A iff for all X: $X \in Q$ iff $X \cap A \in Q$. For example, if $Q = \{X : |A \cap X| > |A - X|\}$, then Q lives on A. In this example, Q is just the standard denotation for a noun phrase paraphrasable as "most members of A"; a set X will be in Q just in case more than half the members of A are in X. If the noun phrase *most people* in (174) quantifies over people in the village, and not people more generally, it will denote a set of sets which lives on the set of people in the village — regardless of whether we treat the domain restriction as being on the noun side or the determiner side of the phrase. We can then take the intersection of this set with the set of things satisfying the scope of the quantifier to obtain the set of people in the village who regularly scream — the target interpretation for the pronoun in the next sentence. More generally, we can say that if a speaker uses a sentence containing a

[96] A quantifier is '[*Dx*:*Fx*]' is maximal iff '[*Dx*:*Fx*](*Gx*)' entails '[every *x*: *Fx*](*G*x)' for arbitrary *G*.

quantificational noun phrase α binding variable *x* in scope φ, then a pronoun anaphoric on α in the subsequent discourse will denote the intersection of the smallest set on which ⟦α⟧ lives with the set of things which satisfy φ when fixed as the value of *x*.

Of course in our current framework, quantificational noun phrases do not denote sets of sets, but functions of type ⟨⟨r, t⟩, t⟩; but it is easy enough to translate between the frameworks. Let us say that a function *Q* of this type *lives on A* iff for all *P*: *Q*(*P*) = **truth** iff ∃*i* *Q*(λ*r′* . *P*(*r′*) = **truth** and *r′*(*i*) ∈ *A*) = **truth**. We may then say that a sentence use formed by substituting α for e_i in φ, where *A* is the smallest set on which $⟦α⟧^{u,w,p}$ lives, makes the set[97] $A \cap \{x \mid \exists r \preccurlyeq Ref_{u,\varphi} [⟦\varphi⟧^{u,w,p}(r) =$ **truth** and *r*(*i*) = *x*]} salient, so that it becomes available as the referent of a pronoun in later sentences. (That is, it is a member of $Relevant_{u,\psi}$, where ψ is the next sentence use in the discourse.)

Because the set on which a quantifier lives is recoverable from the quantifier denotation itself, we do not have to rely on a syntactic operation which strips away the determiner in order to interpret discourse-anaphoric pronouns.

Could it be that predicates of personal taste involve constructional, sublexical or compound indexicality? Such an analysis would not involve a hidden element with similar syntactic properties to an overt pronoun, so some of our earlier arguments against such an element might not apply. For example, a lack of crossover effects might be less surprising in an analysis that appealed to constructional, sublexical, or compound indexicality than it would in an analysis which appealed to a hidden pronoun.[98]

However, constructional, sublexical and compound indexicality, as defined here, all involve contextual variation in content. If the idea is that the content of a sentence like *Roller coasters are fun* varies with the value of a contextually supplied parameter, we are still left with our original mystery: why, if one speaker says "Roller coasters are fun" and another says "Roller coasters are not fun," we consider them to be contradicting each other, even when nothing in the contexts of the two utterances would indicate that the same value is supplied for the parameter in both cases. We can consider as many different ways of "hiding" a parameter as we want, but if semantic content varies with the parameter value, we will face problems in accounting for our intuitions about contradiction and disagreement.

5.6. Indexical and quantificational interpretations of hidden elements

In addition to the variety of ways an element may be said to be "hidden" or "disguised," we should recognize the diversity of interpretations of such elements. Our main concern here is with indexically interpreted items, but it must be recognized that not all hidden or disguised elements are interpreted indexically.

An easy example is the underlying subject of a passive sentence with no *by*-phrase. For example, *John was killed* is true iff there is someone (or something) that killed John; the

[97] Or really the "group," in the terminology we have employed so far. Recall that "groups" are the referents of plural expressions; we have made no identification of groups with sets.

[98] To be frank, however, the syntactic ramifications of constructional, sublexical and compound indexicality have not been explored in enough detail to know whether one should be surprised at a lack of crossover effects or not.

underlying subject is interpreted existentially, not indexically.[99] Similarly, sentences where *eat* is used without an overt object are often interpreted as though the object were existentially quantified; that is, *John ate* is paraphrasable as "John ate something," or perhaps as "John ate a meal."

Such examples stand in contrast to cases where a hidden or disguised element is interpreted indexically. For example, the object of the verb *forget* may be omitted, just as the object of *eat* may, but *John forgot* means that he forgot the specific piece of information under discussion, not merely that there is something which he forgot. Likewise *John arrived* is true iff he arrived at the location under discussion, not iff he arrived somewhere or other.

In terms of our current grammar, the difference between indexically and existentially interpreted elements in a sentence use φ corresponds to the difference between unbound syntactic indices which are assigned values by $Ref_{u,\varphi}$ and those which are not. If the index on an expression is assigned a value by $Ref_{u,\varphi}$, then all extensions of $Ref_{u,\varphi}$ will assign it this same value, which therefore serves intuitively as its referent. But if the index is not assigned a value by $Ref_{u,\varphi}$, it will receive different values relative to different extensions of $Ref_{u,\varphi}$, and will ultimately be interpreted as existentially quantified by the *E* operator defined in (58).

Because the choice between indexical and existential interpretations seems to be made largely on a lexical basis, we may enforce the correct reading in the lexical entries of the governing predicates:

(176) a. If USE(α, $forget_{intr_i}$) then $[\![\alpha]\!]^{u,w,p}$ = [λ*r* : *i* is in the domain of $Ref_{u,\alpha}$. λ*x* . λ*t* . *x* forgets *r*(*i*) at *t*]
b. If USE(α, eat_{intr_i}), then $[\![\alpha]\!]^{u,w,p}$ = [λ*r* : *i* is not in the domain of $Ref_{u,\alpha}$. *x* eats *r*(*i*) at *t*]

In principle, one could also have a predicate which did not impose any condition whether or not one of its indices was in the domain of $Ref_{u,\varphi}$ or not; in this case an indexical reading would be available in some contexts, and an existential reading in others. I know of no clear examples of such predicates, though one might argue that an indexical reading is sometimes available for *eat*, in examples like *John sat down in front of the ham. He picked up his fork and knife, and ate*.

It will be useful to have some kind of diagnostic to help in ascertaining whether hidden and disguised elements are interpreted indexically or existentially. Here, as in the discussion of tense in Section 3.4. above, we may appeal to sluicing or pre-sluicing sentences, following Fillmore (1986).[100] Existentially interpreted hidden and disguised elements license such sentences:

(177) a. John was killed, but I don't know who by.
b. John ate, but I don't remember what he ate.

Indexically interpreted elements do not:

(178) a. ?John arrived, but I don't remember where.
b. ?John forgot, but I don't know what he forgot.

[99] The quantification may be over a contextually defined domain, but this is a separate matter.
[100] Fillmore discusses the issue in terms of "definite" vs. "indefinite" readings, rather than "indexical" and "existential," but this difference in terminology does not seem of much consequence for our current concerns.

By this test, some elements which have been discussed in the literature (notably Perry (1986) and subsequent works) as contextually interpreted appear actually to be existential:

(179) a. It's raining, but I don't know where.
b. The sun is shining, but I don't know where.

This suggests that uses of *It's raining* do not express the proposition that it is raining at *l*, where *l* is some particular location determined in context, but merely that it is raining somewhere — though of course the quantification is over some contextually limited domain of relevant locations.

5.7. Conclusion

In this chapter we have reviewed various ways indexicality can be hidden or disguised, along with related phenomena. I sketched various ways of incorporating hidden or disguised elements into our current grammar: through phonological reduction, pro-drop, syntactic ellipsis, constructional indexicality, and sublexical indexicality. Although there are interesting and important differences among all these cases, and they do not all always involve indexicality, when they do, they all have a similar effect on interpretation, and one which is analogous to the effect of overt indexicals: they produce a variation in semantic content according to pragmatic context.

To the extent that an analysis positing hidden or disguised contextual sensitivity is correct for a particular example, this variation in content may be seen by checking speaker intuitions about contradiction, entailment, and related notions. For example, if in one context John says "Mary doesn't work hard, but Bill does" and in some other context Susan says "John sings beautifully but Bill doesn't," we do not take them to be contradicting each other, even though John says *Bill does* and Susan says *Bill doesn't*. Likewise, if in one context John reports the performance of his class at the University of Illinois by saying "Every student passed the exam," and in a different context Mary clarifies the status of her friend Bill at the University of Rochester by saying "Bill is a student," we do not take these assertions as jointly licensing the inference that Bill passed the exam. In both these cases we can see that the content of at least some of the expressions are different in the different contexts described, so no contradiction or entailment relation is produced, despite the syntactic forms of the sentences involved.

Returning briefly to the issue of personal taste sentences and other examples concerning matters of opinion rather than fact, we can see that appealing to hidden or disguised indexicality leaves us with precisely the same problem we saw when we first consider indexical analyses in Section 2.1. — if we try to account for the intuition that a sentence like *Licorice is tasty* varies in truth value from person to person by claiming that it involves some sort of indexical element which can be fixed differently for different people, we create a problem for ourselves in explaining why, if one person says "Licorice is tasty" and another says "Licorice is not tasty," we understand them to be contradicting each other. Conversely, if we account for the intuition of contradiction by saying that each speaker asserts the negation of the content asserted by the other speaker, then we cannot also explain the intuition of interpersonal variation in truth value by appealing to indexicality.

Beyond these by now familiar problems, we also found that an analysis which appeals to hidden but syntactically present indexicals faces difficulties in that these putative hidden elements

do not produce crossover effects. It is sometimes argued that such elements are needed in syntax to account for "bound" readings of sentences like *Every customer received a tasty dish*, but these readings can be obtained without any appeal to hidden arguments through the use of a μ-operator. A separate argument claims that a syntactically represented hidden argument to *tasty* and *fun* is needed to control PRO in infinitival complements, but on closer inspection the pattern of interpretation turned out not to require PRO at all. I conclude that there is no compelling syntactic reason to posit such hidden arguments to predicates like *fun* and *tasty*.

Chapter 6 : Pragmatics of Truth Assessment

The analysis given in Chapter 4 presents a system in which (uses of) sentences are assigned truth values relative to contexts of assessment, contexts of assessment are assumed to provide values for a perspective parameter, and each perspective consists of an individual, a time and a world. Such a system must be complemented with an appropriate pragmatic theory before it can be evaluated against linguistic data. Our data in this case consists of speaker intuitions about the truth conditions of English sentences (and related matters such as contradiction, entailment, etc.) Until the system is paired with some account clarifying which contexts of assessment provide which values for the parameters, and of how this specification relates to the practical task of assessing whether someone has spoken truthfully or not, we cannot test whether it conforms to speaker intuitions or not. That is, we have to pair our semantic theory with a pragmatic theory that explains how contexts of assessment provide particular values for their parameters, and how people go about assessing the truth values of each other's assertions.

6.1. Contexts of use and parameter values

The need for an account of how contexts "provide" values for parameters is not specific to relativism, but general to any semantic theory which makes use of contextually fixed indices relative to which denotations and truth values are assigned. If we assume that each context of use provides values for parameters like $speaker_{u,\varphi}$, $time_{u,\varphi}$, etc., and that these values play a role in the interpretation of indexical expressions such as *I*, *now*, *here*, etc., the issue naturally arises of what principles determine which contexts will provide which values for these parameters.

The simplest hypothesis would claim that these are fixed "automatically": $speaker_{u,\varphi}$ is the person who says φ, $time_{u,\varphi}$ is the time when he or she says it, etc. But we have already seen that some parameters have their values fixed at least in part by the actions or intentions of the speaker: $Ref_{u,\varphi}$, $addr_{u,\varphi}$, $proximity_{u,\varphi}$, etc. For this reason Kaplan (1989) drew a terminological distinction between "true demonstratives," which must be accompanied by a demonstration (such as a pointing gesture) to determine their referents, and "pure indexicals," for which any demonstration is superfluous or merely for emphasis; Perry (2001) draws an analogous distinction between "discretionary" and "automatic" indexicals.

However, even so-called pure indexicals sometimes seem sensitive to speaker intentions in a way that suggests that their referents are not fixed so automatically as one might at first suppose — a point already forcefully argued by Nunberg (1993), Mount (2008) and others. The much-discussed[101] example of a recorded answering-machine message "I am not here right now" shows that the present tense and adverb *now* are not always fixed to the time when these expressions are spoken.[102] Special usages like the "historical present" suggest much the same thing. The short discourse in (180), though written entirely in present tense, is naturally interpreted as narrating a series of past events:

[101] See Sidelle (1991), Predelli (1998), Predelli (2005), Weatherson (2002), Cohen (2013) among others.

[102] Sidelle and Cohen's view that such examples involve making an utterance "at a distance," so that the time of utterance is the time when the recording is played back rather than the time when the recording is made, still denies that the present tense and *now* are automatically anchored to the time when the person using these expressions engages in the actual phonetic activity of speaking.

(180) So then I say to him, "John, why don't you just shut up?" But he doesn't pay any attention to me at all. He just keeps talking, and pretty soon I'm ready to punch him in the nose…

Even the first person singular pronoun is sometimes used to refer to someone other than the person who speaks or writes it. If I put a tag reading "Eat me" on a piece of cake in the refrigerator, the message is not in any sense a suicidal invitation to cannibalism. Such uses of language are playful, and perhaps in some sense even involve pretense; but it can hardly be denied that the correct interpretation of the tag is one in which the pronoun *me* refers to the cake.

Or, consider the case of a memoir written by John Jones, "with the assistance of Mary Smith."[103] The memoir contains multiple sentences containing first person singular pronouns, all obviously intended to refer to John Jones. Some of these sentences were actually composed by Mary Smith, but one hesitates to say that the memoir is true only if *she* has the properties they ascribe to the referent of the pronoun. The hesitation is even stronger if the wording of these sentences was composed jointly by John and Mary in collaboration, rather than by Mary alone. Do we count any such sentence as exhibiting presupposition failure, because it contains a first-person *singular* pronoun but was jointly authored? It seems unlikely that we can draw a clear, principled distinction between cases where John composes a sentence, taking Mary's advice into account as he does so, and cases where they compose the sentence jointly, so that the two of them serve collectively as author of the sentence; but such a distinction seems crucial to determining what content is expressed, if we take the referent of a first-person pronoun to be fixed automatically to the author.

If, on the other hand, we allow for the possibility that a context of use might provide John as the value for the $speaker_{u,\varphi}$ parameter, even though Mary actually composed (or helped compose) the wording of some sentence use φ containing a first-person pronoun, we can obtain intuitively correct truth conditions for the memoir, and the theoretical question of how to distinguish collective authorship from individual authorship with outside aid and advice loses some of its urgency.

Likewise, if we allow for the possibility that John records his answering machine message in some context of use which yields a value for the $time_{u,\varphi}$ parameter which is later than the time when he records the message,[104] we can obtain intuitive truth conditions for the answering machine case; and if we allow the value of the $time_{u,\varphi}$ parameter to come earlier than the time when at utterance is produced, we can interpret examples like (180) properly.[105]

Even in the example of the note on the cake, I think it is not unreasonable to claim that the $speaker_{u,\varphi}$ parameter is fixed to the cake. Indeed, it seems we *must* take this position if we believe that the speech act which was performed directly is an invitation to eat the cake (not to eat the person who places the card), if what we mean by "content" is the object of the direct speech act, and if we assume that content is compositionally assigned. Certainly in this case the author

[103] The quote marks indicate that this assistance is publicly acknowledged.

[104] Or perhaps *multiple* values for this parameter, corresponding to various times when the message might be played back. See Predelli (1998) for a brief but useful discussion of this complication.

[105] It should be noted that temporal expressions do not always shift uniformly in the historical present, as evidenced by the use of *then* and *pretty soon* in (180). It would take us too far afield to give a detailed analysis of the this pattern here, but I do not think it threatens the point that the parameter-values yielded by a context of use are not a simple, automatic reflex of who authors or pronounces a sentence, at what time, and in what location.

composes the wording on the card "as though" the cake had written it — a fact we can account for simply by acknowledging that the value of the $speaker_{u,\varphi}$ parameter is normally or conventionally fixed to the author/performer of an utterance. But recognizing that such norms or conventions are sometimes violated does not require us to claim that the "real" semantic content of the note in context is derived by treating the pronoun *me* as denoting the person who wrote the note, or that the act of leaving the note is a direct invitation to eat that person. Language use is a creative activity, and speakers often violate norms for special effect. Recognizing this fact does not place this example outside the subject matter of rule-governed, compositional semantics — on the contrary, the reading in question can be derived in a perfectly compositional way, in complete conformity to the semantic rules of the language. The norm which is violated is not part of the grammatical system *per se* — it is not a rule for compositionally assigning denotations relative to indices — but is instead a principle of how that system is deployed in the practice of speaking: a conventional way of relating the parameter values associated with an index to the real, concrete situation in which a linguistic expression is used.

This conception of the relation between parameter values and contexts, in which the formal semantic system of a language simply assigns denotations relative to indices, and the fixing of values for the corresponding parameters in real contexts is seen as a separate issue governed by its own set of pragmatic norms and conventions, is, I think, essentially the position argued at length in Predelli (2005). But something like this distinction is implied by any linguistic theory which differentiates the abstract grammatical system of a language and the principles by which speakers make use of this system in actual speech, and which assigns denotations relative to contextually determined parameter values.[106]

How then, are parameter values fixed, if not "automatically"? As the examples just discussed suggest, an account of this must address the practice of speaking "as though" the context were different than it actually is — as though a piece of cake were the author of a note, or John were the sole author of his memoirs, or events in the past were happening now, etc. This kind of practice, I claim, involves the purposeful violation of pragmatic norms governing the contextual determination of parameter values. These pragmatic norms must not be seen as constitutive rules of the linguistic system, but simply as conventionally established defaults, whose status as norms may be exploited to produce particular rhetorical effects: One can narrate past events more vividly and with a greater sense of immediacy by fixing $time_{u,\varphi}$ to the time when those events took place instead of the time when one composes the utterance; one can convey humor or whimsy by setting $speaker_{u,\varphi}$ to a piece of cake, etc. When such practices occur frequently enough, they may become conventional themselves, and the pragmatic norms governing the language are updated to accommodate them.[107]

To some extent, then, parameter values are under the intentional control of the speaker (that is, of the person who actually authors or performs the utterance, whether or not that person

[106] For more on the distinction between grammatical theory and the theory of what people do when using language, see Lasersohn (2009a), Lasersohn (2012). The fact that the theory developed in this book assigns contents and denotation to "uses" of expressions rather than to the expressions themselves does not in any way threaten the distinction between the theory of grammar (including denotation assignment) and the theory of how the grammatical system of a language is employed in the practice of speaking, listening, etc.

[107] The historical present, for example, seems clearly established as a conventional usage, and should probably not be considered a violation of pragmatic norms in present-day English.

serves as the value of $speaker_{u,\varphi}$). But it would be too simple to say that indexicals always simply refer to whatever the speaker intends. If on September 2nd John mistakenly believes it to be September 3rd, and uses the word *tomorrow* intending to refer to September 4th, this intention will not cause him to speak truthfully when he says *Tomorrow is September 4th*. Nor could he use this sentence truthfully if he knew it was September 2nd when he spoke, but intended to use *tomorrow* to refer to September 4th in an attempt at deceiving his audience into believing it was currently September 3rd. In either of these cases, John's intentions fail, and he actually refers to September 3rd in using the world *tomorrow*, so that he says something false.

These examples do not show that *tomorrow* is immune to the influence of speaker intentions in fixing its reference. If John writes a note on September 2nd, intending it to be read on September 3rd, it would be perfectly acceptable for him to use the word *tomorrow* to refer to September 4th (and doing so would decrease the sense of separation between himself and his intended reader). The difference appears to be that in this last example, John violates a parameter-setting norm in a way which is *purposeful*, *public*, and *cooperative*. The note is worded deliberately to reflect a value for $time_{u,\varphi}$ other than the time when the note is actually written, in order to create a specific rhetorical effect. Moreover John need not hide this fact from his addressee(s); even if he does not specifically intend the audience to recognize that the norm is being violated, if they do recognize this fact, his communicative purpose is not undermined. Nor is the norm violation intended to deceive, or to be otherwise uncooperative (in the sense of Grice (1975)).

Let us suppose, then, that for any expression use α, normally $speaker_{u,\alpha}$ is the person who authors and performs[108] the utterance of α in *u*, $time_{u,\varphi}$ is the time when this utterance takes place in *u*, $addr_{u,\varphi}$ is the person(s) to whom it is addressed in *u*, etc. Unless, in *u*, the speaker purposefully, publicly and cooperatively intends otherwise, the values are fixed by default in accordance with these norms. But if the speaker does appropriately intend a value other than the normal one for any such parameter, then simply by virtue of this intention, the value is fixed in accordance with that intention. In this case, the utterance is made "as though" the parameter value fit the normal pattern, even though the norm is violated.

6.2. Contexts of assessment and parameter values

We have no reason to expect that contexts of assessment will provide values for parameters in any more of an automatic way than contexts of use. I suggest that in fact, contexts of assessment provide values for parameters in much the same way as contexts of use do: a system of pragmatic norms establishes default values, but in appropriate circumstances, the person performing a truth assessment may intentionally assess a sentence use relative to non-normal parameter values, in effect assessing "as though" the context of assessment were different than it actually is.

It seems clear that in the typical case, when an individual *x* assesses a sentence content Φ for truth or falsity at time *t* in world *a*, the context of this assessment provides *a* as the value of its world parameter, and $\langle x, t, a\rangle$ as the value of its perspective parameter — the judge is the person performing the assessment; the time is the time when the assessment is performed, and the world is the world where the assessment is performed.[109] But certain examples seem to involve assessment

[108] Normally, of course, the same person both authors and performs an utterance.

[109] Of course a sentence content might be assessed in more than one world, but we are assuming here that each act of

to parameter values other than the normal ones. To see this, consider two ways of assessing a use of the sentence *The berries stank*:

Suppose at 3:00 pm on July 17, 2012, the berries give off an odor which John, at that time, finds quite pleasant. However, as time passes, John's tastes change; by his standards of 3:00 pm, July 17, 2013, that very same odor is disgusting — although the berries have long ago stopped emitting it. (We may suppose they have been eaten in the meantime and no longer even exist as berries.) By the rules of Chapter 4, a use φ of the sentence *E The berries*$_{1,2}$ *stank*$_3$ is true relative to a perspective with John as the judge and 3:00 pm, July 17, 2013 as the time (and $w_@$ as the world). This seems right — one can easily imagine John sincerely saying (181), as he recollects the now-disgusting odor of the berries:

(181) The berries stank, although at the time I liked that kind of smell.

We may assume that an assertion is sincere only if the speaker assesses the asserted content as true, so this example illustrates truth assessment in a context where the time parameter is fixed to the time when the assessment takes place.

However, a second way of assessing the truth of the sentence also seems to be available, according to which John may judge the sentence true if, at some time *t* prior to the time of use, the berries stank according to John's tastes at *t* — even if John is assessing the sentence at some time other than *t*. This way of assessing the sentence can lead to a judgment that it is *false*: if the pragmatically relevant past times only include 3:00 pm on July 17, 2012 and other times at which John's tastes are as they were then, there will be no *t* among them at which the berries give off an odor which is unpleasant according to John's tastes at *t*. In this case, one can imagine John sincerely saying:

(182) The berries smelled good, although I don't like that kind of smell any more.

This is so even though this assertion appears to contradict (181), which it was claimed he could also sincerely assert.

If we make the assumption that whenever an individual *x* assesses a sentence use φ for truth or falsity at time *t* in world *a*, *x*'s assessment is correct iff it assigns the sentence the truth value which φ has relative to a context of assessment whose world is *a* and whose perspective is $\langle x, t, a\rangle$, then examples like (182) become mysterious. In contrast, if we assume that language users can assess a sentence *as though* from an earlier time, we can make perfect sense of examples like (182).

There are many details to be worked out and alternative analyses to be considered, of course; but it may be worth considering a couple more cases before turning to these matters. In some examples, it is not the time parameter which appears to be fixed in a non-normal way, but the judge parameter. For example, when a person A performing a truth assessment must make a decision for another person B, based on the testimony of a third person C, it may be appropriate for

---

assessment takes place in a unique context, and that each context of assessment provides a unique value for the world parameter, so assessments in our sense are world-bound — the relevant notion is that of an assessment-in-a-world, not that of an assessment with trans-world identity, taking place differently in different possible worlds. This is not to deny that the latter notion might also be legitimate and useful, of course.

A to assess the truth of that testimony relative to B, and not just relative to himself or herself. If Mary tells John that the merry-go-round is fun, and John is trying to decide whether to buy his young son Bill a ticket, he will naturally want to consider whether what Mary said is likely to be true relative to *Bill*, not relative to himself. When he discovers that Bill does in fact enjoy the merry-go-round, he may even thank Mary for telling the truth (and not lying just to get him to waste his money, as she sometimes does). He might also pass along the information, sincerely telling another parent of a young child that the merry-go-round is fun — all the while recognizing that he himself would find the merry-go-round tedious and unpleasant. In such a case, truth of Mary's claim relative to John is not pragmatically relevant to the matter at hand, while truth relative to Bill is relevant; so it is natural for John, in assessing the sentence, to perform the assessment with Bill as the judge. That is, we may regard the context of assessment as fixing Bill as the value of the $judge_c$ parameter, even though it is John, not Bill, who performs the assessment.

Or, consider a case where an assessor must evaluate a claim that some particular event was fun, but where the assessor did not personally participate in the event. For example, suppose Mary waits to the side while John rides the roller coaster. Afterwards, he reports to her "That was fun!" How will Mary assess the truth of this claim? If we assume that for any context of assessment *c*, $judge_c$ is the person performing the assessment, then Mary must assess John's claim according to her own tastes; that is, she must assess whether the event of John riding the roller coaster was fun by her tastes. Note that this is a different question from whether Mary would find it fun to ride the roller coaster herself — that would involve a different, hypothetical event. Nor is it the same question as whether Mary enjoyed watching John ride the roller coaster (or having him absent, etc.) — these are not questions about whether the event of John riding the roller coaster was fun, but about whether the event of Mary watching it (or having a respite from John's annoying presence, etc.) was fun. In fact it is quite hard to see how one could evaluate, relative to one's own tastes, the claim that an event that one did not participate in was fun — I would suggest it is impossible: where USE(α, *fun*) and *e* is a particular event, $[\![\alpha]\!]^{u,w,\langle x,t,a\rangle}(e)$ is defined only if *x* is a participant in *e* in *w*.[110]

However, if we allow for the possibility that the value of the $judge_c$ parameter can sometimes be fixed to a non-normal value — that a person can perform a truth assessment relative to someone else's tastes — then it is easy to see how Mary might assess John's claim: We simply regard Mary's assessment as taking place in a context *c* such that $judge_c$ = John (and $time_c$ is the time during which he rode the roller coaster), even though it is Mary who does the actual assessing in *c*. In fact this conforms to the assessment most of us would make in Mary's place: we regard John's claim as true iff riding the roller coaster was fun *for him*.

6.3. Adopting a stance

In the normal case, where a person assesses the truth or falsity of a sentence content relative to his or her own current perspective, let us say that the assessment is performed *autocentrically*, or that

[110] Of course many things other than particular events can be fun: general types of events, concrete objects like roller coasters, etc. At the most basic level, perhaps, it is *experiences* which are fun, and the predicate holds of these other kinds of things only by virtue of their association with certain kinds of experience; but this does not mean that *fun* "properly" applies only to experiences, or that talk about other things being fun is elliptical for talk about experiences being fun, etc.

the assessor adopts an *autocentric stance* in performing the assessment.[111] If a truth assessment is performed relative to any other perspective, let us say that it is performed *exocentrically*, or that the assessor adopts an *exocentric stance*.[112] If a person *x* performs an autocentric assessment of a sentence content Φ in some context *c*, at time *t* and world *a*, this means that $judge_c = x$, $time_c = t$, and $world_c = a$. But if *x* performs an exocentric assessment of Φ in *c* at *t* and *a*, then $judge_c \neq x$, or $time_c \neq t$, or $world_c \neq w$.

Just as we saw that there are constraints limiting when a context of use may provide non-normal values for parameters like $speaker_{u,\varphi}$ and $time_{u,\varphi}$, we should expect that there may also be constraints on when a context of assessment may provide non-normal values for parameters like $judge_c$ and $time_c$. The two cases are not perfectly analogous, however. Expressing semantic content is normally a purposeful, public, cooperative act, and the content of an uttered expression must be fixed in a purposeful, public, cooperative way. For that reason, any departure from the normal values for parameters involved in fixing content must also be purposeful, public and cooperative. But assessing the truth value of a sentence content is not by nature a public or cooperative act. We should therefore not expect that the constraints governing these parameter settings will parallel too closely those involved in assigning content.

Even if not necessarily public or cooperative, truth assessment is generally a purposeful act. We should therefore expect that exocentric assessment should only be possible when it serves some specific purpose. Where evaluating the truth of a statement relative to some other person or some other time does not serve the purposes at hand, we may regard it as illegitimate. If it does serve those purposes — particularly if it serves them better than an autocentric assessment would — then exocentric assessment is allowed. Thus we find examples like the one discussed in the last section, where a father assesses a claim like *The merry-go-round is fun* relative to his son, because he must decide whether or not to buy his son a ticket to ride it. An autocentric assessment would have been irrelevant to the task at hand. In contrast, if the father were trying to decide whether he himself should ride the merry-go-round (alone), assessing this claim relative to his son would be irrelevant to his decision, so an exocentric assessment with his son as the judge will not be appropriate; in this case his judgment will be based on his own tastes, not his son's.

In addition to requiring exocentric assessment to serve the pragmatic purposes for which the assessment is made, we must regard it as legitimate only when the assessor reasonably expects the content being assessed has a truth value relative to the perspective employed in the assessment. Thus one cannot assess a claim that some particular event was fun relative to a person who did not participate in it (or relative to a time when the event did not occur, etc.)

These two conditions — that an assessment may be exocentric only if its exocentricity serves the purposes for which the assessment is performed, and that the assessed content must be reasonably expected to have a truth value relative to the perspective employed in the exocentric assessment — are really just cases of broader conditions on truth assessments as pragmatic acts

---

[111] In Lasersohn (2005) I used the term *autocentric perspective*, but I now prefer *stance* as better capturing the idea I had in mind, especially since we are already using the term *perspective* in a very different sense. The choice of *autocentric* over the more conventional *egocentric* was made initially because *autocentric* lacks the negative connotations which *egocentric* typically carries in non-technical usage. We use it here to remain consistent, and because we will use the term *egocentric* in Chapter 10 for a somewhat different concept.

[112] Here again, *allocentric* might be a more conventional choice of terminology, but I prefer *exocentric* because some authors use *allocentric* in a way that is closer to what I will later call *acentric* than to what I intend by *exocentric*.

more generally. Even autocentric assessments are disallowed when they would be irrelevant to the purposes at hand, or involve contents which clearly lack truth values relative to the assessor's current actual perspective.

But if these conditions are not special constraints on exocentric assessment, but broad conditions on assessment in general, why do we regard autocentric assessment as the norm, and exocentric assessment as exceptional? I suggest it is not because there are any special constraints which apply only to exocentric assessments, but because autocentric assessment is in some sense more cognitively primitive, more closely related to a fundamental concept of *reliability*, and relies less on counterfactual imagination than exocentric assessment. These matters will be the topic of Chapter 10. For now, let us simply assume that autocentric truth assessment is the norm, so that people will generally assess relative to their own current actual perspectives unless they have some special reason to do otherwise.

6.4. Truth assessment and the adicity of *true*

So far I have been somewhat obscure about what, exactly, it means to assess a sentence content as true or false, and careful readers may have noticed an apparent inconsistency between this kind of wording and the semantic theory developed in Chapter 4. There, we employed several different notions of truth[113] depending on whether they applied to contents, sentences or uses, and according to whether they were monadic or relativized in various ways — to contexts of use, contexts of assessment, indices, etc. Which of these, if any, is the relevant notion for "truth assessment" in the sense discussed in this chapter? If we take seriously the idea of assessment as judging something to be "true" or "false" (rather than, say, "true relative to context of assessment *c*" or "false relative to context of assessment *c*"), then it would appear that truth assessment of a content Φ means judging Φ to be true or false *monadically* — that is, with *true* and *false* functioning as one-place predicates. But so far we have been reserving monadic truth predicates for cases where a sentence content (or use) is true or false objectively, the same for everyone. "Subjective" truth — the kind that applies to personal taste examples — was analyzed in terms of predicates of the form "true relative to" — that is, multi-place predicates.

We need a notion of assessment that makes sense for contents which are true or false only "subjectively," and consistent with the analysis of such contents developed in Chapter 4. But it makes no sense to judge a content to be "true" if *true* is a predicate with more than one argument place. In general, it makes sense to judge that *x* is *P* if *P* is a one-place predicate and *x* is the sort of thing *P* might apply to, but it makes no sense at all to judge that *x* is *R* where *R* is a 2-place predicate (or a 3-place predicate, etc.) We can solve this problem by introducing a one-place truth predicate which applies to the contents of sentences of personal taste (in addition to other sentences). Such a predicate should fit the pattern of **true** in (183):

(183) For any context of assessment *c*,
  a. ⌜Φ is **true**⌝ is true relative to *c* iff Φ is true relative to *c*.
  b. ⌜Φ is **false**⌝ is true relative to *c* iff Φ is false relative to *c*.

[113] See (139), (140), (141), (142), (143), (144).

This pattern renders "is **true**" redundant, roughly as expected under deflationary, minimalist or disquotational theories of truth. To judge that Φ is **true** is to judge that Φ. I suggest that this is the relevant notion of truth in developing an account of what it means to assess sentence contents (included the contents of personal taste sentences) as "true." But since this predicate is redundant, it is eliminable in principle, and does not do any real theoretical work. Performing a truth-assessment of a sentence content Φ just means assessing whether Φ; there is no essential appeal to the predicate **true**.

It is also worth noting that this predicate was not used at all in the compositional semantic theory given in Chapter 4. The compositional apparatus of our semantic theory was described in a completely non-relativistic metalanguage. This is as it should be: the analysis should be stated in clear, objective terms; it is not a matter of taste how the semantic rules of the language operate.[114] It is only in explaining the pragmatic act of assessing whether a sentence content is true that we have appealed to a truth predicate which itself behaves in a relativistic fashion — and even then only in a way that shows that it is completely redundant, hence theoretically dispensable.

6.5. Exocentricity and indexicality

In the analysis advocated here, contents expressed by sentences are sometimes assessed as true or as false based on parameter values provided "non-automatically" — that is, based on the intentions of the person performing the assessment. The value of the judge parameter need not be the person performing the assessment; the value of the time parameter need not be the time when the assessment is performed, etc. This allows us to account for examples like those discussed in Section 6.2. where the truth value of a statement seems to depend on the tastes of someone other than the assessor, or the assessor's tastes at some time other than the time of the assessment.

This approach may be contrasted to one in which sentences containing predicates like *fun* or *tasty* are ambiguous between two readings: a relativist reading in which the truth value depends on parameters fixed by the context of assessment, and a contextualist reading in which the parameters on which the truth value depends are all fixed by the context of use.
Such an analysis is developed in Stephenson (2007a), Stephenson (2007b), for example.

In Stephenson's analysis, predicates of personal taste have an argument place which can be filled by prepositional phrase headed by *for* (as in *The roller coaster is fun for Mary*) or by either of at least two hidden pronoun-like elements. The first of these, notated PRO_J (and identified with the PRO subject of certain kinds of infinitival clauses) receives the semantics in (184):[115]

(184) $[\![\text{PRO}_J]\!]^{c,w,t,j} = j$

Here, *j* is the value of the individual index, fixed to the judge of the context, as in Lasersohn (2005). Hence if *fun* or *tasty* takes PRO_J as its argument, a sentence like *Roller coasters are fun* or *The chili is tasty* will receive a relativistic reading, under which it is true relative to a context (of assessment) iff roller coasters are fun or the chili is tasty for the judge of that context.

But this argument place can also be filled with a hidden pronoun with a more conventional

[114] Of course one analysis of the rules might be more to the taste of some theorists and less to the taste of others, but this does not mean that it is a matter of taste what the correct analysis is.
[115] See Stephenson (2007a), Stephenson (2007b).

semantics like that of an ordinary indexical pronoun. In that case, *Roller coasters are fun* will express a content which is true iff roller coasters are fun for the referent of the hidden pronoun. As with other indexicals, this referent is fixed by the context (of use) by the referential intentions of the speaker; the sentence will express a non-relativistic content which is true or false (for everyone) according to whether the roller coaster are fun for the individual or group intended by the speaker.

Sentences containing predicates of personal taste are thus analyzed as systematically ambiguous in Stephenson's analysis, between a relativistic reading and an indexical reading. The kinds of examples for which I have suggested here an analysis in terms of exocentric truth assessment are treated instead as involving indexical reference. For example, in the case where John considers Mary's claim that the merry-go-round is fun as he tries to decide whether to buy a ticket for his son Bill, the analysis would claim that Mary's use of the sentence expresses a content which is true or false (absolutely) depending on whether the merry-go-round is fun for *Bill* (or, perhaps more plausibly, for some class of people which is likely to include Bill). Because the truth of this content does not vary with the judge parameter, the value of this parameter can be ignored in truth assessment. John might assess Mary's claim in a context which supplies himself as the judge — that is, he might assess is autocentrically — but it will still be Bill (or the class of people intended by the speaker) for whom the merry-go-round must be fun in order for the claim to be true. This technique thus allows Stephenson to eliminate the category of exocentric truth assessment and claim instead that all assessment is autocentric. This in turn allows a simpler account of how contexts of assessment relate to parameter values: the value of the judge parameter may be stipulated always to be the person performing the assessment.

Extending this analysis to the full range of examples for which I have suggested exocentric assessment would require some adjustments. Example (182), in which John assesses a claim relative to his own earlier tastes rather than his current tastes, would require some sort of temporal indexicality in addition to the nominal indexicality which Stephenson discusses.[116] But such adjustments would be straightforward to make: for example we could treat the hidden argument place of predicates of personal taste as filled not with an individual or group, but with a "perspective" in our current sense of a triple $\langle x, t, a \rangle$, where t is a time. Assuming speakers can make indexical reference to perspectives, the intended reading falls out straightforwardly.

The two analyses make different claims about the contents expressed by sentences of personal taste, and about the reasons why someone other than the person performing a truth assessment sometimes seems to be the one whose tastes matter to that assessment. In Stephenson's analysis, personal taste sentences are systematically ambiguous, because there are two distinct covert elements which can enter the syntactic derivation — an indexical element whose referent is fixed like those of other indexicals, and PRO_J, which induces a relativistic interpretation. The semantic content varies with this syntactic distinction, but the pragmatics of truth assessment is kept simple and automatic: the judge of a context is the person performing a truth assessment. In the analysis I have advocated, there is no syntactic or semantic ambiguity like that suggested by Stephenson; whether assessed autocentrically or exocentrically, the syntactic representation and semantic content is the same. However, the pragmatics of truth assessment is analyzed as more complex: the judge of a context may sometimes be a person other than the one performing a truth

[116] Stephenson's analysis, like my own in Lasersohn (2005), does not make provision for changes in a person's taste over time — presumably as a simplification and not because of any serious theoretical commitment.

assessment (but still, of course, the person on the basis of whose tastes the assessment is performed).

It is perhaps worth emphasizing that both analyses admit a relativist reading conforming to the main thesis of this book, namely that sentences concerning matters of opinion (including matters of taste) express contents which are true or false only relative to certain parameters. That thesis does not depend on the choice between the two options under consideration, and in that sense, the present issue is relatively minor. Nonetheless it may be interesting to consider what arguments may be given to support either analysis.[117]

As Stephenson points out, her analysis predicts that faultless disagreement should never arise in cases where the speaker intends his or her assertion to be assessed based on the tastes of someone other than the person performing the assessment. In such a case, because exocentric assessment is unavailable in her system, any sensitivity of truth value to the tastes of another person must be because the sentence contains some overt or covert expression referring to that person, rather than PRO_J. In this case the sentence will express a content whose truth value does not vary with the judge, and we may say that it is true or false absolutely (that is, objectively), and not just relative particular values of the judge parameter. Since the content is true or false objectively, any disagreement about whether or not it is true cannot be faultless. In contrast, under the analysis advocated here, a speaker might assert a sentence with a relativistic content, expecting and intending that it may be assessed by one person on the basis of the tastes of a different person. Because the truth value of that content varies with the judge, two parties may faultlessly disagree about it, provided their assessments occur in contexts that provide different values for the judge parameter.

Stephenson suggests that observed pattern is the one predicted under her analysis: faultless disagreement is impossible when the speaker intends his or her assertion to be assessed on the basis of the tastes of someone other than the person performing the assessment. She offers the following example to illustrate:[118]

(185) Sam: The tuna is tasty.
Sue: No it isn't! It isn't tasty at all!

In this example, Sam and Sue might be disagreeing whether the tuna is tasty for their cat, or they might be disagreeing based on their own tastes. However, Stephenson suggests, there can be no disagreement in the mixed case, where Sam means that the tuna is tasty for the cat, and Sue is expressing her own taste in her response. If we read the example this way, Sue's utterance sounds infelicitous, since it takes the form of a disagreement without actually disagreeing.

I agree that it would be odd for Sue to argue with Sam based on her own tastes, if he were making the assertion based on his assessment of *The tuna is tasty* as true relative to a context which provided the cat, not Sam, as the judge. But I do not believe it is correct to say that faultless disagreement is impossible whenever the speaker intends assessment relative to someone other than the person performing the assessment. Suppose John and Mary are discussing how their young son Bill liked the amusement park. Their obnoxious neighbor Fred overhears and butts in:

[117] The next few paragraphs repeat with minor adjustments an argument from Lasersohn (2008).
[118] Stephenson (2007a), Stephenson (2007b).

(186) Mary: How did Bill like the rides?
John: Well, the merry-go-round was fun, but the water slide was kind of scary.
Fred: Oh, it was not! Your kid is just a weenie!

Here Fred disagrees with the claim that the water slide was scary, even though John clearly intends it to be assessed relative to Bill. We get no sense in this example that Fred has misunderstood who John was referring to, and no sense that he believes that Bill actually enjoyed the water slide.

The difference between examples (185) and (186), I think, is not that the speakers contradict each other in one case but not the other, but merely that it is easier in (186) than in (185) to imagine a plausible pragmatic purpose to expressing the disagreement. Fred might be the sort of person who tries to make his own children look good by pointing out the shortcomings of other people's children, for example; or perhaps he has some vested interest in making sure that the water slide is seen as an appropriate ride for children Bill's age. But what would Mary have to gain by pointing out that her tastes conflict with the cat's? To the extent that we can imagine a plausible reason, (185) will begin to sound less anomalous, even in the mixed case. For example, if Bill is in the habit of deciding what to serve Mary based on his observations of what the cat seems to enjoy, it would be perfectly natural for Mary to express her disagreement with the cat as in (185).

Even aside from issues of disagreement and contradiction, the two analyses differ in that Stephenson's analysis predicts that a single use of a sentence must be interpreted either indexically or relativistically, with no possibility of being assessed multiple times, sometimes on the basis of the tastes of the person performing the assessment and sometimes on the basis of some other person's tastes (except if in cases of misunderstanding and the like). Yet examples of this kind are not hard to construct. Consider the case again where John assesses Mary's claim that the merry-go-round is fun, in order to decide whether to buy a ticket for his son Bill. John assesses Mary's claim relative to Bill's tastes, not his own. Yet Mary's own daughter Sue, overhearing the conversation, might assess the very same use of the sentence *The merry-go-round is fun*, based not on Bill's tastes but on her own. (Having heard her mother make the claim, she now wants to ride the merry-go-round.)

The two analyses also differ in their claims about whether such sentences are ambiguous. More specifically, under a Stephenson-style analysis, a personal taste sentence expresses a different content when intended to be assessed on the basis of one's own tastes than it does when intended to be assessed on the basis of someone else's tastes; but under the analysis I have been advocating, there is no such ambiguity; a sincere assertion might be made because the speaker assesses the content as true relative to his or her own tastes, or someone else's, but the content is the same in either case. As discussed in connection with examples (14) and (15) in Section 1.5. we can use coordinate constructions like Right Node Raising as a heuristic for evaluating claims that an expression ambiguously expresses multiple contents. When applied to our current issue, this heuristic seems to favor the claim that there is no ambiguity. For example, suppose John reports that the merry-go-round is fun (on the basis of observing his son Bill enjoying the merry-go-round, even though John knows that he himself would find it tedious). Having ridden the merry-go-round and found it too scary to be enjoyable, Sue reports that it is not fun. We may describe John's and Sue's utterances using a sentence like (187):

(187) John claimed, but Sue denied, that the merry-go-round is fun.

The clause *the merry-go-round is fun* appears just once in this sentence, and therefore should be interpreted as having a single denotation which "distributes" over the conjuncts *John claimed* and *Sue denied* — not two different denotations which serve as arguments to the two verbs. The context is opaque, so the denotation of the subordinate clause is what would normally be its content. The sentence describes this (one) content as claimed by John and denied by Sue. But in the scenario outlined, John makes his claim based on a third person's tastes, and Sue makes hers based on her own tastes. Because (187) seems like an accurate way of describing the scenario, we may take it as evidence in favor of an analysis which does not treat the sentence *the merry-go-round is fun* as ambiguous in the way Stephenson claims.

6.6. Acentric stances

If the content of a sentence concerns a matter of taste, so that it is true relative to some perspectives and false relative to others, then in order to assess it for truth or falsity, one must choose a perspective from which to perform the assessment — that is, one must adopt a stance, autocentric or exocentric. But it is also possible simply to consider a sentence content in the abstract, without performing a truth assessment. In this case, there is no need to choose a perspective; one may adopt an *acentric* stance.

Adopting an acentric stance toward a sentence content affords one a "bird's eye view" of that content — a detached, neutral prospect which does not favor any one perspective over another. It is, I think, only because we have the cognitive capacity to adopt an acentric stance that we can recognize that some disagreements are faultless. A non-human agent with less imaginative power than we have, locked in its own perspective and unable to contemplate claims except from that perspective, could draw no distinction between faultless and non-faultless disagreements. Indeed, it sometimes requires concerted imaginative effort to maintain an acentric stance, and regard the tastes of others as bearing fully the same degree of legitimacy as one's own.[119]

When we do maintain such a stance, truth-assessment of personal taste sentences becomes impossible. Assessing such sentences requires a value for the judge parameter, and settling on a value means abandoning an acentric stance. If one's stance were always acentric, and never autocentric or exocentric, it would presumably appear that personal taste sentences simply lacked truth values, and that no monadic truth predicate, not even a relativistic one as in (183), could apply to them.

[119] Because recognition of faultless disagreement depends on the possibility of adopting an acentric stance, it is perhaps worth considering the possibility that the word *faultless* (in at least one of its senses) requires a relativist analysis. We might treat a uses of sentences of the form ⌜**δ** is faultless⌝ (where **δ** is a term denoting some disagreement) as true relative to a perspective $\langle x, t, a\rangle$ only if x adopts an acentric stance towards Φ at t in a, for some Φ at issue in the disagreement denoted by **δ**. I forego working out the details of such analysis here, because it seems mistaken to me; it is a matter of fact, not opinion, whether a given disagreement is faultless.

Chapter 7 : Attitude Predicates in Relativist Semantics

In this chapter, we extend the semantic and pragmatic theory of previous chapters to deal with sentences ascribing mental attitudes. The basic strategy is simple and familiar: We treat attitude predicates as creating intensional contexts; the phrase with which an attitude predicate combines therefore denotes what would otherwise be its content. Contents are functions from world-perspective pairs to ordinary denotations. Hence a verb like *believe*, when it combines with a (finite) subordinate clause, in effect denotes a relation between individuals and sets of world-perspective pairs.

As just described, this strategy seems to differ from standard possible-worlds accounts of attitude reports only in the addition of the perspective index. There is an additional complication, however, at least if we assume the kind of pragmatics for truth assessment developed in Chapter 6: Many attitudes involve truth assessment, and truth assessment, on this account, involves adopting a stance — assessing either autocentrically based on one's own current perspective, or exocentrically, based on another perspective which serves the purposes for which the truth assessment is performed. We may expect that the bearer of a mental attitude involving truth assessment must therefore also adopt a stance in performing that assessment, and that the semantics of attitude predicates will reflect this selection of stance in some way.

7.1. Stance and belief

Consider again the case of John, who reports on his young son Bill's trip to the amusement park by saying *The merry-go-round was fun*. We may suppose that John is sincere in making this assertion — that is, he believes what he says — hence that he believes that the merry-go-round was fun. But of course this does not mean that he believes that the merry-go-round was fun for *himself*, or based on his own tastes; it may simply be that he has assessed the content of the clause *The merry-go-round was fun* as true, exocentrically based on Bill's perspective.

But now we face a problem. John might also assess the content of *The merry-go-round was fun* as false, autocentrically based on his own perspective. Having assessed this content as false, he disbelieves it. If someone simultaneously believes and disbelieves the same content, that is usually taken to mean that his or her beliefs are incoherent. But in the scenario just described, it does not seem intuitively that John is incoherent in his beliefs — the story gives us no reason to think that John is irrational, or fails to think through the relevant consequences of what he believes, or anything of the kind.

In Lasersohn (2005) I suggested an analysis of this problem based on the idea that the verb *believe* has a hidden argument place for a context. Contexts, in the framework of that paper, were like pairs of a context of use and a context of assessment; each context specified a value for the judge parameter. Because this hidden argument specified a judge, it effectively encoded the stance adopted by the believer: If the judge of the context serving as hidden argument to *believe* was identical to the subject, the sentence was interpreted as claiming that subject believed autocentrically; if it was different, the interpretation was that the subject believed exocentrically. In this way it might happen that **believe(j, c_1, fun(mgr))** and that **believe(j, c_2, ~fun(mgr))**, particularly if the judge of $\mathbf{c_1}$ is different from the judge of $\mathbf{c_2}$. In effect, John stands in a different attitude to **fun(mgr)** than he does to **~fun(mgr)** — even though both attitudes are describable

using the verb *believe* — so neither one of these attitudes must be regarded as incoherent.

I still believe something roughly like this analysis is correct: Belief involves truth assessment, which requires adopting a stance — autocentric or exocentric, according to the context of assessment. The contents which an agent believes by virtue of assessments relative to a given context *c* form a unified "body" of belief and may be regarded as coherent or incoherent, depending on whether they contradict each other. But it is part of our intellectual capacity to entertain, assess, and even believe sentence contents from multiple perspectives — hence to maintain multiple such bodies of belief. In a limited way, contents belonging to different bodies may contradict each other, but we do not generally regard this as incoherence or irrationality.

However, there is a serious problem with the analysis of Lasersohn (2005). The analysis appealed to an extra argument place in the verb *believe*, filled by a hidden indexical making reference to a context. There is, as far as I can see, no syntactic evidence at all for such an argument. We might explain the absence of syntactic evidence by appealing to sublexical indexicality (in the sense of Section 5.5.) so that the hidden argument formed part of the internal lexical semantics of *believe* rather than functioning as a separate syntactic unit, but then, the content of *believe* should vary with the referent of the hidden argument. Unfortunately, there is considerable evidence against such variation in content. If the content varied in the way suggested, then any single occurrence of *believe* should be interpreted with the hidden indexical fixed to one particular value. In this case, the reported attitude should be one based on assessment relative to one particular context, hence one particular judge. Yet we find that when *believe* appears in a coordinate construction with multiple subordinate clauses, it may be used to describe beliefs based on autocentric truth assessment with respect to one of the clauses, and exocentric with respect to another:

(188) John believes wholeheartedly that the chili will be tasty, and halfheartedly that the hot dog will be tasty.

Suppose John is preparing dinner for himself and his son Bill. He is planning to eat the chili himself (and not serve any to Bill), and to serve the hot dog to Bill (not eating any of it himself). He is making the chili just the way he likes it, and confidently expects he will enjoy the taste. He knows Bill is not a big fan of hot dogs, but tries to convince himself that Bill will like this particular one — that is, he half-heartedly believes that Bill will find it tasty. In this scenario, John (wholeheartedly) believes the content of *the chili will be tasty* autocentrically based on an assessment relative to himself, and (halfheartedly) believes the content of *the hot dog will be tasty* exocentrically based on an assessment relative to Bill. But the verb *believe* appears only once in the sentence, and should contribute its content only once in determining the content of the sentence as a whole. We ought not analyze the content of *believe* as varying with the stance ascribed to the subject.

Let us suppose, then, that *believe* does not have a hidden argument place as suggested in Lasersohn (2005), or otherwise make indexical reference to the context, judge or stance adopted by the subject. We add the various forms of *believe* to our lexicon as in (189), along with the complementizer *that*:

(189) a. $believe_{inf} \in (\text{ACC}\backslash(\text{V}\cap\text{INF}))/\text{C}$

b. $believe_{1sg_i} \in (((\text{NOM} \cap 1\text{P} \cap \text{SG})\backslash \text{V})/\text{C}) \cap \text{PRES}$
c. $believe_{2sg_i} \in (((\text{NOM} \cap 2\text{P} \cap \text{SG})\backslash \text{V})/\text{C}) \cap \text{PRES}$
d. $believes_i \in (((\text{NOM} \cap 3\text{P} \cap \text{SG})\backslash \text{V})/\text{C}) \cap \text{PRES}$
e. $believe_{pl_i} \in (((\text{NOM} \cap \text{PL})\backslash \text{V})/\text{C}) \cap \text{PRES}$
f. $believed_i \in ((\text{NOM}\backslash \text{V})/\text{C}) \cap \text{PAST}$
g. *that* $\in$ C/V

Adapting the technique of Hintikka (1969) to our current framework, let us assume that for each individual *x*, at a given time *t* and world *w*, certain worlds are compatible with everything *x* believes at *t* in *w*, and others are not. Let us call the set of worlds which are compatible with what *x* believes at *t* in *w* the "belief set" of *x* (at *t* in *w*) and notate it $\mathbf{B}_{x,t,w}$. As far as *x* is concerned at *t* in *w*, all the worlds in $\mathbf{B}_{x,t,w}$ are candidates for the actual world, and all the other worlds are not.

Now we can say that *x* believes Φ from a perspective *p* at *t* in *w* iff for all $w' \in \mathbf{B}_{x,t,w}$, $\Phi(w', p)$ = **truth**. (If *x* believes Φ autocentrically at *t* in *w*, then for all $w' \in \mathbf{B}_{x,t,w}$, $\Phi(w', \langle x, t, w \rangle)$ = **truth**.[120]) To say that *x* believes Φ (at *t* in *w*) is simply to say that there is some perspective from which *x* believes Φ (at *t* in *w*); this yields the following lexical definition:[121]

(190) If USE(α, $believe_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda q : q \in \mathbf{D}_{\langle s,t \rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . \exists p' \forall w' [w' \in \mathbf{B}_{x,t,w} \rightarrow q(w', p') = \textbf{truth}]]$

Now we need to identify the conditions under which an agent's beliefs are coherent. The idea is to sort the various contents which the agent believes into separate "bodies" of belief according to the perspective from which he or she believes them; the agent's beliefs are coherent iff none of these bodies is internally inconsistent.

For any perspective *p*, let *x's p-body* (at *t* in *w*) be the set $\{\Phi \mid \forall w' \in \mathbf{B}_{x,t,w} \Phi(w', p) = \textbf{truth}\}$ — in other words, the set of sentence-contents which *x* believes from *p* at *t* in *w*. We say that *x*'s *p*-body (at *t* in *w*) is *coherent* iff there is at least one *w'* such that for all Φ in *x*'s *p*-body (at *t* in *w*), $\Phi(w', p)$ = **truth**. We say that *x* believes coherently (at *t* in *w*) iff for every perspective *p*, *x*'s *p*-body (at *t* in *w*) is coherent.

Note that under this definition, if we limit attention to contents which are objectively true or false, coherence requires global consistency. To say that Φ has its truth value *objectively* just means that for all worlds *w* and all perspectives *p*, *p'*: $\Phi(w, p) = \Phi(w, p')$. But if this equality holds

[120] This is merely an *if*, and not an *iff*, because there is a *de se* aspect to autocentric assessment which this condition fails to capture. We will address the *de se* nature of autocentric assessment in Section 7.6. below.

[121] In common with other analyses broadly in the tradition of Hintikka (1969), this analysis has the effect that if an agent believes one sentence-content, he or she must also believe all contents which it entails. I believe this is correct, and that intuitions to the contrary can be explained as the result of contextual variation in meaning. Readers who are skeptical on this point are invited to adapt some other account of mental attitudes to our current framework in place of Hintikka's account; my concern in this section is not to defend a Hintikka-style analysis, but that coherence of belief does not require global consistency of the contents believed, but only consistency within each body of belief keyed to a particular perspective.

The analysis also fails to account for "sequence of tense" phenomena — for example, the absence of a reading of *John believed that Mary slept* which places the time when Mary slept (according to John's belief) later than the belief itself, but before the time of utterance. An account of such matters would take us too far afield to pursue here.

for all worlds w, it certainly holds for all worlds w in $\mathbf{B}_{x,t,w'}$ (for any x, t, w'). In other words, if a content has its truth value objectively, then (at a given time and world) one believes that content from all perspectives or none. Now suppose that Φ and Ψ contradict each other — that is, that there is no pair $\langle w, p\rangle$ such that $\Phi(w, p) = \mathbf{truth}$ and $\Psi(w, p) = \mathbf{truth}$. If an agent x believes Φ from perspective p (at t in w') and believes Ψ from perspective p' (at t in w'), it will automatically be the case that x believes Φ from p' and Ψ from p (at t in w') — so x's p-body and p'-body will both be incoherent, and we may conclude that x does not believe coherently (at t in w').

But if we consider contents which are true or false only subjectively, global consistency is *not* required for coherence. To say that Φ has its truth value only subjectively just means that there are perspectives p and p' and a world w such that $\Phi(w, p) \neq \Phi(w, p')$. This allows for the case where $\forall w \in \mathbf{B}_{x,t,w'}\, \Phi(w, p) = \mathbf{truth}$, but $\forall w \in \mathbf{B}_{x,t,w'}\, \Phi(w, p') = \mathbf{falsehood}$ — that is, where x believes Φ from perspective p but not from perspective p' (at t in w'). So if Φ and Ψ have their truth values only subjectively, it may happen that x believes Φ from p (at t in w') and believes Ψ from p' (at t in w'), without it being the case that x believes Ψ from p (at t in w') or that x believes Φ from p' (at t in w'). In this case x believes Φ and also believes Ψ (at t in w'); but even if Φ and Ψ contradict each other, x's p-body and p'-body may both remain coherent, and x may believe coherently (at t in w').

In summary, because belief is based on truth-assessment, and assessment is performed from a perspective, an agent's beliefs are compartmentalized into separate "bodies," according to the perspective from which they are believed. People have the imaginative capacity to maintain several such bodies at once. Coherence of belief is defined body-by-body; contradictions between contents belonging to different bodies is not regarded as incoherence, providing those contents involved are true or false only subjectively (as in matters of taste). But contents which are true or false objectively do not respect the compartmentalization; if two such contents contradict each other, believing them both is incoherent.

7.2. Factives, relativism and speaker commitment

Factive attitude predicates provide an interesting test case for distinguishing among different approaches to the semantics of personal taste sentences.[122] A factive[123] predicate is one which takes a clausal complement, and whose felicitous use presupposes that this complement is true. Standard examples include *regret*, *recognize*, *forget* and many others.[124] A speaker who asserted (191), for example, would normally be understood to presuppose that John's theory had flaws.

(191) John recognizes that his theory has flaws.

That this information is presupposed is easily seen by its preservation under negation:

[122] Here, I repeat an argument and incorporate text from Lasersohn (2009b).

[123] The term *factive* is due to Kiparsky and Kiparsky (1970).

[124] Factive predicates are also sometimes defined syntactically, in terms of their ability to combine with the noun *fact* and with gerundive complements, their inability to combine with infinitival clauses, etc. As Kiparsky and Kiparsky (1970) point out, a few verbs, including *know* and *realize*, are semantically factive, but do not conform to the normal factive syntactic pattern. For our purposes here, it is the semantic and pragmatic constraints on factive predicates which will be important, not their syntax.

(192) John doesn't recognize that his theory has flaws.

It should be acknowledged that under the right discourse conditions, these predicates may be used *without* commitment on the part of the speaker to the truth of their complements. For example, a speaker may assert (193) without committing to the claim that John stole the money:

(193) John imagines that he stole the money, and regrets that he did it.

I assume that a theory of presupposition projection will account for the pattern of when factive predicates do and when they don't carry this presupposition. The details of such a theory need not concern us here; see Beaver (1997) for an overview of different approaches. The important point is that the existence of examples like (193) should not keep us from recognizing the fact that ordinarily, factive predicates do commit the speaker to the truth of their complement clauses.

Many factive predicates, though not all, also semantically entail that their *subjects* are committed to the truth of their complement clauses. For example, (191) entails that John believes that his theory has flaws; likewise, if John regrets that he stole the money, then John must believe that he stole the money.

We have a variety of options for making these ideas more precise. The presupposition which commits the speaker to the truth of the complement clause could be treated as a truth-conditional entailment, for example, or as a pragmatic implicature, or in any number of other ways. The exact status of this commitment is of relatively little importance for our present purposes; the crucial point is just that the commitment exists, not its precise analysis. For the sake of concreteness, however, it may be worth spelling out a possible line of analysis in a little more detail.[125] Following our analysis of *believe* from Section 7.1. we assume that for any individual x, time t, and world w, there is a set of all worlds compatible with what x recognizes at t in w. We represent this set as $\mathbf{R}_{x,t,w}$. To assure the factivity of *recognize*, we must stipulate that (for all x, t, w) $w \in \mathbf{R}_{x,t,w}$. (If Φ is false, you cannot recognize that Φ.) To capture the fact that the subject of *recognize* is committed to the truth of the complement clause, we also stipulate that $\mathbf{B}_{x,t,w} \subseteq \mathbf{R}_{x,t,w}$. (If you recognize that Φ, you must also believe that Φ.) Now we assign $recognize_{inf}$, like $believe_{inf}$, to syntactic category (ACC\(V∩INF))/C,[126] and define it as follows:

(194) If USE(α $recognize_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda q : q \in \mathbf{D}_{\langle s,t \rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . \exists p' \forall w' [w' \in \mathbf{R}_{x,t,w} \rightarrow q(w', p') = \mathbf{truth}]]$

Finally, to capture the speaker commitment carried by *recognize*, we posit a general felicity condition on assertion, to the effect that a speaker may only felicitously assert what he or she believes:

(195) For any sentence use φ: φ is a felicitous assertion in u only if $speaker_{u,\varphi}$ believes $\langle\!\langle \varphi \rangle\!\rangle^u$ at $time_{u,\varphi}$ in u — that is, only if $\exists p' \forall w' [w' \in \mathbf{B}_{speaker_u, time_u, world_u} \rightarrow \langle\!\langle \varphi \rangle\!\rangle^u(w', p') = \mathbf{truth}]$

[125] The brief sketch of an analysis which follows will leave some important features of factivity unexplained — for example, preservation under negation — but a detailed consideration of these issues would take us too far afield.

[126] I forego the obvious listings of the other morphological variants of the verb: $recognize_{1sg_i}$, $recognize_{2sg_i}$, etc.

Now it can be easily seen that if a use of a sentence like *John recognizes that Mary slept* is true and felicitous in *u*, both John and the speaker must believe in *u* that Mary slept. John's believing it follows directly from the stipulation that $\mathbf{B}_{x,t,w} \subseteq \mathbf{R}_{x,t,w}$; the speaker's believing it follows from the requirement in (195) that felicitous assertion requires belief, combined with the lexical meaning for *recognize* given in (194), which guarantees that a use of a sentence of the form ⌜**α** *recognizes* **φ**⌝ can only be true if its constituent use of **φ** is true.

The fact that some factive predicates commit both the subject and the speaker to believing the content of the complement clause makes them an interesting test case for the semantic analysis of predicates of personal taste. Belief involves truth assessment, and if truth assessment proceeds as suggested in Chapter 6 above, both the subject and the speaker must adopt a stance in assessing the complement clause. Nothing prevents them from adopting stances with different judges—in fact, since an autocentric stance is the norm, we should expect that most typically, the subject and the speaker would each assess the subordinate clause as true using himself or herself as the relevant judge. Our expectation, then is that (196) will have a reading which suggests that *Licorice is tasty* as an expression both of John's taste and of the speaker's—and in fact, that this will be the default interpretation:

(196) John recognizes that licorice is tasty.

This expectation is borne out; a speaker who asserted this sentence would ordinarily be understood as liking the taste of licorice, and as claiming that John liked it too.

In contrast, in an analysis where *tasty* has a hidden argument, whose denotation is fixed in the same way as an ordinary indexical pronoun, this reading is predicted to be unavailable. In such an analysis, the content of a clause containing *tasty* will vary with the denotation of the indexical; *Licorice is tasty* expresses a different content when the hidden indexical is fixed to John than it does when the hidden indexical is fixed to Mary. In a sentence like (196), where this clause appears in an intensional context so that it denotes what would otherwise be its content, the clause must have different denotations depending on whether the value of the hidden argument is fixed to John or to the speaker. Since the clause only appears once in the sentence, it can contribute just one denotation to the semantic composition, and we predict that (196) can have a meaning which commits both John and the speaker to the belief that licorice is tasty for the speaker, or one which commits them both to the claim that licorice is tasty for John (but not necessarily for the speaker) — but not to a "mixed" reading which commits the speaker to the belief that licorice is tasty for the speaker, and commits John to the belief that it is tasty for John. This prediction is wrong, so we may regard factive predicates as providing some evidence in favor of a relativist analysis of predicates of personal taste over an analysis which appeals instead to hidden indexicals in the way suggested.

7.3. Infinitival clauses and time reference

We now take up attitude reports in which the subordinate clause is infinitival, such as *John believes Mary to sleep*. These show some interesting differences from examples where the subordinate clause is finite (such as *John believes that Mary sleeps*), especially in time reference. We limit attention in this section to examples in which the infinitival phrase is a whole clause with an overt subject, as opposed to examples like *John wants to sleep,* in which the subject of the

infinitival verb is not overtly expressed. Such "control" examples will be the topic of Section 7.4.

Most attitude verbs, when they combine with an infinitival complement, require the preverbal marker *to*.[127] *Believe* is typical in this way:

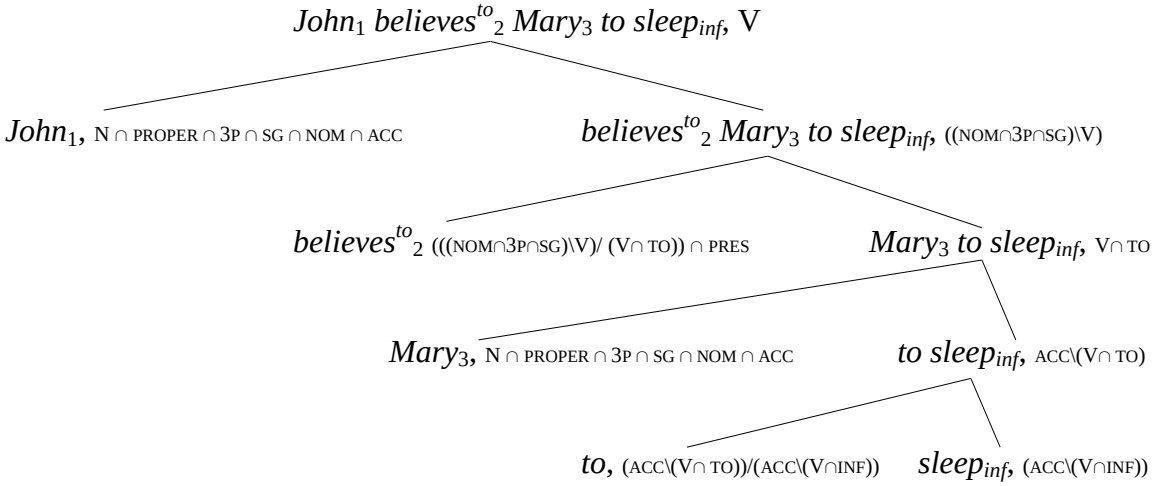
(197) a. John believes Mary to sleep.
b. *John believes Mary sleep.

Syntactically, let us assign *to* to category (ACC\(V∩ TO))/(ACC\(V∩INF)). Alongside the various forms of our existing verb *believe* (which takes finite *that*-clause complements), we introduce a variant *believe*^{to} which takes infinitival complements:

(198) a. *believe*$^{to}_{inf}$ ∈ (ACC\(V∩INF))/ (V∩ TO)
b. *believe*$^{to}_{1sgi}$ ∈ (((NOM∩1P∩SG)\V)/ (V∩ TO)) ∩ PRES
c. *believe*$^{to}_{2sgi}$ ∈ (((NOM∩2P∩SG)\V)/ (V∩ TO)) ∩ PRES
d. *believes*$^{to}_{i}$ ∈ (((NOM∩3P∩SG)\V)/ (V∩ TO)) ∩ PRES
e. *believe*$^{to}_{pli}$ ∈ (((NOM∩PL)\V)/ (V∩ TO)) ∩ PRES
f. *believed*$^{to}_{i}$ ∈ ((NOM\V)/ (V∩ TO)) ∩ PAST

We can now derive sentences such as *John believes Mary to sleep* as shown in the following tree:

(199)



Turning now to the semantics, a crucial issue is whether infinitival clauses have truth values. Most analyses seem to assume they do, and perhaps this is to be expected, given their grammatical status as clauses and their semantic role as expressing the objects of attitudes. However, truth predicates do not apply naturally at all to infinitival clauses:[128]

[127] Exceptions include perceptual verbs such as *see* and *hear*, and certain verbs of assistance and permission such as *help* and *let*. We will not deal with such verbs here.
[128] The complementizer *for* appears at the beginning of the infinitival clause in (200) because the clause is in subject position; *for* is also used with infinitival clauses serving as complements to adjectives and certain verbs. The usual

(200) *For Mary to sleep is true.

I will therefore *not* analyze infinitival clauses (or their contents) as the sorts of things that can be true or false, though naturally we will want to relate their contents to the contents of corresponding finite clauses in a systematic way.

Before considering the semantics of *to*, it should be noted that our semantic rules already assign denotations to "bare" infinitival clauses without *to*, treating them as predicates of times — that is, of type ⟨i, t⟩ in extensional contexts, or of ⟨s, ⟨i, t⟩⟩ in intensional contexts (which would be the typical contexts in which such phrases appear). For example, a use φ of [$Mary_1$ $sleep_{inf}$] consisting of α (a use of $Mary_1$) and β (a use of $sleep_{inf}$), where φ, α, β are all intensional, will receive a denotation as follows (relative to arbitrary worlds *u*, *w* and perspective *p*):

(201) $[\![\alpha]\!]^{u,w,p}$ = [λ⟨*w′*, *p′*⟩ . *Lex*(α, *u*, *w′*, *p′*)], by (128).
= [λ⟨*w′*, *p′*⟩ . $Ref_{u,\alpha}(1)$], provided that $Ref_{u,\alpha}(1)$ is named *Mary* in *u*, by (114).

$[\![\beta]\!]^{u,w,p}$ = [λ⟨*w′*, *p′*⟩ . *Lex*(β, *u*, *w′*, *p′*)], by (128).
= [λ⟨*w′*, *p′*⟩ . λ*y* : *y* ∈ $\mathbf{D}_e$. λ*t* ∈ $\mathbf{D}_i$. *y* sleeps at *t* in *w′*], by (122).

$[\![\varphi]\!]^{u,w,p}$ = APPLY($[\![\beta]\!]^{u,w,p}$, $[\![\alpha]\!]^{u,w,p}$), by (131).
= [λ⟨*w′*, *p′*⟩ . APPLY($[\![\beta]\!]^{u,w,p}$ (*w′*, *p′*), ($[\![\alpha]\!]^{u,w,p}$(*w′*, *p′*))], by (130), Case D(i).
= [λ⟨*w′*, *p′*⟩ . λ*t* ∈ $\mathbf{D}_i$. $Ref_{u,\alpha}(1)$ sleeps at *t* in *w′*], provided that $Ref_{u,\alpha}(1)$ is named *Mary* in *u*, by (130), Case A.

Put more informally: in an intensional context, the bare infinitival clause *Mary sleep* denotes the function which maps each pair of a world and a perspective to set of times when Mary sleeps in that world.

Now we must consider the semantics of *to*. As has long been observed, infinitives with *to* normally serve as a complement to an attitude verb only if the attitude is "non-backward-looking."[129] *Regret*, for example, can take a finite or gerundive clause as its complement, but not an infinitive, since one regrets what has already been the case; but *expect* can take an infinitival complement, since one expects what is yet to come:[130]

(202) a. John regretted that Mary stole the money.
b. John regretted Mary's stealing the money.
c. *John regretted Mary to steal the money.

---

analysis is that *for* is needed in such examples to assign case to the subject of the infinitive, but is omitted in infinitival complements to verbs which themselves can assign case to the subject of infinitive. Our current grammar oversimplifies pattern by letting the infinitival verb itself select for accusative case.

[129] See, e.g., Curme (1931) p. 467. Wurmbrand (2014) provides some more recent discussion.

[130] Some verbs, such as *remember* and *forget* take an infinitival complement only if it describes an action which is supposed to occur at or after the time of the attitude: *John remembered to steal the money* means that he stole the money after he remembered to do so; to say that he had a memory of a previous action, one must use a finite or gerundive clause: *John remembered that he stole the money* or *John remembered stealing the money*.

d. John expected that Mary would steal the money.
e. John expected Mary's stealing the money.
f. John expected Mary to steal the money.

To get this effect, let us define *to* so that an infinitival clause with *to* can take a time *t* as it argument, returning a partial function from times to truth values, defined only for times at least as late as *t*, but otherwise returning the same values as the corresponding bare infinitival clause:

(203) If USE(α, *to*) then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle e,\langle i,t\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t : t \in \mathbf{D}_i . \lambda t' : t' \in \mathbf{D}_i \,\&\, t \leq t' . f(x)(t') = \mathbf{truth}]]$

It may easily be confirmed from this that an extensional use of $Mary_1$ *to* $sleep_{inf}$ will denote a function which maps any given time *t* onto a partial function which is undefined for times earlier than *t*, but which is otherwise identical to the function denoted by a corresponding use of the bare infinitival clause $Mary_1$ $sleep_{inf}$. (Intensional uses, of course, will denote a function mapping world-perspective pairs to such functions.)

The utility of treating *to* in this way is best understood by considering a "forward-looking" attitude predicate such as *expect*. Let $\mathbf{E}_{x,t,w}$ be the set of worlds compatible with what *x* expects at time *t* in world *w*. We let $expect^{to}{}_{inf}$ be of syntactic category (ACC\(V∩INF))/ (V∩ TO), with morphological variants $expects^{to}$, etc. assigned to categories in the same way as the various forms of $believe^{to}$ in (198). We define $expect^{to}{}_{inf}$ as follows:

(204) If USE(α, $expect^{to}{}_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle i,\langle i,t\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_e \exists p' \forall w' [w' \in \mathbf{E}_{x,t,w} \rightarrow \exists t' [t < t' \,\&\, f(w', p')(t)(t') = \mathbf{truth}]]]$

Readers may confirm that under this definition, a use φ of $John_1$ $expects^{to}$ $Mary_2$ *to* $sleep_{inf}$ will be true in *u* iff $\forall w' [w' \in \mathbf{E}_{Ref_{u,\varphi}(1),time_{u,\varphi},u} \rightarrow \exists t' [time_{u,\varphi} < t' \,\&\, Ref_{u,\varphi}(2)$ sleeps at t' in $w']]$, provided $Ref_{u,\varphi}(1)$ is named *John* in *u* and $Ref_{u,\varphi}(2)$ is named *Mary* in *u*. That is, the sentence is true iff in all the worlds compatible with what John expects at the time of utterance, Mary sleeps at a later time.

But if there were a hypothetical verb *schmexpect*, defined like *expect* but with the ordering of *t* and *t'* reversed as in (205), it could not sensibly combine with an infinitival clause with *to*, because *to* would impose the opposite restriction on their ordering.

(205) If USE(α, $schmexpect^{to}{}_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle i,\langle i,t\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_e \exists p' \forall w' [w' \in \mathbf{E}_{x,t,w} \rightarrow \exists t' [\boldsymbol{t' < t} \,\&\, f(w', p')(t)(t') = \mathbf{truth}]]]$

Returning now to $believe^{to}$, we define it as follows:

(206) If USE(α, $believe^{to}{}_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle i,\langle i,t\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_e \exists p' \forall w' [w' \in \mathbf{B}_{x,t,w} \rightarrow f(w', p')(t)(t) = \mathbf{truth}]]$

Under this definition, *believe*^{to} is neither forward-looking nor backward-looking, but fills both temporal arguments of the infinitival clause with the same time, namely the time of the attitude. Hence *John believes Mary to sleep* means that in the worlds compatible with John's believes at time *t*, Mary sleeps at time *t*.

7.4. Centered attitudes and control of infinitival subjects

Now we turn to infinitival phrases with no overt subjects, such as the boldface phrases in (207)a. and b.:

(207) a. John expects **to arrive**.
b. John told Mary **to jump out the window**.

Most frequently, such infinitival phrases are syntactically analyzed as whole clauses with hidden pronoun subjects (notated as 'PRO'), so that these sentences may be represented as in (208), where the bracketed portions are taken to be infinitival clauses (and not simply verb phrases):

(208) a. John expects [PRO to arrive]
b. John told Mary [PRO to jump out the window]

The semantics of PRO is a matter of some debate. It is clear that (in many constructions), PRO is related to some phrase in a superordinate clause — for example to *John* in (208)a., *Mary* in (208)b. — usually termed the *controller* of PRO. One idea which should be rejected right away is that PRO is interpreted just like an overt pronoun, anaphoric to its controller. On this view, (207)a. and b. would be equivalent to (209)a. and b. respectively, aside from semantic differences traceable to the modals *will* and *should* in place of *to*:

(209) a. John_i expects that he_i will arrive.
b. John told Mary_i that she_i should jump out the window.

But in fact (207)a. and b. are not equivalent to (209)a. and b. As has been clear at least since Morgan (1970), examples like (207)a. unambiguously describe the subject as having a *de se* attitude (to use the terminology of Lewis (1979)), while (209)a. allows a reading which is not *de se*. For example, if John hears about a certain person that is scheduled to arrive — not realizing that that person is himself — and forms his expectation on that basis, (209)a. is true but (207)a. is false. Analogously, (207)b. unambiguously describes John's speech act as *de te*, while (209)b. does not: If John is talking with Mary at the party without recognizing her, and says to her "You know, Mary really ought to jump out that window," then (209)b. is true but (207)b. is false, even though (207)b. would be true if John said instead "You know, you really ought to jump out that window."[131]

A detailed exploration of the issues surrounding *de se* and other "centered" attitudes is well beyond the scope of this study. However, some consideration of such matters is necessary, first because the notion of an autocentric stance (as developed in Chapter 6 above) requires some

[131] In the latter case John *directly* performs a statement (that Mary should jump out the window), thereby indirectly telling her to jump out the window.

account of *de se* attitudes, and second because it had been argued that PRO should be analyzed using a relativist, assessment-sensitive semantics analogous to our treatment of predicates of personal taste (Stephenson (2007b), Stephenson (2008), Stephenson (2010)).

We turn first to Stephenson's account. As already mentioned in Section 6.5. above, Stephenson suggests that the PRO subject of certain infinitival clauses — including examples like those in (208)a. — is actually PRO_J, with the semantics given in (184) and repeated here as (210):

(210) $[\![\text{PRO}_J]\!]^{c,w,t,j} = j$

In other words, PRO_J denotes the value of the individual index; hence it has as its content a function mapping each triple of a world, a time and an individual onto that individual. Unlike our analysis in Section 7.3. above, Stephenson treats infinitival clauses as having truth values; hence a clause like *PRO$_J$ to arrive* has as its content a function mapping each world-time-individual triple onto **truth** iff that individual arrives in that world at that time. Because some individuals arrive (in a given world at a given time) and others don't, this function will yield different truth values as its individual argument varies. In other words, this clause receives a relativistic content, analogous to the contents assigned to personal taste sentences: Just as *Licorice is tasty* expresses a content which is true relative to some people and false relative to others, according to their tastes, *PRO$_J$ to arrive* expresses a content which is true relative to some people and false relative to others, according to whether they arrive or not. Sentences like *John expects to arrive* describe John as bearing a mental attitude to such a content.

Next, Stephenson analyzes attitude predicates as quantifying over world-time-individual triples,[132] in much the same way as we have been analyzing them as quantifying over worlds (see (190), (194)): for example, a sentence of the form ⌜**α** *believes that* **φ**⌝ is true (relative to a world *w*, time *t*, and individual *x*) iff for all *w′*, *t′*, *x′* compatible with everything the individual denoted by α believes at *t* in *w*, φ is true relative to *w′*, *t′*, *x′*. Compatibility is taken to be a matter of self-location, essentially as in Lewis (1979); to say that *w′*, *t′*, *x′* are compatible with everything John (for example) believes at *t* in *w* is to say that for all John believes (at *t* in *w*), *w′* could be the actual world, *t′* could the current time, and he himself could be *x′*.

Adapting this idea to the verb *expect*, with the semantics for PRO_J just given, it is apparent that *John expects PRO$_J$ to arrive* will be true (at *t* in *w*) iff for all *w′*, *t′*, *x′* compatible with everything John expects (at *t* in *w*), *PRO$_J$ to arrive* is true relative to *w′*, *t′*, *x′* — that is, *x′* arrives at *t′* in *w′*. In other words, any world, time and individual which (given John's expectations) are candidates for the actual world, the current time, and who he himself might be, are such that the individual arrives in that world at that time.

This analysis accounts straightforwardly for the obligatorily *de se* interpretation of *John expects to arrive*, and its contrast with *John$_i$ expects that he$_i$ will arrive*. The latter sentence will be true (relative to *w*, *t*, *x*) iff every triple *w′*, *t′*, *x′* compatible with everything John expects (at *w*, *t*) is such that *he$_i$ will arrive* is true relative to *w′*, *t′*, *x′*. Because *he$_i$* denotes John, it must be the case that John arrives in *w′*, presumably at some time later than *t′* (because of the future auxiliary *will*). But there is no restriction on *x′*, so it might well happen that John does not self-identify as *x′* in *w′* at *t′*. Hence he could expect that he will arrive simply by expecting (*de re*) of himself that he will arrive,

[132] This is the way such predicates are analyzed in Stephenson (2007b); in Stephenson (2010) the times are omitted, apparently just as a convenient simplification.

without realizing that this expectation is self-directed. In contrast, *John expects to arrive* will only be true if John self-identifies as x' by the semantics of PRO_J, so a *de se* reading is the only available interpretation.

This analysis is easily recognized as a "centered worlds" analysis of *de se* attitudes, and as such is similar in many respects to the approach advocated in Lewis (1979), according to which attitudes are relations to predicate contents, and bearing an attitude involves an act of self-predication. But it should be noted that Stephenson's analysis — unlike Lewis' — does *not* actually claim that the phrase expressing the object of an attitude is a predicate. In both approaches, the content of such a phrase is modeled as a function which takes both a world argument and an individual argument to yield a truth value, but the two analyses differ as to the status of the individual argument. A predicate has, at the level of ordinary *extension*, an open argument place. For example, a 1-place predicate of individuals may be taken as denoting a function from individuals to truth values. The *intension* (content) of a 1-place predicate of individuals may therefore be modeled as a function which takes a possible world (and/or other indices) as its argument, yielding a function which takes an individual as its argument, to yield a truth value. In Stephenson's analysis, however, the ordinary extension of an infinitival clause (even with a PRO subject) is simply a truth value — there is no open argument place for an individual. The intension is a function that takes both a world and an individual as arguments,[133] but these are both indices, not arguments to the extension of the infinitival phrase.

Treating the individual argument as an index — particularly, as an index whose value is fixed to the judge (or any other aspect of the context of assessment), as Stephenson does — has some unusual and highly questionable effects. Of course in this book I have been advocating for an individual index whose value is fixed in the context of assessment in just this way, but only for sentences whose truth value is at least in part a matter of opinion, rather than a matter of fact. Drawing this distinction between matters of opinion and matters of fact was the primary motivation for introducing the individual index in the first place. If we allow the truth value of an infinitival clause to vary with the value of this index simply by virtue of its PRO subject, and not because the infinitival clause contains any predicates of personal taste or similar expressions, then we must either regard the truth of such clauses as likewise dependent on matters of opinion, or regard clauses whose truth value depends purely on matters of fact as nonetheless varying from person to person. Neither of these options seems to me to be defensible.

If John expects to arrive, for example, his expectation will eventually either be met or not, according to the facts; opinions play no role in deciding the issue.[134] My personal preferences, tastes and affective state are irrelevant to determining whether John's expectation has been met, hence also irrelevant to the issue of whether the object of John's expectation is true or false (in a semantic theory which assigns this object a truth value).[135] If, following John's arrival, you believe that his expectation has been met and I believe it has not, then you are right and I am wrong; there is no possibility of faultless disagreement.[136] It seems clear that an infinitival clause

---

[133] And a time, at least in Stephenson (2007b).

[134] Of course in some sense John's expectation is itself an opinion, and we may all have our own opinions as to whether John will arrive or not; but this is an entirely separate matter.

[135] Our current theory, of course, allows at least two distinct type of things to serve as objects of expectation, corresponding to the contents of finite clauses and infinitival clauses; only the former have truth values. But in Stephenson's theory, finite and infinitival clauses have the same type of content, which is truth-value-bearing.

[136] Because John's expectation concerns a future contingency (future, that is, at the time of the expectation), one might

cannot be taken as expressing a matter of opinion simply because its subject is PRO, independently of whether it contains any predicates of personal taste or other expressions specifically concerned with matters of opinion.

But if the truth value of a sentence content is determined on a factual basis and not by opinion, we cannot let it vary from person to person. If we do, we grotesquely undermine the whole notion of objective fact, and leave it utterly mysterious what distinguishes matters of fact from matters of opinion. It is better to treat the truth values of factual claims as invariant.

Let us, therefore, set aside Stephenson's analysis of the PRO subject of infinitival clauses as judge-denoting, and consider other techniques for obtaining *de se* and other centered readings. It should perhaps be emphasized at this point that our project is *semantic* — that of giving a compositional interpretation of attitude reports in which the complement of the attitude predicate is an infinitival phrase with no overt subject. This task must be distinguished from the related philosophical project of explaining how attitudes can be *de se*. Considerations of compositionality will be a much more central concern for us in constructing a semantic theory of attitude reports than they would be for someone giving a (non-semantic) philosophical theory of attitudes.

In particular, we should construct our semantic theory bearing in mind that a single infinitival phrase may appear in construction with conjoined subjects, verbs, or larger phrases:

(211) a. John and Mary (each) expect to fly.
b. John hopes and expects to fly.
c. John hopes, and Mary fears, to fly.

To fit with our general assumptions about compositionality,[137] we must assign the infinitival phrase *to fly* a single denotation in each of these examples which will combine appropriately with the conjoined structures to give readings in which John and Mary's attitudes are both *de se* (in a. and c.) and in which both the hope and the fear are *de se* attitudes (in b. and c.) This turns out to be problematic under certain assumptions about the semantics of infinitival phrases.

For example, suppose that instead of the possible-worlds approach to intensionality we have adopted here, we were to assume a Fregean theory in which expressions in intensional contexts denote modes by which their customary denotations are presented to the mind. Frege (1918) suggested that "everyone is presented to himself in a special and primitive way, in which he is presented to no one else." Such special modes of self-presentation are entirely private; only the person presented by such a mode can "grasp" it.[138] A natural hypothesis about the semantics of infinitival phrases serving as complements to attitude verbs would be that their PRO subjects

---

claim that if two people disagree about his expectation *before* his arrival, the disagreement is faultless due to a general variation in the truth values of future contingent claims according to the time of assessment, as in MacFarlane (2003). We will consider whether a relativist semantics is appropriate for future contingents in Section 9.2. In the meantime, we may note that even in such a theory, once all the relevant facts are settled, no faultless disagreement is possible; hence the object of expectation ultimately receives a truth value on a purely factual basis, and in that sense appears difference from the contents of sentences about matters of opinion.

[137] See Section 1.5. above.

[138] Frege writes: "So, when Dr. Lauben has the thought that he was wounded, he will probably be basing it on this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts specified in this way." (p. 333 of the Beaney edition of Frege (1918))

denote such "special and primitive" modes of presentation, of the subject of the attitude expressed by the verb. For example in (212) we would take PRO as denoting a special way by which John is presented to his own mind, by which John is not presented to anyone else's mind. We would count an attitude as *de se* iff it involves this kind of special self-presentation, hence capturing the obligatory *de se* interpretation of infinitival PRO:

(212) John hopes PRO to fly

Unfortunately, this approach does not extend straightforwardly to examples like (211)a. and c. In each of these examples, we would have a single occurrence of PRO, appearing in an infinitival phrase which must denote the object both of John's attitude and of Mary's attitude. If PRO denotes a mode of presentation which is private to John, we do not correctly capture the fact that these sentences ascribe a *de se* attitude to Mary; and conversely, if it denotes a mode which is private to Mary, we do not capture the fact that it ascribes a *de se* attitude to John. Of course we could get the correct reading by "unpacking" these sentences to something like (213)a. and b., and interpreting the two occurrences of PRO in each example differently:

(213) a. John expects PRO to fly and Mary expects PRO to fly.
b. John hopes PRO to fly and Mary fears PRO to fly.

But such an unpacking operation violates the pattern we observe generally in other examples — normally, if a single occurrence of a phrase appears in construction with a conjunction structure, it must be interpreted identically with respect to all the conjuncts, as already mentioned in connection with (14) on p. 15.[139]

Similar concerns weigh against an analysis like that of Higginbotham (2003). Higginbotham adopts a semantic theory in which verbs have a hidden argument place for events as in Davidson (1967). Gerundive clauses (like *John flying*) are treated as denoting (in intensional contexts) functions from possible worlds to sets of events, and attitude verbs which take gerundive complements are analyzed as denoting relations between individuals and such functions. Hence *Mary remembers John flying* is represented as in (214):[140]

(214) $\exists e$ remember(m, ^$\lambda e'$ fly(j, e'), e)

Although Higginbotham generally follows Davidson in treating the subject and object of a clause as arguments of the verb, he also makes use of thematic relations roughly in the style of "neo-Davidsonian" semantic theories as developed by Parsons (1980), Parsons (1990) and many others. We let σ be the thematic relation corresponding to the subject of an attitude verb: if e is an

[139] As Stanley (2011) correctly points out, the Lewisian strategy of analyzing *de se* attitudes in terms of self-ascription of properties does not solve the problem of explaining what it is to take a first-person perspective on the world. It may be that an elucidation of the concept of self-ascription will ultimately require something like Frege's "special and primitive" modes of presentation. As I see it, however, the primary advantage of a self-ascription-based account of *de se* attitude reports is not that it better explains what a first-person perspective is, or eliminates the theoretical appeal to modes of presentation, but simply that it allows a compositional semantics.

[140] The ^-operator is a standard Montague-style intension-forming operator: for any expression α, ^α denotes that function which maps any possible world w onto the extension of α relative to w.

event (or state) of some individual *x* bearing an attitude such as remembering something, we say that $\sigma(e) = x$. The same individual may stand in thematic relations to multiple events; where σ and θ are thematic relations, we construe '$\sigma(e)$ & $\theta(e')$' as a term presupposing that $\sigma(e) = \theta(e')$, and (on the assumption that this presupposition is met) standing for the unique *x* such that $x = \sigma(e) = \theta(e')$. When PRO appears as the subject of a gerundive or infinitival clause serving as complement to an attitude, it is interpreted as term of the just this form, where σ is the thematic relation which the controller of PRO bears to the attitude verb, and θ is the thematic relation which the gerundive or infinitival verb assigns to its subject.

For example, if σ is the thematic relation associated with the subject of *remember*, and θ is the thematic relation associated with the subject of *fly*, we may interpret *Mary remembers flying* as in (215):

(215) $\exists e$ remember(m, $\hat{}\lambda e'$ fly($\sigma(e)$ & $\theta(e')$, e'), e)

Here, the object of Mary's remembering is $\hat{}\lambda e'$ fly($\sigma(e)$ & $\theta(e')$, e') — the property of being an event in which the subject of that very remembering event flies. This is supposed to give a *de se* reading: if Mary remembers some event in which she flew, but doesn't recognize that the person doing the flying was herself, she would presumably not identify that flying person as the subject of her remembering event, so (215) would be false.

But now consider how we might analyze, in this approach, an example like (216):

(216) John anticipates, and Mary remembers, flying.

This sentence describes a pair of attitude events corresponding to the two main verbs, one with John as the subject and one with Mary. But the sentence contains just a single gerundive phrase, presumably with just a single occurrence of PRO as its subject. If we represent this subject as '$\sigma(e)$ & $\theta(e')$', we must consider which event variable '*e*' is — is it the variable which serves as argument to *anticipates*, or the one which serves as argument to *remembers*?[141] Neither answer gives the right results; we don't want John's anticipation to be directed toward the subject of Mary's remembering or vice versa. Of course we could get the intended interpretation by claiming that the gerundive phrase is interpreted twice with a different content each time; but once again this conflicts with our assumptions about compositionality, and with the pattern we observe more generally in examples where a single phrase appears in construction with a coordinate conjunction structure.[142]

[141] We must have two different variables serving as arguments to the two verbs: If we instead had two occurrences of the same variable bound by the a single quantifier, John's anticipating would have to be the same event as Mary's remembering; and if we had two occurrences of the same variable bound by different quantifiers, only one of those quantifiers could bind the occurrence of this variable which serves as part of the interpretation of PRO, which would clearly be the wrong result.

[142] Aside from this problem, it is far from clear that Higginbotham's approach adequately captures the *de se* interpretation of PRO. Notice that (215) is true only if there is some remembering event *e* toward which Mary has a *de re* attitude — the attitude of remembering the type of event in which whoever is the subject of *e* flies. But it is possible to construct examples in which the subject has such a *de re* attitude, without the corresponding English sentence being true. Here is an example involving the verb *anticipate* rather than *remember*: Suppose John has the power of telepathy, and is putting on a show with Mary as a member of the audience. He says: "By my telepathic powers, I am detecting an attitude of anticipation from a member of the audience! What that audience member is anticipating is that she will find

It appears we must analyze infinitival and gerundive phrases (without overt subjects) in such a way that their content does not vary with the denotation of the controller, or of the predicate of which the infinitival or gerundive phrase serves as complement. This is most easily and straightforwardly accomplished if we take the infinitive or gerundive phrase as a predicate, with an open argument place corresponding to the subject, as in Bach (1979), Williams (1980), Dowty (1985), Chierchia (1989) and others.

There are a variety of ways this idea can be worked out. One possibility (advocated by Dowty (1985) and others) would be to treat the infinitival phrase as a bare verb phrase, with no syntactically present subject at all. But this approach faces well-known problems accounting for syntactic patterns in which infinitival phrases without overt subjects pattern with whole clauses rather than finite verb phrases. For example as pointed out by Chomsky (1980), interrogative pronouns normally appear at the beginning of a clause, not a verb phrase; yet they can appear at the beginning of an infinitival phrase as in (217):

(217) John doesn't know what to wear.

Koster and May (1982) provide a long catalog of additional generalizations which hold both of infinitival phrases and of whole clauses, but not of finite verb phrases.

An alternative way of analyzing infinitival phrases as predicates (suggested for example by Chierchia (1989)) is to claim that such phrases are full clauses, but also to require that PRO be bound by a hidden operator scoping over the infinitival phrase, with the semantics of a λ-operator. Thus *John wants to fly* is represented essentially as *John wants* O_i[PRO_i *to fly*], where *O* is interpreted as λ-operator, so that O_i[PRO_i *to fly*] denotes the function [λx . fly(x)]. But this approach also faces difficulties. The use of such an operator is independently motivated in certain infinitival phrases with non-subject gaps; for example (218)a. is analyzed as in (218)b.:[143]

(218) a. The spaghetti is ready for John to serve.
b. The spaghetti is ready O_i[for John to serve e_i]

But if attitude verbs like *want* simply select for predicates formed via such an operator, nothing would seem to prevent infinitival phrases in which the operator binds a non-subject gap as complements to such verbs. Hence we should expect (219) to be well-formed, with the interpretation that Mary wants to be served by John. In fact the sentence lacks this reading, and means merely that Mary wants John to perform some service:[144]

(219) Mary wants O_i[for John to serve e_i]

$100 under her seat!" John's power is authentic; the attitude he detects is Mary's. Having seen the show before and knowing how it goes, Mary does in fact anticipate that the person whose anticipation John detects will find $100 — but she does not realize that person is herself. In this case, it seems to me, Mary does anticipate *de re* of that very anticipation event that its subject will find $100; but the sentence *Mary anticipates finding $100 under her seat* is not true, because her attitude is not *de se* and the PRO subject of the gerundive clause requires a *de se* attitude.

[143] I believe this analysis originates with Chomsky (1977).

[144] Essentially this argument is given in Landau (2000), though presented rather differently. On the reading in which Mary simply wants John to perform some service, the object of *serve* is interpreted existentially as in Section 5.6. above, rather than bound by *O*.

The solution to this problem, I suggest, is not to treat PRO like an ordinary variable and bind it with an operator, but instead to assign PRO a semantics which keeps an argument place open from the start, with no operator involved. In that case, the infinitival clause can function as a predicate without any operator-binding, and the syntactic structure of clauses like [PRO *to fly*] will be different enough from that of clauses like O_i[*for John to serve* e_i] that is should come as no surprise that certain predicates may select complements with one of these structures but not the other.

To work out this idea, let PRO bear a referential index (similarly to pronouns, names, and common nouns) and belong category (V∩ TO) /(ACC\(V∩ TO)). That is, it takes an infinitival verb phrase as its argument to yield an infinitival clause:[145]

(220) $PRO_i \in$ (V∩ TO) /(ACC\(V∩ TO))

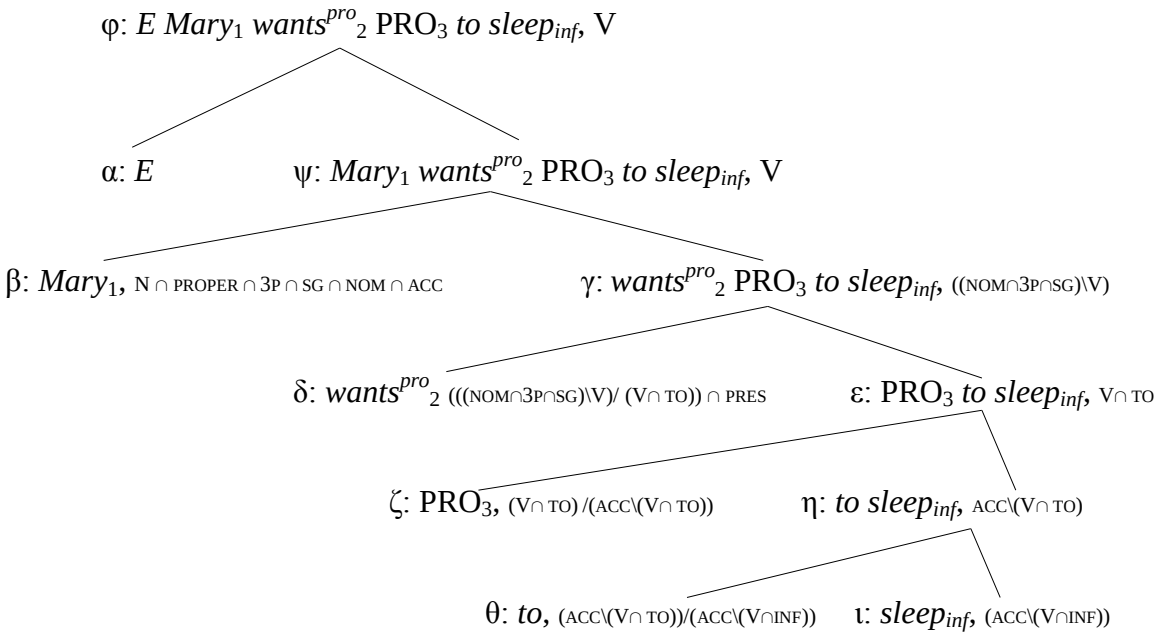
We also add various forms of *want* to our lexicon, categorized to take infinitival clause complements:

(221) a. $want^{pro}_{inf} \in$ (ACC\(V∩INF))/ (V∩ TO)
b. $want^{pro}_{1sgi} \in$ (((NOM∩1P∩SG)\V)/ (V∩ TO)) ∩ PRES
c. $want^{pro}_{2sgi} \in$ (((NOM∩2P∩SG)\V)/ (V∩ TO)) ∩ PRES
d. $wants^{pro}_{i} \in$ (((NOM∩3P∩SG)\V)/ (V∩ TO)) ∩ PRES
e. $want^{pro}_{pli} \in$ (((NOM∩PL)\V)/ (V∩ TO)) ∩ PRES
f. $wanted^{pro}_{i} \in$ ((NOM\V)/ (V∩ TO)) ∩ PAST

Our grammar will now generate sentences like *Mary wants to sleep* with the following structure:

[145] This analysis is intended only for PRO subjects of infinitival clauses serving as complements to attitude predicates. PRO subjects of infinitival clauses in other positions may require a different analysis, but consideration of such cases here would take us too far afield.

(222)



Now for any individual *x*, time *t* and world *w*, let $\mathbb{W}_{x,t,w}$ be the set of world-individual pairs compatible with what *x* wants at *t* in *w*; that is, $\langle w', x' \rangle \in \mathbb{W}_{x,t,w}$ iff being *x′* in *w′* satisfies everything *x* wants at *t* in *w*. We may define the actual verb *want* in terms of $\mathbb{W}$ as follows:

(223) If USE(α, $want^{pro}_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle e,\langle i,\langle i,t\rangle\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_e .$
$\exists p'\forall w'\forall x'[\langle w', x'\rangle \in \mathbb{W}_{x,t,w} \rightarrow \exists t'[t<t' \,\&\, f(w', p')(x')(t)(t') = \textbf{truth}]$

In defining a semantics for PRO, we should keep in mind the possibility of "partial control," as in (224):[146]

(224) The students were trying to decide where to meet. John wanted to meet in the library, but Mary wanted to meet in a café.

The meaning here is that John wanted the group to meet in the library, and Mary wanted the group to meet in a café, not that John wanted it to be the case that *he* met in the library and Mary wanted it to be the case that *she* met in a café (which would make no sense anyway, since *meet* is a collective predicate). The attitudes described are still *de se* — if John simply desires the group to meet in the library, but does not recognize that he himself is a member of the group, (224) is false. But the verb phrase *meet in the library* (as opposed to the whole infinitival clause [PRO *to meet in the library*]) does not express the property which John wants to be applicable to himself. The possibility of partial control is most obvious when the verb phrase of the infinitival clause is a collective predicate such as *meet* in (224), but is available much more generally. Even in examples like *Mary wanted to sleep*, partial control is possible, though it may take some context to make this interpretation salient:

[146] See Landau (2000) for a useful discussion of partial control and the conditions under which it is allowed.

(225) The hiking party were trying to decide whether to press on or to stop and catch some sleep. John wanted to press on, but Mary wanted to sleep.

On one prominent reading of (225), Mary is described as wanting the group to stop and sleep, not merely herself. Since *sleep* is a distributive predicate, the group cannot sleep without Mary sleeping, so her desire for the group to sleep implies a desire that she herself sleep; this reading must nonetheless be distinguished from the reading where Mary's desire is purely self-directed.

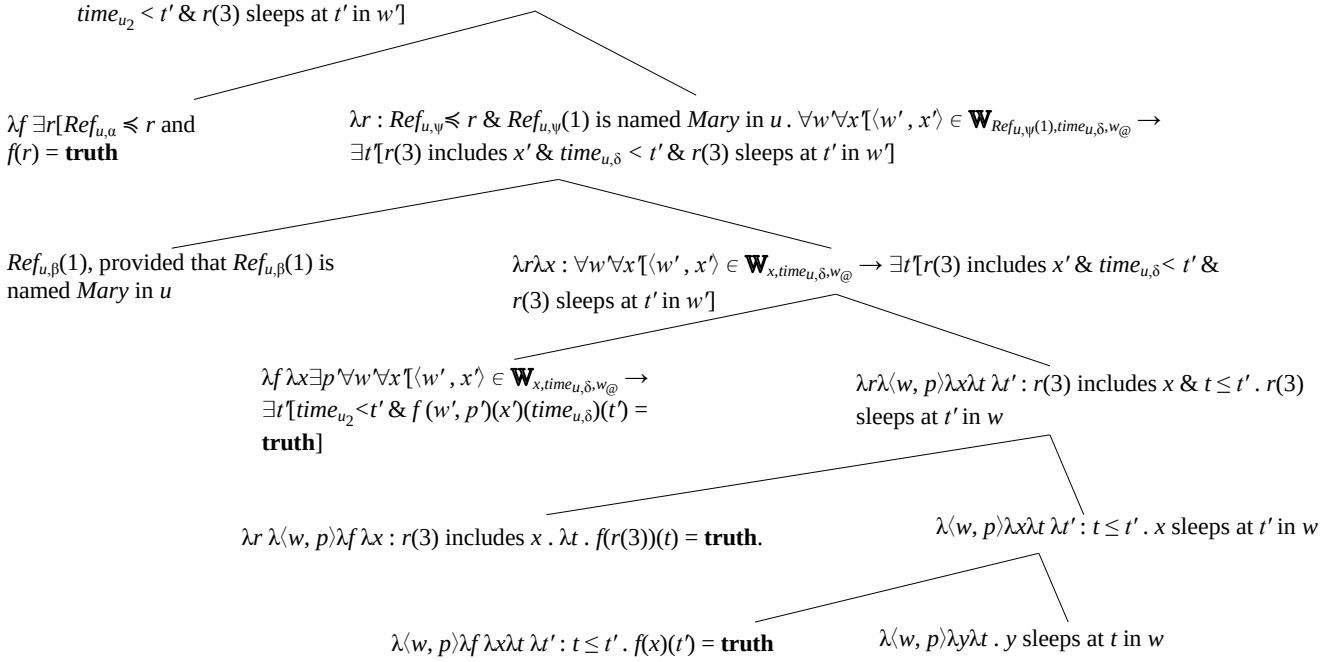
To deal with the possibility of partial control, we must distinguish the predicate expressed by an infinitival clause, which serves as the object of a *de se* attitude, from the predicate expressed by the verb phrase which combines with PRO to form that infinitival clause. The argument place kept open by PRO is *not* identical to the subject argument place for the infinitival verb phrase. For that reason, we do not let PRO_i denote the value of i (its referential index). Instead, we let the value of i serve as the argument to the constituent verb phrase, but define PRO in such a way that it keeps open an argument place for groups or individuals included in that value. Specifically (given an assignment function and world-perspective pair) let of PRO_i denote a function taking the denotation of an infinitival verb phrase π as an argument, yielding a function that maps any given x (and two times) onto **truth** iff the denotation of π maps y (and those times) onto **truth** and y includes x, where y is the value of i:[147]

(226) If $\text{USE}(\alpha, \text{PRO}_i)$ and α is intensional, then $[\![\alpha]\!]^{u,w,p} = [\lambda r : \mathit{Ref}_{u,\alpha} \preccurlyeq r \,.\, \lambda\langle w', p'\rangle \,.\, \lambda f : f \in \mathbf{D}_{\langle e,\langle i,t\rangle\rangle} \,.\, \lambda x : x \in \mathbf{D}_e \,\&\, r(i) \text{ includes } x \,.\, \lambda t : t \in \mathbf{D}_i \,.\, f(r(i))(t) = \textbf{truth}]$. (If α is extensional, $[\![\alpha]\!]^{u,w,p}$ is undefined.)

With these definitions in place, example (222) may now be interpreted as in (227):

[147] We construe "includes" here so that any group includes its members, and any individual includes itself, so that *Mary wants to sleep* may mean either that Mary wants (*de se*) that she herself sleep, or that Mary wants (*de se*) that a group of which she is a member sleep.

(227) $\exists r$ [$Ref_{u,\varphi} \preccurlyeq r$ & $Ref_{u,\varphi}(1)$ is named *Mary* in u . $\forall w' \forall x' [\langle w', x' \rangle \in \mathbf{W}_{Ref_{u,\varphi}(1),time_{u,\delta},w_@} \rightarrow \exists t' [r(3)$ includes x' & $time_{u_2} < t'$ & $r(3)$ sleeps at t' in $w']$



In effect, this says that there is some group or individual $r(3)$, and every world w' and individual x' such that being x' in w' satisfies everything Mary wants are such that which x' is included in $r(3)$ and $r(3)$ sleeps in w'. In a context whose reference assignment already gives a value for index 3, the existential quantification is superfluous, and we obtain a reading involving indexical reference to a group or individual meeting this condition.

Before leaving this topic behind, we should give an account of how the interpretation of attitude predicates taking infinitival complements with PRO relates semantically to the interpretation of such predicates when they take infinitival complements without PRO. Because we are treating infinitival clauses with and without PRO as belonging to different logical types — with an open argument place in the former case which is filled in the latter — we cannot use our existing lexical entry for *want* in (223) and our existing compositional rules to derive an interpretation for (228)a., for example; nor can we use our lexical entry for *expect* in (204) and our existing rules to derive an interpretation for (228)b.:

(228) a. Mary wants John to sleep.
b. Mary expects to sleep.

To allow for like examples like (228)a., we add various additional forms of *want* to our lexicon:[148]

[148] Note that we cannot adopt a general principle that any predicate which takes infinitival complements with PRO can also take infinitival complements without PRO; *try* is an obvious exception. Nor do predicates which take infinitival complements without PRO all take complements with PRO; *believe* is an exception. (Note that the Italian counterpart to *believe*, the verb *credere*, does take infinitival complements with PRO; see Chierchia (1989).) I see no alternative but to stipulate lexically which predicates combine with complements with PRO and which do not.

(229) a. $want^{to}_{inf} \in (\text{ACC}\backslash(\text{V}\cap\text{INF}))/ (\text{V}\cap \text{TO})$
b. $want^{to}_{1sgi} \in (((\text{NOM}\cap 1\text{P}\cap\text{SG})\backslash\text{V})/ (\text{V}\cap \text{TO})) \cap \text{PRES}$
c. $want^{to}_{2sgi} \in (((\text{NOM}\cap 2\text{P}\cap\text{SG})\backslash\text{V})/ (\text{V}\cap \text{TO})) \cap \text{PRES}$
d. $wants^{to}_{i} \in (((\text{NOM}\cap 3\text{P}\cap\text{SG})\backslash\text{V})/ (\text{V}\cap \text{TO})) \cap \text{PRES}$
e. $want^{to}_{pli} \in (((\text{NOM}\cap\text{PL})\backslash\text{V})/ (\text{V}\cap \text{TO})) \cap \text{PRES}$
f. $wanted^{to}_{i} \in ((\text{NOM}\backslash\text{V})/ (\text{V}\cap \text{TO})) \cap \text{PAST}$

Now, recalling that $\langle w', x'\rangle \in \mathbb{W}_{x,t,w}$ iff being x' in w' satisfies everything x wants at t in w, let us define a set $\mathbf{W}_{x,t,w}$ (of worlds, rather than world-individual pairs) in terms of $\mathbb{W}_{x,t,w} : w' \in \mathbf{W}_{x,t,w}$ iff for all $y \in Relevant_u$, $\langle w', y\rangle \in \mathbb{W}_{x,t,w}$. In other words, if no matter who y is (from among the relevant individuals), being y in w' satisfies everything x wants at t in w, we may say that w' satisfies everything x wants at t in w. Now we can define $want^{to}_{i}$ in terms of $\mathbf{W}$, on the model of our definition of $expect^{to}_{inf}$ in (204):

(230) If $\text{USE}(\alpha, want^{to}_{inf})$ then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle i,\langle i,t\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i \exists p'\forall w'[w' \in \mathbf{W}_{x,t,w} \rightarrow \exists t'[t<t' \,\&\, f(w', p')(t)(t') = \textbf{truth}]]]$

This gives the effect that *Mary wants John to sleep* is true iff in all the worlds satisfying everything Mary wants, John sleeps (at a later time).

In the same way, we can let $\mathbb{E}_{x,t,w}$ be the set of world-individual pairs such that $\langle w', x'\rangle \in \mathbb{E}_{x,t,w}$ iff being x' in w' is compatible with everything x expects at t in w. The set $\mathbf{E}_{x,t,w}$ used in (204) may then be defined in terms of $\mathbb{E}_{x,t,w}$ in the same way we defined $\mathbf{W}_{x,t,w}$ in terms of $\mathbb{W}_{x,t,w} : w' \in \mathbf{E}_{x,t,w}$ iff for all $y \in Relevant_u$, $\langle w', y\rangle \in \mathbb{E}_{x,t,w}$. This allows us to add PRO-controlling forms of *expect* to our lexicon and define them as in (232):

(231) a. $expect^{pro}_{inf} \in (\text{ACC}\backslash(\text{V}\cap\text{INF}))/(\text{V}\cap \text{TO})$
g. $expect^{pro}_{1sgi} \in (((\text{NOM}\cap 1\text{P}\cap\text{SG})\backslash\text{V})/(\text{V}\cap \text{TO})) \cap \text{PRES}$
h. $expect^{pro}_{2sgi} \in (((\text{NOM}\cap 2\text{P}\cap\text{SG})\backslash\text{V})/(\text{V}\cap \text{TO})) \cap \text{PRES}$
i. $expects^{pro}_{i} \in (((\text{NOM}\cap 3\text{P}\cap\text{SG})\backslash\text{V})/(\text{V}\cap \text{TO})) \cap \text{PRES}$
j. $expect^{pro}_{pli} \in (((\text{NOM}\cap\text{PL})\backslash\text{V})/(\text{V}\cap \text{TO})) \cap \text{PRES}$
k. $expected^{pro}_{i} \in ((\text{NOM}\backslash\text{V})/(\text{V}\cap \text{TO})) \cap \text{PAST}$

(232) If $\text{USE}(\alpha, expect^{pro}_{inf})$ then $Lex(\alpha, u, w, p) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle e,\langle i,\langle i,t\rangle\rangle\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . \exists p'\forall w'\forall x'[\langle w', x'\rangle \in \mathbb{E}_{x,t,w} \rightarrow \exists t'[t<t' \,\&\, f(w', p')(x')(t)(t') = \textbf{truth}]]]$

More generally, for any set $\mathbb{A}$ of world-individual pairs and expression use α, we may define a corresponding set $\mathbf{A}$ of worlds such that $w \in \mathbf{A}$ iff for all $y \in Relevant_{u,\alpha}$, $\langle w, y\rangle \in \mathbb{A}$. The PRO-controlling and non-controlling predicates corresponding to these sets may then be defined on the model just illustrated.

7.5. Effective centering without PRO

The analysis just given derives *de se* interpretations for attitude reports in which the object of the

attitude is expressed by an infinitival clause with a PRO subject. However, it seems clear that in some cases, sentences may be used to express or report *de se* or other centered attitudes, even if they do not contain such a clause. For example, consider the case of the amnesiac Rudolf Lingens lost in the Stanford Library, from Perry (1977).[149] In the library, he reads all about Rudolf Lingens, and acquires a detailed knowledge of his life, including information about Lingens' current whereabouts. Yet it seems that no matter how much such information he finds in the library volumes, there is still an extra step to the realization he might express to himself by saying "*I* am Rudolf Lingens!" Here there is no infinitival clause and no PRO; yet the sentence seems to express a centered content. In this scenario, Lingens comes to have a *de se* attitude.

A related class of examples involve *de nunc* attitudes. As Prior (1959) points out, a use of the sentence *Thank goodness that's over* on June 15, 1954 is not accurately paraphrased as "Thank goodness that ended before June 15, 1954." Rather, the sentence seems to express relief that the event referred to[150] is over "now" — the time the speaker is experiencing, regardless of what time that might turn out to be. Prior argues on the basis of this sort of example that sentences like *That's over* must have contents whose truth values vary over time, rather than expressing contents which are true or false eternally — precisely the kind of analysis which I argued against in Section 3.4. above.

I will argue in this section[151] that in all such cases, the linguistic expression in question does not actually have a centered content — that is, the sort of content toward which one might bear an authentically *de se* or similar attitude. Rather, the centered "flavor" which such examples seem to have is due to a pragmatic assumption that the relevant attitude-holder bears a *de se* attitude toward some other content, which is not the content of the linguistic expression in question, but which relates to it in a systematic way.

This is not to say that in all such cases, linguistic expressions are interpreted in the normal way. On the contrary, various special interpretations may be involved. For example, in many such cases the content of the relevant expression may be derived via *diagonalization*, roughly as developed in Stalnaker (1999). The basic idea of diagonalization was briefly sketched above in Section 2.5. I will suggest that this technique does not, by itself, provide an adequate account of centered attitudes. However, in combination with the analysis of attitude reports using infinitival clauses with PRO just presented in Section 7.4. an appeal to diagonalization can explain how sentences with no infinitives and no PRO can nonetheless be interpreted in an "effectively" centered way.

Diagonalization is perhaps most easily understood by considering first its application to puzzles about the informativity of identity statements involving demonstrative pronouns. Proper names and demonstrative pronouns both ordinarily denote rigidly, so that *Chris Barker* denotes Chris Barker relative to every world, and *this*, used while pointing at Chris Barker, also denotes Chris Barker relative to every world. But this creates a theoretical problem: If I point at Chris Barker and say "This is Chris Barker," the content I express is true relative to a world *w* iff the

[149] The scenario in which Lingens is lost in the library is from Perry's article, but the character of Rudolf Lingens is originally from Frege (1918).

[150] Many discussions of this example assume a dental root canal procedure as the relevant event.

[151] The discussion in this and the next section includes arguments and (with some modification) analyses from Lasersohn (2013).

referent of *this* relative to *w* is the referent of *Chris Barker* relative to *w*; but if both terms rigidly denote Chris Barker, the content is true in all worlds, and my utterance should be completely uninformative. But in fact a sentence of this kind can be used to convey new and useful information.

Diagonalization offers a solution to this problem. If (prior to any decision whether to believe what I say) you are unsure whether the person I am pointing at is Chris Barker or Chris Kennedy, this amounts to uncertainty on your part whether we are in a world w_1 where my use of *this* (interpreted in the ordinary way) rigidly denotes Chris Barker, or in a world w_2 where this same use (interpreted in the ordinary way) rigidly denotes Chris Kennedy. If you accept my assertion, this means you accept that we are in a world like w_1, and not in w_2 or any similar world. That is, you accept a proposition which is true in worlds like w_1 but false in worlds like w_2 — and this is distinct from both from the proposition my use of the sentence, interpreted in the ordinary way, expresses in w_1 (which is true in all worlds) and from the proposition which this same use, again interpreted in the ordinary way, expresses in w_2 (which is false in all worlds). The information you accept — and the information I convey — is the "diagonal" proposition for this use of the sentence: the proposition which is true in a world *w* iff ordinary, "horizontal" proposition expressed in *w* is true in *w*.

Let us accept, then, that in certain contexts, expressions may be interpreted as expressing special diagonal contents, rather than their ordinary horizontal contents. We need to clarify exactly how this interpretation fits with our formal grammatical theory, and what kinds of contexts trigger it.

Stalnaker's original presentation gives the impression that diagonalization takes place at the level of whole clauses; in fact, he defines a diagonalization sentence-operator. It will be considerably easier in our current system, however, to diagonalize at the level of lexical items. Our existing system assigns denotations to uses of lexical items based on the "customary" assignment by *Lex*, but in the case of intensional uses, assigns denotations which differ from those assigned by *Lex*; see (128). We continue this strategy, but revise (128) to make denotations sensitive not just to whether a use is intensional or extensional, but also to whether it is horizontal or diagonal:

(233) For any lexical item **α**, if USE(α, **α**), then:
- a. If α is extensional-horizontal in *u*, then $[\![\alpha]\!]^{u,w,p} = Lex(\alpha, u, w, p)$;
- b. If α is intensional-horizontal in *u*, then $[\![\alpha]\!]^{u,w,p} = \lambda\langle w', p'\rangle\ .\ Lex(\alpha, u, w', p')$;
- c. If α is extensional-diagonal in *u*, then $[\![\alpha]\!]^{u,w,p} = Lex(\alpha, w, w, p)$;
- d. If α is intensional-diagonal in *u*, then $[\![\alpha]\!]^{u,w,p} = \lambda\langle w', p'\rangle\ .\ Lex(\alpha, w', w', p')$.

For example, a diagonal-extensional use α of the pronoun *I* will denote, relative to any given world *w*, the speaker of α in *w*. In effect, diagonalization eliminates the distinction between the "*u*" and "*w*" world parameters, derigidifying those aspects of interpretation which are normally held rigid by the context of use.

As an example, suppose Lingens in the library has not yet fully deduced his identity, but has narrowed it down to the point where he knows he is one of three individuals: Rudolf Lingens, Gustav Lauben, or Leo Peter. That is, he does not know whether he is in world w_1, where his use of the pronoun *I* customarily denotes (relative to w_1, w_2, w_3, or any other world) Rudolf Lingens, or in world w_2 where it customarily denotes (relative to w_1, w_2, w_3, or any other world) Gustav Lauben,

or in world w_3 where it customarily denotes (relative to w_1, w_2, w_3, or any other world) Leo Peter. In this situation an extensional-diagonal use by Lingens of *I* (as it occurs in any of w_1, w_2, w_3) will denote Lingens relative to w_1, Lauben relative to w_2, and Peter relative to w_3. These denotations can enter the semantic derivation of a use of the sentence *I am Rudolf Lingens* to yield a value of **truth** relative to w_1 and **falsity** relative to w_2 and w_3. When Lingens finally realizes which of these three people he is, the information he acquires is precisely that he is in w_1 and not in w_2 or w_3 — in other words, the information that the content of such a use of the sentence is true.

This seems right as far as it goes, but saying this much will not fully explain the *de se* nature of this example. The reason is similar to that pointed out in Footnote 142 above for the analysis in Higginbotham (2003): The reading we derive via this technique is one which makes a *de re* claim about a particular use of the pronoun *I*. In effect, the analysis assigns a use φ of the sentence *I am Rudolf Lingens* a reading which can be paraphrased as "The speaker of α is Rudolf Lingens," where α is the particular use of *I* which forms a constituent of φ. Normally, this is an inconsequential point, because normally we can recognize ourselves as the speakers of the words we use. But in principle, one could be the speaker of some particular use α, and have a *de re* belief about α, without recognizing oneself as the speaker of α. For example, suppose that before discovering his true identity, Rudolf Lingens tells someone "I am Rudolf Lingens," intending to deceive his addressee. Unbeknownst to Lingens, we secretly tape his utterance. Later, we play back the recording to him. Lingens fails to recognize that the recording is of his own earlier utterance; he believes it was produced by the "real" Rudolf Lingens, whom he does not yet identify as himself. In this case there is a use α of the pronoun *I*, of which Lingens believes that its speaker is Rudolf Lingens; this belief is, moreover, *true* — Lingens is the speaker of α. Yet Lingens has still not self-identified as Rudolf Lingens; he hasn't acquired the *de se* belief we are trying to account for.

It seems clear that the content of an extensional-diagonal use of *I am Rudolf Lingens* does not, by itself, give us a *de se* reading. However, this does not mean we need to revise our semantic analysis of this sentence at all. If Lingens comes to believe this content, this *does* amount to acquiring the *de se* belief that he is Rudolf Lingens, provided he *also* believes himself (*de se*) to be the speaker of α. In order to get the correct effect, we need only find some way of representing Lingens' belief when he believes himself (*de se*) to be the speaker of α.

Fortunately, we have an analysis of *de se* attitudes in place, which we developed in order to give a semantics of predicates which take infinitival complements with PRO subjects. Just as we did with *want* and *expect*, we may assume for *believe* a corresponding set $\mathbb{B}_{x,t,w}$ (for each individual *x*, time *t* and world *w*): the set of world-individual pairs such that $\langle w', x' \rangle \in \mathbb{B}_{x,t,w}$ iff being x' in w' is compatible with everything *x* believes at *t* in *w*. To say that in world *w* at time *t* Lingens believes himself (*de se*) to be the speaker of α is just to say that for all w', x' such that $\langle w', x' \rangle \in \mathbb{B}_{\text{Lingens},t,w}$: $speaker_{w',\alpha} = x'$.[152]

[152] One way to think of this is to regard Lingens as standing in a belief relation to the content of the infinitival clause *PRO to be the speaker of α*. It is a little bit of a distortion to think of it in these terms, since the English verb *believe* does not take complement clauses with PRO subjects and therefore does not denote the kind of function which could take such a content as its argument. But as pointed out in Footnote 148, this is a quirk of English: the Italian verb *credere*, which is normally translated as *believe*, does allow such complements. It would be better to say that for Lingens to believe himself (*de se*) to be the speaker of α is for him to stand the in the *credere* relation to the content of *PRO to be the speaker of α*.

To summarize to this point: I am suggesting that the sentence *I am Rudolf Lingens* does not have an authentically *de se* reading by virtue of its conventional, internal semantics. Normally, where the speaker of a use φ of this sentence is *x*, φ expresses a content which is true in a world *w* iff *x* is Rudolf Lingens in *w* (hence, true in all worlds or none); but if φ occurs in the right pragmatic circumstances, it can express a special "diagonal" content which is true in *w* iff the speaker of φ in *w* is Rudolf Lingens (a contingent proposition). In principle, Lingens can believe this content without believing anything *de se*; but if he also believes himself *de se* to be the speaker of φ, then in effect he believes himself *de se* to be Rudolf Lingens. The *de se* belief that he is the speaker of φ can be characterized by quantifying on world-individual pairs, roughly as in Lewis (1979), but this is entirely separate from the content of the sentence *I am Rudolf Lingens*.

Under what pragmatic circumstances is a diagonal reading available? Stalnaker (1979) originally suggested that such a reading was available when the ordinary, horizontal reading of a sentence was uninformative in context, in the sense that it would not trigger any change to the "common ground" of the discourse participants. I have little to add to this idea, though I suspect that all kinds of pragmatic anomaly may trigger diagonal readings, not just underinformativeness. I will not pursue this idea here, but simply note that horizontal readings are generally preferred, and that diagonal readings generally are available only when horizontal readings are pragmatically anomalous.

The example just discussed, although it concerns the expression of a *de se* belief, does not involve any attitude predicates; we were discussing the free-standing sentence *I am Rudolf Lingens*, not a subordinate clause serving as complement to *believe* or any other predicate. Once we turn to subordinate clauses, we face some additional complications, particularly because the subject of a *de se* attitude may be referred to in the subordinate clause by the third person pronoun, to which that subject need not bear any sort of metalinguistic attitude.

For example, consider the case of mad Humeson, suffering under the delusion that he is Irene Heim.[153] In describing Humeson's delusion, John might accurately say "Humeson believes that he is Irene Heim," but this does not mean that in all the worlds compatible with what Humeson believes, John's use of the pronoun *he* refers to Irene Heim. On the contrary, if Humeson is even familiar with John's use of *he*, would more likely take it (and John's use of the name *Humeson*) to refer to someone other than Irene Heim.

Something like the technique developed in Higginbotham (2003), rejected in Section 7.4. above in the analysis of PRO, seems more appropriate here. Humeson is in a particular experiential state. His belief is that the subject of that experiential state (hence the subject of that very belief) is Irene Heim. By itself, this does not qualify Humeson's belief as *de se*,[154] but if he also knows himself (*de se*) to be the subject of that experiential state, then it seems fair to say that Humeson believes himself *de se* to be Irene Heim.

To implement this idea, let us optionally label each attitude verb with a two positive integers, called its *logophoric indices*. Tensed verbs may therefore bear three indices: a temporal

[153] I owe this twist on Perry's (1977) example of Heimson and Hume to Liam Moran.

[154] This is for the same reasons as were pointed out in Footnote 142 above for the analysis in Higginbotham (2003) and in the discussion just above for the diagonal reading of *I am Rudolf Lingens*. Humeson might hear about a particular experiential state and form attitudes towards it, including the belief that its subject is Irene Heim, without self-identifying as the subject of that state.

index and two logophoric indices. By convention, let us write the logophoric indices as superscripts on the verb. One of the logophoric indices will serve to bind individual variables, and the other to bind temporal variables. Let us write the individual index to the left and the termporal index to the right. Thus in $believes_i^{j,k}$, i is the ordinary temporal index associated with tense, j is individual logophoric index, and k is the temporal logophoric index.

Now where x is any individual, let us write '$\text{EXP-STATE}_{w,t}(x)$' for the experiential state of x in world w at time t. Our analysis depends on the idea that a given experiental state may have different subjects in different worlds and times, so for any experiental state s, let us write '$\text{SUBJ}_w(s)$' for the subject of s in world w and $\text{TIME}_w(s)$ for the time of s in w. We require that for all w, t, x: $\text{SUBJ}_w(\text{EXP-STATE}_{w,t}(x)) = x$ and $\text{TIME}_w(\text{EXP-STATE}_{w,t}(x)) = t$ That is, if s is x's experiential state in w at t then x is the subject of s in w and t is the time of s in w.

Now we need to give a semantics for $believe_{inf\,i}$. Our earlier definition for $believe_{inf}$, given in (190) is repeated here as (234):

(234) If $\text{USE}(\alpha, believe_{inf})$ then $Lex(\alpha, u, w, p) = [\lambda q : q \in \mathbf{D}_{\langle s,t\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . \exists p' \forall w' [w' \in \mathbf{B}_{x,t,w} \rightarrow q(w', p') = \mathbf{truth}]]$

We may define $believe_{inf}^{j,k}$ as in (235):[155]

(235) If $\text{USE}(\alpha, believe_{inf}^{j,k})$ then $Lex(\alpha, u, w, p) = [\lambda q : q \in \mathbf{D}_{\langle r,\langle s,t\rangle\rangle} . \lambda x : x \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . \exists p' \forall w' [w' \in \mathbf{B}_{x,t,w} \rightarrow \exists r' [Ref_{u,\alpha} \preccurlyeq r \,\&\, \lambda y \lambda t' [q(w', p')(r'[y/j][t'/k])](\text{SUBJ}_{w'}(\text{EXP-STATE}_{w,t}(x))(\text{TIME}_{w'}(\text{EXP-STATE}_{w,t}(x))) = \mathbf{truth}]]]$

This gives a rather complex appearance, but the idea behind it is much simpler, and is perhaps most easily understood by working through an example. Suppose φ is a use of the clause $he_1\ is_5\ Irene\ Heim_4$, appearing as part of the larger sentence $Humeson_2\ believes_3^{1,5}\ he_1\ is_5\ Irene\ Heim_4$. Given its syntactic position, we may assume that φ is an intensional use, and we have no reason to assume it is diagonal, so we may identify its denotation (relative to a given triple u, w, p) as follows:[156]

(236) $[\![\varphi]\!]^{u,w,p} = [\lambda \langle w', p' \rangle . [\lambda r : Ref_{u,\varphi} \preccurlyeq r . r(3) = Ref_{u,\varphi}(4)]]$, provided $Ref_{u,\varphi}(4)$ is named *Irene Heim* in u (undefined otherwise).

This function may be regarded as substituting for q in (235). Thus, '$\lambda y \lambda t' [q(w', p')(r'[y/j][t'/k])]$' in (235) may be reduced to $\lambda y \lambda t' [y = Ref_{u,\varphi}(4)]$ — the function which maps any individual and time onto **truth** iff that individual is Irene Heim and onto **falsity** otherwise.[157]

Next '$\lambda y \lambda t' [y = Ref_{u,\varphi}(4)](\text{SUBJ}_{w'}(\text{EXP-STATE}_{w,t}(x)))(\text{TIME}_{w'}(\text{EXP-STATE}_{w,t}(x)))$' will reduce to **truth** iff the subject in w' of the experiential state of x in w at t is Irene Heim.

---

[155] $r[y/i]$ is that assignment like r except that it assigns y as the value of i.

[156] Obviously, we would need to augment our lexicon to include the name *Irene Heim* and the *is* of equality (as opposed to the copula *is* which takes a predicative complement). I forgo an explicit formulation of these lexical entries here (or that for the name *Humeson*).

[157] I assume that identity is time-insensitive, so the time argument is superfluous in this example.

The long condition '$\exists r'[Ref_{u,\alpha} \preccurlyeq r$ & $\lambda y \lambda t'[q(w', p')(r'[y/j][t'/k])](\text{SUBJ}_{w'}(\text{EXP-STATE}_{w,t}(x))$ $(\text{TIME}_{w'}(\text{EXP-STATE}_{w,t}(x))) =$ **truth**]' in (235) existentially closes the subordinate clause (which in this case is vacuous). It thus amounts to a condition that the subject in w' of the experiential state of x in w at t is Irene Heim.

The sentence use φ as a whole will therefore be true in a world w iff for every world w' in the set of worlds consistent with what Humeson believes at the time of φ in w, the subject in w' of Humeson's experiential state in w at the time of φ in w is Irene Heim. Humeson believes *de re* of his own current experiential state that its subject is Irene Heim.

So much for the semantics of the attitude verb and the subordinate clause. The reading produced does not quite have the truth condition that Humeson believes himself *de se* to be Irene Heim. But this is correct; the sentence does not literally require him to do so. We naturally interpret it that way only because we typically assume that Humeson believes himself *de se* to be the subject of his own experiential state. That is, we assume:

(237) For all w', x' such that $\langle w', x' \rangle \in \mathbb{B}_{\text{Humeson},t,w}$: $\text{SUBJ}_{w'}(\text{EXP-STATE}_{w,t}(\text{Humeson})) = x'$.

This authentically *de se* belief appeals to $\mathbb{B}_{\text{Humeson},t,w}$ rather than $\mathbf{B}_{\text{Humeson},t,w}$; that is, it depends on Lewis-style self-predication, and does not have a truth-value-bearing content as its object. But if we assume that Humeson has such a belief, then if he also stands in the belief relation to the truth-value-bearing content of the clause *he is Irene Heim* as just sketched, it is fair to conclude that he believes himself *de se* to be Irene Heim.

As a second example using the same basic technique, let us consider the semantics of *de nunc* attitude expressions, such as Prior's example *Thank goodness that's over*. Prior writes:[158]

> One says, e.g. 'Thank goodness that's over!', and not only is this, when said, quite clear without any date appended, but it says something which it is impossible that any use of a tenseless copula with a date should convey. It certainly doesn't mean the same as, e.g. 'Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954', even if it be said then.

The argument here is compact and not very explicit, but the point is clear, I think: one can have and express different attitudes toward the contents of sentences like *That's over* and *The date of the conclusion of that thing is Friday, June 15, 1954*, even when it is Friday, June 15, 1954. Therefore we must distinguish these contents in some way. Prior advocates differentiating these contents by analyzing *That's over* as expressing a content that is was false but has been true since Friday, June 15, 1954, and analyzing *The date of the conclusion of that thing is Friday, June 15, 1954* as expressing a content which was, is, and always will be true. In a theory in which truth is relativized to indices, we might do something similar by analyzing the first sentence as expressing a content that is true relative to some times but not others, but the second as expressing a content that is true relative to all times (though Prior himself would not have approved of such a move).

I will argue to the contrary that both sentences should be assigned contents which are true or false eternally; the impression of a centered reading for *That's over* comes from the plausible

[158] Prior (1959) p. 17.

assumption that the speaker has a *de se* attitude which is *not* toward the content of this sentence, but which relates to it in a systematic way.

Let us assume that the sentence *Thank goodness that's over* is roughly paraphrasable as "*I thank goodness that it is over*." We may assume *over* has the following lexical semantics:

(238) If USE(α, *over*) then $Lex(\alpha, u, w, p) = [\lambda x : x \in \mathbf{D}_e \,.\, \lambda t : t \in \mathbf{D}_i \,.\, x$ is over in w at $t]]$

An intensional-horizontal use φ of it_1 is_2 *over* will therefore have the denotation shown in (239):

(239) $[\![\varphi]\!]^{u,w,p} = [\lambda\langle w', p'\rangle \,.\, [\lambda r : Ref_{u,\varphi} \preccurlyeq r \,.\, r(1)$ is over in w at $r(2)]]$

As with *believe*, we let *thank* be labeled with logophoric indices, and define it as follows:

(240) If $\text{USE}(\alpha, thank_{inf}^{j,k})$ then $Lex(\alpha, u, w, p) = [\lambda q : q \in \mathbf{D}_{\langle r,\langle s,t\rangle\rangle} \,.\, \lambda z : z \in \mathbf{D}_e \,.\, \lambda x : x \in \mathbf{D}_e \,.\, \lambda t \in \mathbf{D}_i \,.\, \exists p'\forall w'[w' \in \mathbf{T}_{x,z,t,w} \rightarrow \exists r'[Ref_{u,\alpha} \preccurlyeq r \,\&\, \lambda y\lambda t'[q(w', p')(r'[y/j][t'/k])](\text{SUBJ}_{w'}(\text{EXP-STATE}_{w,t}(x))(\text{TIME}_{w'}(\text{EXP-STATE}_{w,t}(x))) = \mathbf{truth}]]]$

Here $\mathbf{T}_{x,z,t,w}$ is the set of worlds consistent with everything for which x thanks z at t in w. For simplicity adopting the (debatable) assumption that *goodness* should be analyzed like a proper name (and assuming an appropriate lexical entry for the first-person present form of *thank*), we may see that a use ψ of I_4 $thank_5^{3,2}$ $goodness_6$ $that\ it_1$ is_2 *over* will have the truth condition in (241):[159]

(241) $[\![\psi]\!]^{u,w,p} = \mathbf{truth}$ iff $\exists p'\forall w'[w' \in \mathbf{T}_{speaker_{u,\psi}(4),\ Ref_{u,\psi}(6),time_{\alpha,u,w}} \rightarrow Ref_{u,\psi}(1)$ is over in w at $(\text{TIME}_{w'}(\text{EXP-STATE}_{w,t}(x)))]$

The appeal to experiential states distinguishes this analysis from one which Prior rejects, in which *Thank goodness that's over* means something like "Thank goodness the conclusion of that thing is contemporaneous with this utterance." As Prior puts it, "Why should anyone thank goodness for that?" The mere fact that the conclusion of some unpleasant event temporally precedes a particular experiential state does not by itself provide any more of a reason to be thankful than the fact that it precedes a particular utterance; but this fact in combination with knowing that experiential state to be one's own current experiential state does provide such a reason. Knowing a particular experiential state to be one's own can be treated Lewis-style as an attitude toward a non-truth-bearing property, rather than a proposition: $\lambda x[x$ is the subject of experiental state $s]$ — or to frame it in the terms of our analysis here: in w at t, x knows himself or herself (*de se*) to be the subject of state s iff for all w', x' such that $\langle w', x'\rangle \in \mathbb{K}_{x,t,w}$: $\text{SUBJ}_{w'}(s) = x'$.

7.6. *De se* belief and autocentric truth assessment

With the account of infinitival control and *de se* attitude reports developed in earlier in this chapter

[159] Provided $Ref_{u,\psi}(1)$ is sexless in u and $Ref_{\psi,u}(6)$ is named *goodness* in u; undefined if these conditions are not met, **falsity** otherwise.

now in place, we can draw out more clearly the connection between autocentric truth assessment and *de se* belief.

Recall from Section 7.1. that an individual x believes a content Φ from a perspective p at time t in world w iff for all $w' \in \mathbf{B}_{x,t,w}$, $\Phi(w', p) =$ **truth**, where $\mathbf{B}_{x,t,w}$ is the set of worlds x regards at t in w as candidates for the actual world. Belief involves truth assessment — we believe what we have assessed as **true** (with no intervening reassessment) — so we can also say in this case that at t in w, x has assessed Φ as **true**, adopting a stance with p as the relevant perspective.

In autocentric belief and truth assessment, the adopted stance uses the assessor's own perspective: If x believes Φ autocentrically at t in w, then for all $w' \in \mathbf{B}_{x,t,w}$, $\Phi(w', \langle x, t, w\rangle) =$ **truth**. But as pointed out in Footnote 120, this does not fully capture what it means to believe Φ autocentrically. Autocentric assessment (and belief) is *de se*, something this statement does not capture. For example, suppose that John is in a charitable mood: he wants to buy a tasty piece of cake for a random stranger. He catches a glimpse of his own reflection in the mirror, but fails to recognize himself; he thinks the reflection is that of a suitable random stranger and decides to buy that person some cake. After buying the cake, he absentmindedly takes a bite, and dislikes it — but puts on a smile as he swallows in order not to offend anyone. As he does so he sees himself in the mirror, and — still failing to recognize his own reflection — takes the "stranger's" smile as sufficient evidence to assess the cake as tasty from his perspective, so he decides to go ahead with the plan and buy the "stranger" another piece of cake. John (wrongly) assesses the content of *The cake is tasty* as **true**, based on a perspective $\langle x, t, w\rangle$, where unbeknownst to John, x is John himself. In this kind of case, we should count the assessment as exocentric, not autocentric, even though it is based on a perspective which turns out to be John's own.

To say whether x autocentrically believes Φ at t in w, we must appeal not to $\mathbf{B}_{x,t,w}$ — a set of worlds — but to $\mathbb{B}_{x,t,w}$, the corresponding set of world-individual pairs. We let $\mathbb{B}_{x,t,w}$ be the set of all pairs $\langle w', x'\rangle$ such that being x' in w' is compatible with x's belief-state at t in w; that is, such that x regards being x' in w' as a candidate for his or her own situation. Now we can say:

(242) x autocentrically believes Φ at time t in world w iff for all $\langle w', x'\rangle \in \mathbb{B}_{x,t,w}$, $\Phi(w', \langle x', t, w'\rangle) =$ **truth**.[160]

To autocentrically assess (at t in w) Φ as **true** is simply to implicitly or explicitly consider Φ and come to autocentrically believe it (at t in w).

In Section 7.1. I suggested that coherence of belief depended on a division of an agent's beliefs into separate "p-bodies" corresponding to the various perspectives used in assessing sentence contents. In addition, we should now recognize, for any individual x, the body of beliefs which x has autocentrically — what we might call x's *auto-body*. As with the p-bodies, internal coherence of an agent's auto-body is required for global coherence of belief; but contents belonging to an agent's auto-body may contradict contents belonging to a p-body without resulting in incoherence.

7.7. Against non-indexical contextualism

[160] Rephrased in more Lewisian terms, x autocentrically believes Φ at time t in world w iff at t in w, x self-predicates $\lambda w' \lambda x'.\ \Phi(w', \langle x', t, w'\rangle)$.

It may be instructive to consider our present system in light of the typology of contextual sensitivity presented in MacFarlane (2009) and related work. Like us, MacFarlane advocates a semantic theory where the truth value of a sentence may vary with the context of use and/or with the context of assessment — indeed, the explicit idea that truth may be sensitive to contexts of assessment, and the characterization of semantic relativism in these terms, are both due to MacFarlane. Also roughly like us, MacFarlane assigns contents to linguistic expressions and denotations to contents, and allows that both assignments may be sensitive to context.

If, as theorists, we countenance sensitivity both to contexts of use and to contexts of assessment, both in the assignment of contents to expressions and the assignment of denotations to contents, this suggests four possible ways in which truth and denotation may ultimately depend on context, summarized in the table below:[161]

(243)

	Assignment of *contents* to expressions is sensitive to context	Assignment of *denotations* to contents is sensitive to context
Assignment is sensitive to context of *use*	**Indexical contextualism**	**Non-indexical contextualism**
Assignment is sensitive to context of *assessment*	**Indexical relativism**	**Non-indexical relativism**

In this typology, the upper left-hand box, *indexical contextualism*, represents perhaps the least controversial category — ordinary indexicality of the classical sort represented by personal pronouns like *I* and *you*, analyzed more-or-less along the lines of Kaplan (1989), with contents that vary according to the context of use. The lower right-hand box, *non-indexical relativism*, represents relativism of the kind I have been advocating in the analysis of personal taste sentences, where the truth value of such a sentence depends on the context of assessment, but the content remains constant even as the truth value varies. The lower left-hand box, *indexical relativism*, represents a kind of context-sensitivity which has been comparatively little advocated (but see Weatherson (2009), Parsons (2011)); it would call for a single use of an expression to vary in content, depending on the context of assessment. This kind of variation would seem to be incompatible with our assumption on p. 8 that each sentence has exactly one content in context, aside from special cases like puns and deliberate double-entendres.

The upper right-hand box, *non-indexical contextualism*, is our main concern in this section. A sentence exhibiting this kind of context-sensitivity would vary in truth value according to the context of *use*, but its content would remain constant even as the truth value varies. Analyses which appeal to this sort of context-sensitivity have been advocated with some frequency in the treatment of tense, and also perhaps in the analysis of *de se* attitudes — though this is less clear. I

[161] Some of the terminology here is taken from Weatherson (2009), who gives a similar table.

will argue in this section that this style of analysis is not only unnecessary, but also faces significant conceptual difficulties; and should therefore not be used in semantic theory.[162]

Non-indexical contextualism is perhaps best understood by example. Consider the standard semantics for tense logic, as implemented in Montague (1968; 1973), Lewis (1972), Kaplan (1989), etc.[163] Here, the truth values of sentences are assigned relative to world-time pairs, so a single sentence might vary in truth value (even if we hold the world constant) with the passage of time: a sentence like *John saw Mary* might be false at time t_1 (in a given world) but true at some later time t_2 (in the same world), if he saw her in the meantime. This is variation in truth value according to the context of *use*, not assessment: in judging whether someone has spoken truthfully or not, we consider the truth value of the sentence relative to the time at which the speaker used it, not relative to the time when we perform the truth assessment. (Hence if Bill said *John saw Mary* at t_1, he spoke falsely, no matter if we consider the issue at t_2.) The content (or intension) of a sentence $\boldsymbol{\varphi}$ in these analyses is the function which maps any world-time pair $\langle w, t \rangle$ onto the truth value of $\boldsymbol{\varphi}$ relative to $\langle w, t \rangle$ — a single function which yields different values for different world-time pairs, not different functions at different times — so it is just the denotation (truth value) which varies with the context of use, not the content. In other worlds, this analysis treats tense in a non-indexical contextualist fashion.

But we rejected this approach to the semantics of tense in Section 3.4. , primarily on the grounds that it complicated the definition of contradiction. Instead, we gave an indexical analysis, in which *John saw Mary* expresses a different content when used at t_1 than when used at t_2.

Positive arguments in favor of a non-indexical contextualist analysis of tense have certainly been advanced, but few if any of these are very compelling. One frequently cited argument comes from a brief passage in Kaplan (1989):

> If we build the time of evaluation into the contents..., it would make no sense to have temporal operators. To put the point another way, if what is said is thought of as incorporating reference to a specific time, or state of the world, or whatever, it is otiose to ask whether what is said would have been true at another time, in another state of the world, or whatever. Temporal operators applied to eternal sentences ... are redundant.

Roughly, the argument is that we need time indices so that temporal operators will have something to operate on. This is only a weak argument for temporal indices, because it presupposes that tense markers are to be treated as intensional operators — that is, as operators that manipulate indices involved in the assignment of truth values to sentence contents. But this is a highly debatable assumption, and a number of alternative approaches have been proposed and defended in some detail. The analysis given above in Section 3.4. and Section 7.5. provides just one example.

To make an argument like Kaplan's stick, then, it would be necessary to show that a treatment of tenses as intensional operators is correct, or at least superior to the alternatives. We

---

[162] I repeat here rather compactly some arguments made in more detail in Lasersohn (2013), incorporating some text from that article.

[163] Of course these authors were building directly on Prior (1957) and related work; but their formalism is different, particularly in its use of indices.

might do this by showing that tenses display systematic scope ambiguities, for example — but as Enç (1981) has argued in detail, examples that initially appear to display scope ambiguities involving tense turn out on closer inspection to be quite problematic to analyze in this way. Moreover, even if compelling reasons could be found to treat tense markers as having scope, this would still be compatible with treating them as variable-binding operators rather than intensional operators. Familiar examples show that we cannot assign truth-values to sentences relative to just a single time parameter, and analyze tense markers as manipulating that parameter (Kamp (1971)):

(244) A child was born who will become ruler of the world
$\neq$ **P** $\exists x$[born(x) & **F** ruler(x)]

But such examples are easily handled in a logic with an arbitrary number of time parameters or variables, and quantifiers which bind them:

(245) $\exists x \exists t[t < t^* \ \&\ \text{born}(x, t) \ \&\ \exists t'[t' > t^* \ \&\ \text{ruler}(x, t')]]$

Tenses often seem to be interpreted more like bindable variables than like operators (as observed by Partee (1973) and many others). Adverbial quantifiers can bind variables over times:

(246) Whenever John smoked, Mary got angry.

If we analyze *John smoked* on the model of '**P** smoke(j)' and *Mary got angry* on the model of '**P** angry(m)', and then treat sentences of the form ⌜*Whenever* **φ**, **ψ**⌝ as true iff at every time at which **φ** is true, **ψ** is also true, we get the wrong results. If past tense morphemes are just variables over past and future times, however, the semantics is straightforward.

A sufficiently sophisticated version of the operator analysis can deal with such examples (Ludlow (1999), Ludlow (2006)), so they should not be seen as incompatible with an operator-based approach. But why appeal to such an analysis if a variable-binding analysis is available? Philosophers sometimes argue for an operator-analysis of tense on the grounds that such an analysis allows us to avoid committing ourselves to an ontology of past and future times, but a theory that claims that sentence contents have truth-values only relative to times must surely be committed to an ontology of times anyway, so presentist arguments for an operator analysis of tense will be of no help to the defender of such a theory.[164]

[164] A presentist might here object that what is fundamentally at issue is not whether sentences express contents that are true or false relative to times, but whether they may express contents that have one truth value now, but had a different one earlier, or will have a different one later. For the metaphysically scrupulous, the claim that sentences do express such contents may be attractive because it suggests a way out of the ontological commitment to past and future times. The basic idea behind such arguments is that a T-sentence such as (i) does not contain any overt quantification over times — and can be derived compositionally without assigning a time as the semantic value of any constituent of **Pφ** — and therefore does not involve a commitment to them (in contrast to (ii), which does):

(i) ⟨**Pφ**⟩ is true iff ⟨**φ**⟩ was true.
(ii) ⟨**Pφ**⟩ is true relative to t iff there exists some t earlier than t such that ⟨**φ**⟩ is true relative to t.

Personally, I am not so confident that (i) does not involve an ontological commitment to times. Questions of ontological commitment are best resolved by presenting an explicit model-theory, not by appeals to the syntactic form

A somewhat different argument for a non-indexical contextualist analysis of tense may be given based *de nunc* attitudes, as in Prior's example *Thank goodness that's over*, discussed in Section 7.5. above. As Prior pointed out, this sentence is not accurately paraphrased as "Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954," even when used on that date. If, however, the sentence expresses a time-neutral content — true relative to some times and false relative to others — we might regard that content as an appropriate object for which one might bear an attitude of thankfulness (or so the argument goes).

But we already saw in that section that such examples may be treated using a temporal logophoric index on the verb, binding the temporal variable expressed by the tense in the subordinate clause. This analysis did not make any appeal to contents which are true or false relative to times. There was no claim that any sentence expressed a content which was true relative to some contexts of use and false relative to others; so no appeal to non-indexical contextualism.

Likewise, there seems no reason to suppose that *de se* or other centered attitude sentences provide us with a reason to appeal to a non-indexical contextualist style of semantic analysis. The idea here would be that *de se* attitudes are attitudes toward contents which are true or false relative to centered worlds, where centered worlds may be identified with ordered pairs of worlds and spacetime locations, or ordered triples of worlds, times and individuals, or something similar. But even if we appeal to centered worlds in the analysis of attitude reports, it is far from clear that we must regard any sentence as expressing a content which varies in truth value according to the context of use; in centered attitude reports it is the context of the attitude-holder, not the context in which the report is made, which is crucial. To my knowledge, no one has argued that in a sentence like *John expects to win*, the clause *PRO to win* expresses a content which varies in truth value according to who uses that clause.

Even with regard to free-standing sentences like Rudolph Lingens' utterance "I am Rudolph Lingens!" we have seen how to use diagonalization to represent the content of the sentence as a simple set of possible worlds, not centered worlds. In *some* sense this content varies in truth value according to the context, since worlds determine contexts for particular utterances; but there is no reason to consider this a non-indexical contextualist analysis, any more than possible-worlds analyses in general should be considered non-indexical contextualist.

If tense and centered attitudes fail to provide any positive reasons in favor of non-indexical contextualism in semantics, we may ask whether there are any specific reasons which militate *against* making use of such an approach. I believe there are — not just in the analysis of tense and centered attitudes but in semantic theory more generally.

We may note first that a non-indexical contextualist semantic theory will greatly complicate the task of defining disagreement and contradiction. For the most part, semanticists have assumed (implicitly or explicitly) a fairly simple account of what it means to disagree — usually something along the lines of the following definition, which as formulated but rejected by MacFarlane (2007):[165]

---

of T-sentences.

[165] I have slightly compressed MacFarlane's wording here and in other definitions, in the interest of readability. In this definition, 'accepts' and 'rejects' should be understood broadly enough to cover both speech acts such as assertion and denial, and mental acts or states such as belief and disbelief. MacFarlane uses the term 'proposition' while leaving open the issue whether propositions might have truth values relative to indices other than worlds; because some people understand 'proposition' as by definition implying that all contextual effects have been resolved so that an absolute

(247) Two parties disagree just in case there is a proposition that one party accepts and the other rejects.

MacFarlane rejects this as an adequate account of disagreement, but apparently for the sole reason that it conflicts with the possibility of a non-indexical contextualist semantic analysis:

> Consider, for example, tensed propositions, which have truth-values relative to world/time pairs. One such proposition is the proposition that Joe is sitting. (Do not confuse this with the proposition that Joe is sitting now, or at any other time: the tensed proposition is, in Kaplan's terms, 'temporally neutral.') If you asserted this proposition at 2 pm and I denied it at 3 pm, we have not in any real sense disagreed. Your assertion concerned Joe's position at 2 pm, while my denial concerned his position at 3 pm. So accepting and rejecting the same proposition cannot be sufficient for genuine disagreement. (MacFarlane (2007) p. 22)

Of course, the argument from this example is only as strong as the assumption that there are temporally neutral propositions, whose truth-values vary with a time index — but it is precisely this sort of assumption I would now like to call into question. MacFarlane considers several alternative definitions of disagreement, but finds that all of them are problematic. One possibility is to build an appeal to contexts into the definition of disagreement, in the guise of a condition on 'accuracy':

(248) a. An acceptance (rejection) is *accurate* just in case the proposition accepted is true (false) at the circumstance of evaluation that is relevant to the assessment of the acceptance (rejection) in its context (or at all such circumstances, if there is more than one).
c. Two parties *disagree* iff (i) there is a proposition that one party accepts and the other rejects, and (ii) it is not the case that both the acceptance and the rejection are accurate.

But this still will not work, as MacFarlane points out:

> Suppose that at noon Mary accepts the tensed proposition that Socrates is sitting, and at midnight Tom rejects this proposition. And suppose that Socrates was sitting at both noon and midnight. Then Mary's acceptance is accurate, while Tom's rejection is inaccurate, but still there is no disagreement. (MacFarlane (2007) p. 24)

We might try modalizing the accuracy condition:

(249) Two parties disagree iff (i) There is a proposition that one party accepts and the other rejects, and (ii) the acceptance and the rejection *cannot* both be accurate.

But this too fails as a characterization of disagreement:

---

truth value can be assigned, I would prefer the term 'sentence content' here.

> Suppose that at noon Mary accepts the tensed proposition that the number of flies in the room is either odd or even, and at midnight Tom rejects this same tensed proposition. Then Tom's rejection *can't* be accurate (he is, after all, rejecting a necessary truth), so *a fortiori* Mary's acceptance and Tom's rejection can't *both* be accurate. Yet they do not disagree: Mary's thought concerns noon, while Tom's concerns midnight. (MacFarlane (2007) p. 24)

Despite his own identification of this problem, MacFarlane uses (249) as a working account of disagreement, in order to move on and explore other issues. There is nothing wrong with his doing so, but it seems clear that some better account of disagreement is needed — as I think MacFarlane himself would be the first to admit. Especially given the role that disagreement has played in arguments for relativist semantics, it seems vital to the relativist project to provide an account of disagreement that works at least as well as accounts that are available in a non-relativist semantic framework.

Fortunately, relativists do not have much to worry about from the difficulties MacFarlane points out in the definition of disagreement. None of these difficulties arises because of the theoretical possibility of assigning truth values to sentence contents relative to indices whose values are fixed by the context of assessment. They all arise because MacFarlane is deliberately keeping open the possibility that sentence contents may be assigned truth values relative to indices whose values are fixed by the context of use — that is, they arise because he is committed to the possibility of non-indexical contextualist analyses in semantic theory. If we reject the use of indices in the assignment of truth values to sentence contents when the values of such indices are fixed by the context of use, a simple definition of disagreement along the lines of (247) regains its plausibility. By rejecting such indices, we greatly simplify and render less problematic the theory of disagreement.[166]

---

[166] MacFarlane suggests that the same problems he points out for the definition of disagreement if we allow temporally neutral propositions will arise even for propositions whose truth values vary with nothing more than a possible world index:

> Consider Jane (who inhabits this world, the actual world) and June, her counterpart in another possible world. Jane asserts that Mars has two moons, and June denies this very proposition. Do they disagree? Not in any real way. Jane's assertion concerns our world, while June's concerns hers. If June lives in a world where Mars has three moons, her denial may be just as correct as Jane's assertion. (MacFarlane (2007) p. 23)

But as argued by Cappelen and Hawthorne (2009), it does not really make sense to ask whether two people in different worlds disagree with one another:

> …the claim that each of two individuals in different worlds accepts some proposition P is not akin to the claim that two individuals in different countries accept that proposition. The latter claim entails that there are two individuals that accept P. But, even if we unrestrict our quantifiers as far as possible, the former claim does not entail that there are two individuals that accept P.... Thus, possible worlds scenarios of the sort that MacFarlane entertains do not provide examples where a pair of individuals accept some content P but nevertheless fail to agree that P, or examples where a pair of individuals accept P and not-P respectively but nevertheless fail to disagree that P. (Cappelen and Hawthorne (2009) p. 64)

A second reason for rejecting the possibility of non-indexical contextualism in semantic theory is that it forces us to draw an odd distinction between what we accept as true and what we accept someone as having truthfully said.

As an example, consider a non-indexical contextualist analysis of knowledge ascriptions, in which the truth value, but not the content, of a knowledge ascription varies with a standard of evidence fixed by the context of use. John says "Bill knows he has hands" in a context with a sufficiently liberal standard of evidence to qualify him as speaking truthfully. Later, in a much more stringent context, we discuss Bill's evidence and conclude that he does *not* know that he has hands. Nonetheless, we regard John as having spoken truthfully when he said that Bill did know he had hands, because the context in which he spoke involved a more liberal standard. So, we accept both that Bill does not know he has hands, and that John spoke truthfully in saying that Bill knows that he has hands.

This pattern would not be puzzling or troubling in an analysis in which the content of *Bill knows he has hands* is different in the two contexts. But if we claim that John said the very same thing in claiming that Bill knows he has hands as we now deny in saying that he does not, it is odd indeed for us to acknowledge that he was speaking truthfully. If John said something that we agree is false, it would be inconsistent on our part to regard him as having spoken truthfully.

This is a general problem to any analysis using contents that are assigned truth values relative to indices fixed by the context of use, and not just knowledge ascriptions. Assume some sentence **φ** invariantly expresses a content φ which is true relative to index values provided by the context *c* in which **φ** is used. We must regard the speaker of **φ** in *c* as speaking truthfully, even if we are not in *c* ourselves, because the indices are fixed by the context of use, not assessment. But we can also form attitudes toward φ, and if we are in a context *c′* which provides index values relative to which φ is false, we are correct taking an attitude of disbelief toward φ — even though we are also correct in believing the speaker to have spoken truthfully in asserting φ. But it is intuitively inconsistent to regard someone as having spoken truthfully if the content of their assertion is something we disbelieve.

The most one can ask about disagreement in scenarios like the one MacFarlane presents, I think, is not whether Jane disagrees with June, but whether Jane (actually) disagrees with what June would have been saying, if Mars had three moons and June said "Mars does not have two moons." It seems to me that she does, at least if we understand *two* in this example as meaning "exactly two." (This is not to say that Jane would have disagreed with June under such circumstances, of course.) Such examples, then, do not provide us with a reason to regard the possible world index as fixed by the context of use. The idea that it is fixed by the context of use seems to originate with Kaplan (1989) p. 547, according to which a sentence **φ** is true relative to a context *c* iff the content expressed by **φ** in *c* is true relative to the world (and time) of *c*. Here, we have instead adopted the more usual position that the content of a sentence is true iff it is true relative to the actual world — not the world of the context in which the sentence is used. (See (142) above.)

Chapter 8 : Assertion and Other Speech Acts

In this chapter, we consider the notion of relative truth as it relates to the theory of speech acts, especially — though not exclusively — the act of assertion. Relative truth poses challenges not just for the theory of semantics, but also for the theory of linguistic pragmatics: before we can accept it, we must show that it is possible to make sense of the idea of asserting a content which is true relative to some perspectives but not others (and of questioning, ordering, or performing other speech acts with the kinds of contents which a relativist semantics requires).

We may approach such issues in a variety of ways, depending on whether our concern is with fundamental philosophical questions about what it means to assert something, or with more "functional" questions about the role of assertion in discourse structure, or with teleological questions about why someone would make an assertion.

8.1. Assertion, Norms, and Portrayal as True

At a very basic level, asserting a content seems to involve *portraying that content as true*. But in our current system, we have several differently defined truth predicates (see (139), (140), (141), (142), (143), (144), (183)); the question naturally arises which of these is the relevant one. If we take the wording used above seriously, so that assertion involves portrayal of a content as "true" rather than "true relative to such-and-such a parameter," we can eliminate (139), (140), (141), and (144), since those define multivalent truth predicates, not any notion of truth *tout court*. We can also eliminate (143) and (144), since those define truth for uses or sentences, not contents. This leaves (142) and (183). But (142) defines "absolute" truth — contents are not true in this sense unless they are true from all perspectives, with no variation from person to person, time to time, etc. To portray the content of a taste sentence (or any other content which varied in truth value from perspective to perspective) as true in the sense of (142) would be misleading; a speaker with thorough knowledge of the truth conditions of that content could not sincerely portray it as true in this sense. This leaves (183) — which characterizes **true**, the only one of our truth predicates which itself behaves relativistically, but which also appears to be essentially redundant.

What does it mean to portray a content as true in this sense? In general, to portray X as Y is to (purposefully, publicly, cooperatively) behave in a way in which one would *normally* behave only if X were Y.[167] Hence, to assert a content Φ is to purposefully, publicly, cooperatively behave in a way which one would normally behave only if Φ were **true**.[168]

How does one normally behave if a sentence content is **true**? Given the redundancy of **true**, this varies widely by example. To behave as though Φ is **true** is to behave as though Φ. Behaving as though it is **true** that the stove is hot means behaving as though the stove is hot: one refrains from touching it, warns other people not to, and says things like "The stove is hot!" Behaving as though it is **true** that the chili is tasty means behaving as though the chili is tasty: one eats the chili, recommends it to others, and says things like "The chili is tasty."[169]

---

[167] Recall our discussion in Chapter 6 of speaking and assessing "as though" the author, assessor, etc. were different.

[168] This definition is deliberately formulated to include as assertions not only those cases where a speaker uses a sentence whose content is the object of the assertion, but also indirect linguistic assertions and many sorts of non-linguistic signaling.

[169] There is, I think, very little one can say beyond this kind of case-by-case characterization; but one possible

So long, then, as we can sensibly regard people's purposeful, public, cooperative behavior as sometimes conditioned by contents which are true or false relative to some perspectives but not others, it seems that we can make sense of the idea of portraying such contents as **true**, hence of asserting them. We have already seen in some detail what is involved in assessing such contents, believing them, and bearing other sorts of mental attitudes toward them. If such contents can be the objects of attitudes, we can surely allow them a role in reasoning and decision-making, hence in conditioning actions.

If asserting a content Φ means behaving in a way which one would *normally* behave only if Φ were **true**, we may say that the act of asserting Φ is governed by a *norm* that Φ be **true**. This places our approach in the very broad and longstanding line of analysis which considers assertion to be governed, perhaps constitutively, by such a norm.

However, considerable caution is necessary in interpreting the claim that assertion is governed by such a norm, and the related claim that asserting a content Φ means behaving in a way which one would normally behave only if Φ were true. The word *normally* is often understood as a rough synonym of *typically*, or *ordinarily*. That is not the relevant sense of *normally* here, however. It is quite typical and ordinary for speakers to make assertions which are not, strictly speaking, true; and such assertions are not necessary dishonest or deceptive.

This is because there are several other norms which also govern assertions: assertions must be relevant, informative, appropriately worded, etc. — that is, they must conform to something like Grice (1975) Maxims of Conversation. The norm of truthfulness perhaps has a special theoretical status among these, but it by no means takes precedence over the others in affecting speaker behavior. In many cases it is impossible to meet all such norms at once, and the norm enforcing truthfulness is frequently violated in order to avoid more serious violations of the other norms.

For example, as discussed in more detail in Lasersohn (1999), in many contexts it is more appropriate to say "John arrived at three o'clock" than to say "John arrived at 15.9 seconds after 3:02," even if the latter comes closer to being literally true. The latter sentence, in most contexts, would come across as unnecessarily precise, at the cost of undue wordiness and encoding irrelevant information. Falsehood is tolerated as it concerns details which are irrelevant to the matter at hand, which would detract from the overall effectiveness of the assertion if they were accurately reported. In this sense, a false sentence may still come "close enough to true for its context," and we must recognize that what is generally expected of speakers is not that all their assertions be true, but that they be close enough to true for their contexts.

Because of the interplay of conflicting norms, then, we cannot straightforwardly identify ordinary, typical speech behavior with behavior adhering to a norm of truthfulness. This does not mean that no such norm governs the act of assertion, however, and it does not stand in the way of claiming that asserting a sentence content involves behaving in a way which such a norm would dictate.

We cannot review here the extensive literature for and against the claim that assertion is subject to a norm of truthfulness; but it may be worth nonetheless considering one well-known

---

unifying feature might be that asserting a content involves undertaking a commitment to withdraw the assertion if one later assesses it as false, as MacFarlane (2005) argues in some detail.

argument that such a view is incompatible with a relativistic truth predicate, due to Evans (1985).

Evans formulates his objection in terms of the semantics of tense, but the objection itself is sometimes taken to generalize to relativism about truth more generally.[170] Evans asks us to consider a semantics for tense in which "…we abandon the idea that particular historical utterances are assessable, once and for all, as correct or incorrect. Rather, we must acknowledge that the evaluation of particular utterances must change as the world changes" (p. 347). Under this view, if at time *t* it is raining, John says "It is raining," and we evaluate his utterance for "correctness," we should regard his utterance as correct; but if at *t′* it is not raining and we evaluate his utterance then, we should regard it as incorrect, no matter that the utterance itself took place at *t*, when it *was* raining. In other words, the analysis in question is one in which the semantics of tense is sensitive to the context of assessment, not the context of use.

I know of no one who has seriously advocated this analysis of tense.[171] It may nonetheless be worthwhile to consider Evans' arguments against such an analysis:

> If a theory of reference permits a subject to deduce merely that a particular utterance is now correct, but will later be incorrect, it cannot assist the subject in deciding what to say, nor in interpreting the remarks of others. What should he aim at, or take others to be aiming at? *Maximum* correctness? But of course, if he knew an answer to this question, it would necessarily generate a once-and-for-all assessment of utterances, according to whether or not they meet whatever condition the answer gave. In fact, we know what he should do; he should utter sentence types true at the time of utterance. One who utters the sentence type 'It is raining' rules out dry weather only at the time of utterance; he does not rule out later dryness, and hence there can be no argument from the later state of the weather to a re-appraisal of the utterance. Utterances have to be evaluated according to what they rule out, and so different utterances of the same tensed sentence may have to be evaluated (once and for all) differently. (pp. 349–350)

An analogy may perhaps be drawn to a game in which players toss beanbags, and score points according to how close their beanbags are to a target. The game is more difficult, of course, if the target is moving chaotically; still, as long as the scores are calculated based on how close the beanbag is to the target *at the time that it lands*, it is possible to strategize about where to aim. But if the game is not scored in that way, but instead is scored and rescored over and over again, at each time *t* based on how close each beanbag is to the target is at *t*, no matter how close it was when it landed — in such a game there can be no strategizing about where to aim. So likewise if one tries to be "correct" in one's language use, but the correctness of an utterance does not depend on conditions at the time of the utterance, but instead is constantly shifting, there can be no strategizing about what to say in order to meet the goal of "correctness." And if speakers cannot employ any such strategy, neither does it seem that listeners can reconstruct their communicative intent from their behavior.

---

[170] Evans himself was explicit that his arguments did not generalize to relativization of truth to possible worlds.

[171] Evans suggests that such a position may have been held by scholastic and Stoic logicians and perhaps by Prior, but it is far from clear to me that he was correct in this. MacFarlane (2003) argues that the semantics of future tense is sensitive to the context of assessment, but in a different way; see also the treatment of future tense in Section 9.2. below.

We may well ask whether this sort of concern renders problematic the claim that asserting a content Φ involves purposefully, publicly, cooperatively behaving in a way one normally would behave only if Φ were **true**. At first it seems that it might — if Φ is a content which can be legitimately assessed as **true** by one person but as **false** by another, what does it mean to behave in a way one would normally behave only if Φ were **true**? If one is trying to conform to this norm, and assert only those contents which are **true**, is there a strategy one can adopt? What can one "aim for" in choosing one's words? By what standard can we judge whether a speaker is conforming to this norm or not?

In considering these questions, we should be cautious about assuming too close an analogy to the tense cases that Evans considers. If it is raining at *t*, and John says at *t* "It is raining," we judge his utterance to have been correct, even if it is not raining at *t′* when we make the assessment. But if John loves licorice and says "Licorice is tasty," we may judge his utterance to be incorrect, because we find licorice unpalatable. In the analysis of personal taste expressions (unlike the analysis of tense) we should not expect a definite, objective answer as to whether a given utterance was "correct" in Evans' sense.

But this does not preclude all strategizing about how to make a correct utterance. Each of us will strategize to "aim" at correctness according to our own judgment (so long as we are sincere and adopting an autocentric stance): John's strategy leads him to assert that licorice is tasty, while ours does not. This *does not* amount to privileging the speaker, so that an utterance counts as correct once and for all iff the asserted content is true relative to the speaker. Correctness, like **true**, behaves relativistically, so that we may legitimately judge John's strategy to have failed, even has he judges it to succeed. But constructing a strategy is a goal-oriented planning process, and the planner will naturally rely on his or her own judgment in evaluating whether the final step in a plan under consideration constitutes a satisfaction of the intended goal or not. There is, so far as I can see, nothing in the general idea of a relativistic truth predicate which renders impossible all strategizing about what to say in making assertions, or which denies such truth predicates a role in the norms governing the act of assertion.

8.2. Assertion, conversation, and context change

A rather different (though not incompatible) perspective on the act of assertion arises out of the work of Stalnaker (1979) and subsequent developments such as Heim (1982). Here, the central concern is with the effect which making an assertion has on the pragmatic context relative to which subsequent utterances are interpreted. Certain aspects of interpretation, notably presupposition and anaphora, depend on features of the context which are easily altered simply by virtue of the speaker asserting certain semantic contents. There is, for example, something anomalous about using a phrase like *the rabbit* in a context where no rabbit has been previously mentioned in the conversation or made salient in any other way; but if a speaker makes an assertion using the sentence *I found a rabbit under my porch yesterday*, subsequent uses of *the rabbit* are completely unobjectionable.

Following Stalnaker, let us use the term *common ground* for the set of sentence contents which the participants in a conversation treat (at a given time) as public knowledge which they all share. If one of the participants makes an assertion, the asserted content is added to the common

ground.[172] The use of definite noun phrases such as *the rabbit* — and other presuppositional or anaphoric expressions — is conditioned on the information in the common ground, so making an assertion may license subsequent use of such expressions.

In Stalnaker's original conception, the elements of the common ground were propositions, in the sense of *sets of possible worlds*. The intersection of these propositions, termed the *context set*, is the set of worlds which are "live options" as far as the conversation is concerned at that point — the set of worlds which may be regarded as candidates for the actual world, given everything which the conversational participants regard as a public, shared background knowledge.

Stephenson (2007b) has adapted this approach to a relativist framework, in which the common ground is not a set of sets of possible worlds, but of sets of world-time-individual triples, and the context set is the intersection of these sets. We may follow her here, making some minor adjustments in keeping with the details of our formalism: Let us regard a discourse as a sequence $D = \langle \varphi_1, \ldots, \varphi_n \rangle$ of sentence uses. For any φ_i ($1 \leq i \leq n$), $participants_{u,\varphi i}$ is a set containing $speaker_{u,\varphi i}$ and $addr_{u,\varphi i}$, regarded as the participants in the discourse in world u at the time of φ_i. We let $CG_{u,\varphi i}$ be the common ground of φ_i in world u — that is, the set of contents which the members of $participants_{u,\varphi i}$ regard as public knowledge which they all share; $CS_{u,\varphi i}$, the context set of φ_i, is defined as $\bigcap CG_{u,\varphi i}$. Let DIRECT-FORCE be a function assigning a relation between worlds, times, individuals and sentence contents to each use φ_i and world u, such that if DIRECT-FORCE(φ_i , u) = R, then $R(u, time_{u,\varphi i}, speaker_{u,\varphi i}, \llparenthesis \varphi_i \rrparenthesis^u)$. For example, if DIRECT-FORCE(φ_i, u) = ASSERT, then ASSERT(u, $time_{u,\varphi i}$, $speaker_{u,\varphi i}$, $\llparenthesis \varphi_i \rrparenthesis^u$); in other words, if the direct force of sentence use φ_i in world u is that of an assertion, then in u, at the time of φ_i, the speaker of φ_i asserts the content of φ_i relative to u. Now we require:

(250) For any discourse $D = \langle \varphi_1, \ldots, \varphi_n \rangle$, world u, and integer i ($1 \leq i < n$): $CG_{u,\varphi i+1} = CG_{u,\varphi i} \cup \{\Phi \mid$ ASSERT(u, $time_{u,\varphi i}$, $speaker_{u,\varphi i}$, Φ)$\}$

More informally, the common ground for any given sentence use in the discourse is the union of the common ground of the previous sentence use with all the contents asserted at the time of that use.[173]

As mentioned in Footnote 172 above, the asserted content does not remain as part of the common ground unless it is accepted by the participants in the discourse as a body. Thus if the speaker makes a taste claim, for example by asserting that licorice is tasty, the other participants must also accept that licorice is tasty in order for the content of the assertion to remain in the common ground. In a typical context, calling for autocentric assessment, this means that all the

[172] The asserted content does not remain part of the common ground unless the other participants in the discourse accept it (which they may do tacitly by not registering any objections). But even in the case where an objection is raised, I would regard the asserted content as temporarily added to the common ground and then removed, rather than considering it never to have been added in the first place.

[173] We appeal to the set $\{\Phi \mid$ ASSERT(u, $time_{u,\varphi i}$, $speaker_{u,\varphi i}$, Φ)$\}$ rather than simply to $\{\llparenthesis \varphi_i \rrparenthesis^u\}$ because of the possibility that in φ_i the speaker asserts some contents indirectly, in addition to any directly asserted content. The formulation in (250) must be regarded as a simplification in that it ignores the possibility of information being added to the common ground by virtue of becoming salient through some other means than assertion. To use a popular example, if a donkey unexpectedly walks into the room where the conversation is taking place, the fact that the donkey is present becomes part of the common ground regardless of whether anyone asserts that it is there.

participants in the discourse must accept that licorice is tasty according to their own standards of taste — they must all agree that licorice is tasty (or at least act as if they agreed).[174] If any of them disagrees, and regards the matter as being of sufficient importance, they must express their disagreement and raise an objection. The content is then removed from the common ground, and if anyone else presses the issue, a dispute may ensue.

There is nothing very original or surprising in this view of assertion, I think; we have simply adapted the Stalnakerian approach to a relativistic framework, basically following Stephenson. But perhaps it is worth making sure a few more of the details are clear:

Suppose that (in world w at time t) each of the participants in some discourse autocentrically accepts the assertion of some content Φ — for example the content that licorice is tasty. According to our account of autocentric assessment in Section 7.6. this means that for each such participant x, it holds that for all $\langle w', x' \rangle \in \mathbb{B}_{x,t,w}$, $\Phi(w', \langle x', t, w' \rangle) =$ **truth**. Put more informally, it means that for all x' and w' such that being x' in w' is compatible with x's belief-state at t in w, Φ is true relative to w' and the perspective of x' in w' at t. As far as x believes, if p ($= \langle x', t, w' \rangle$) is a candidate for his or her own perspective, and w' is a candidate for the actual world, then Φ is true relative to w' and p.

This has the effect that for each discourse participant x, any content added to the common ground must be true relative to at least one perspective which x takes as a candidate for his or her own perspective. It would rarely be the case that there is a single perspective which simultaneously serves as a candidate for every individual discourse participant; rather, any content added to the common ground must be true relative to multiple perspectives — at least one for each discourse participant, but with no requirement that any participants share any candidate perspectives. Since the context set is simply the intersection of all the contents in the common ground, it will be a set of world-perspective pairs. The world in each pair may be regarded as a candidate for the actual world, given what has been asserted in the discourse. The perspective is a candidate for the perspective of at least *one* of the participants.

8.3. Why assert?

Given the important role which faultless disagreement has played in motivating relativist semantic theories, it is important for us to consider carefully the role of assertion in expressing disagreement about matters of taste. Although I stressed on p. 22 above that the problem of accounting for faultless disagreement must not be conceived as a problem in explaining the conversational structure of disputes, it is still incumbent on us to explain why and how disputes about matters of taste arise. In particular, it is important to clarify why one might make an assertion about a matter of taste, and press that assertion in the face of opposition, if one really does not see any fault in making the *opposite* assertion. Even if you believe roller coasters are fun, why would you enter a dispute with those who claim that they are not fun, if you do not see their claim as faulty in some sense? And why would you press the issue and engage in a dispute, knowing that there are no facts

[174] In discussing this pattern, Stephenson (2007a) p. 509, (2007b) p. 67 suggests that it establishes the discourse participants as group as the "actual" value of the judge parameter. But this is very misleading; we do not consider a taste sentence to be true *simpliciter* iff it is true relative to the discourse participants, the way we consider an objective sentence to be true *simpliciter* iff it is true relative to the actual world. Nor — as Stephenson carefully points out — is it required by a norm of assertion that asserted contents must be true relative to the whole group of discourse participants.

which could be brought to bear to resolve the issue?

These questions arise in part, I think, due to an overly broad reading of the word *faultless*. Faultless disagreement, as this concept was introduced on pp. 4–5, involves two people who assert or believe contents which contradict each other, without either one making an error of fact. But to say that two people faultlessly disagree in this sense (and that they themselves recognize their disagreement as faultless in this sense) does not imply that neither party sees anything wrong at all with the beliefs or assertions of the other. One may regard another person's beliefs or assertions as objectionable — and even *mistaken* — in all sorts of ways which do not involve errors of fact.

In particular, we may recognize "errors of taste" as distinct from errors of fact, and acknowledge that two parties will normally engage in a dispute about a matter of taste only if each of them regards the other as making an error of taste. This in no way represents a retreat from the idea that disagreements over matters of taste are faultless in our original sense, but is simply a clarification of what kind of fault was envisaged.

What is an error of taste? Crucially, this is dependent on perspective: If I believe that roller coasters are fun, and you believe they are not fun, then from my perspective you are making an error of taste, and from your perspective I am making an error of taste. Objectively, there can be no answer, because the error is of taste and not of fact. Putting it more formally:

(251) *x* commits an error of taste in *w* at *t* according to *p* iff *x* asserts or believes Φ in *w* at *t* and Φ(*w*, *p*) = **falsity** and there exists some *p′* such that Φ(*w*, *p′*) = **truth**.

Simply recognizing that people engaged in disputes over matters of taste do regard each other as making errors in some sense will not explain why people make taste assertions, however, or why they persist in disputes over matters of taste when they know that there are no facts which will resolve the issue. Wouldn't it be pointless to argue with someone over matters of taste, if no evidence could be brought to bear on the issue?

In considering this question, we should be careful to recognize that people may make conflicting judgments of taste even when their tastes are not in conflict, simply because one person misjudges whether or not something really is to his or her taste. Suppose, for example, that John tries the chili, finds it unpalatable, and says "Ugh! This is terrible!" If I know that John has just brushed his teeth, I might try to convince him that he is wrong about the tastiness of the chili by pointing out that his flavor perception might be skewed because of his recent toothbrushing. In this case, I am bringing factual evidence to bear on the dispute, but the dispute doesn't really turn on any incompatibility between John's tastes and mine; I am trying to convince John that the chili really does conform to *his* existing standard of taste — not to convince to give up that standard and adopt one more like mine instead. So this is *not*, strictly speaking, a dispute over a matter of taste.

Limiting attention to those cases where a dispute really does turn on incompatible standards of taste advanced by the various parties to the dispute, there can be no objective, factual evidence brought to bear to resolve the disagreement. If we assume that the only rational means of resolving a disagreement is by the evaluation of objective evidence, then such disputes cannot be rationally resolved; and if we assume that it is irrational to enter into a dispute that cannot be rationally resolved, then all disputes over taste are by nature irrational. No doubt it is for this reason that *De gustibus non est disputandum* has the status of a proverb — but in point of fact, normal, reasonable people do enter into disputes over taste quite frequently, and may even devote considerable energy to pressing those disputes. Why?

I think the answer is simply that there is a practical advantage in doing so. If enough of us prevail in arguing that roller coasters are fun, then more roller coasters will be built, and we will have more fun. If enough of us prevail in arguing that licorice is tasty, more licorice will be manufactured for us to enjoy. The more one can advance one's own tastes, convincing others to adopt them, or at least regard them as prestigious or worthy of defense, the more the world will be to one's taste. Because each of us individually — and whole communities collectively — make plans and decisions based in part on our taste judgments, arguing about taste is not at all useless, but is rather a natural part of pursuing one's own self-interest.

"Prevailing" in such disputes cannot mean showing that one's opponent has made some error of fact or logic. The purpose of pressing a dispute over matters of taste is to gain a *social advantage* for one's own tastes: to have them adopted more widely, or to give them priority over the tastes of others in planning and decision-making.[175]

8.4. Questions

Questions present some interesting differences from assertions, particularly with regard to the adoption of stances.[176] As we have seen, truth assessment requires the adoption of a stance — autocentric or exocentric. Because assertion is governed by a norm that the asserted content be **true**, a speaker making an assertion must normally perform a truth assessment, and therefore adopt a stance. Truth assessment is most typically performed from an autocentric stance, with exocentric stances representing a marked option requiring special circumstances. In the default case, then, then a speaker will assert with the intention that the asserted content be true relative to himself or herself; or, as we may put it, assertions are by default made autocentrically.

This pattern appears to be reversed in questions. If, for example, John turns to Mary while they are riding the roller coaster and asks "Is this fun?" he would typically be asking his question from what seems like an *exo*centric stance — he would be inquiring whether *she* found the roller coaster fun, not whether he himself does. This is despite the fact that if he were to make an assertion using the corresponding declarative sentence (*This is fun*) in the same circumstances, he would ordinarily be adopting an autocentric stance, regarding his assertion as warranted because it is true relative to a context in which he himself is the judge.

We can make sense of this effect very straightforwardly if we simply construe questions as *invitations to assert*. This understanding of the illocutionary force of questions fits particularly well with the analysis of the semantic contents of interrogative sentences proposed in Hamblin (1958) and related literature.[177] In this approach, a question is regarded as denoting a partition of the set of possible worlds, so that each cell in the partition corresponds to a complete answer to the question. For example, a simple *yes-no* question like *Does John smoke?* would denote the set {{*w* | John smokes in *w*}, {*w* | John doesn't smoke in *w*}}, with the two cells in the partition corresponding to the two possible answers.

In adapting this approach to our current grammar, we will naturally want to use

[175] This does not, of course, mean that the direct illocutionary act performed in using a taste sentence is one of urging or commending rather than asserting. One can make a direct assertion with the intended perlocutionary effect of spreading a particular standard of taste.

[176] This section repeats a point made in Lasersohn (2005), section 6.1.

[177] See especially Higginbotham and May (1981), Groenendijk and Stokhof (1984).

characteristic functions of sets of world-perspective pairs in place of sets of worlds. Remembering that a perspective is just an individual-time-world triple $\langle x, t, a\rangle$, we let the sentence *Is licorice tasty?* denote the set {[$\lambda\langle w, \langle x, t, a\rangle\rangle$. Licorice is tasty in w by the standards of x at t in a], [$\lambda\langle w, \langle x, t, a\rangle\rangle$. Licorice is not tasty in w by the standards of x at t in a]} — that is, the set whose two members are the two complete answers to the question, namely the contents of the corresponding declarative sentences *Licorice is tasty* and *Licorice is not tasty*.

A full-scale analysis of the syntax and compositional semantics of interrogative sentences in English — even of just those interrogatives corresponding to the limited class of declarative sentences covered by our formal grammar so far — is beyond the scope of this book. We will simply assume that some extension of our grammar can be formulated to generate interrogative sentences and assign them denotations on the model just illustrated.

More important for our current purposes is the way denotations of this sort relate to the illocutionary act of asking. Since each element of the set denoted by an interrogative sentence is the sort of thing which could serve as the content of a declarative sentence, we may regard the act of asking a question as an invitation or request to assert one of these contents. Where Φ is the content of an interrogative sentence, and a speaker x asks an addressee y Φ, we regard x as inviting or requesting y to choose a member of Φ and assert it.

Because assertions are by default made autocentrically, an invitation to assert will normally also be intended and understood as an invitation to assert autocentrically. So if John turns to Mary on the roller coaster and asks “Is this fun?” he is inviting her to answer based on *her* standards of fun, not his.

Of course in a context where it is independently expected that an assertion will be exocentric, questions will also shift their orientation away from the addressee. If Mary asks John “Was the merry-go-round fun?” on his return from a trip to the amusement park with their young son Bill, the question may naturally be understood as a request to assert exocentrically, with John assessing his answer relative to a context with Bill as the value of the judge parameter, rather than himself.

Chapter 9 : Between Fact and Opinion

In this chapter, we consider the possibility that certain perspectives may be ranked as objectively better than others. If we recognize such rankings, then the theoretical possibility exists that certain sentence contents vary in truth value from perspective to perspective, yet also have "objective" truth values with no relativization. Such contents would occupy a middle ground, between fully subjective matters of opinion, and fully objective matters of fact.

I will argue that several classes of sentences should be analyzed this way, including some of the main examples (other than personal taste sentences) for which people have suggested a relativist analysis: contingent futures and epistemic modals.

In this chapter we also consider examples involving aesthetic predicates like *good* or *high quality*, scalar predicates, degree words of sufficiency and excess such as *enough* and *too*, and derogatory epithets such as *jerk* and *asshole*. I argue tentatively that aesthetic predicates fall into the "middle ground" with contingent futures and epistemic modals, but that scalar predicates, *enough*, *too*, and derogatory epithets (other than racial and ethnic slurs) are more like predicates of personal taste.

The wide variety of examples considered in this short chapter will preclude a detailed semantic analysis of any of them. Specific denotation assignments are suggested for future *will* and epistemic modal *may*; but the discussion of all other examples will be kept at a very informal level.

9.1. Aesthetic judgment and refinement of taste

One concern which some readers may have with the idea of faultless disagreement over matters of taste is that it seems to place all standards of taste on the same footing, making no allowance for the possibility that some people have more refined tastes than others, so that their judgments on matters of taste are in some sense more legitimate than the judgments of those whose tastes are less developed. An experienced wine-taster, for example, may base his or her judgment on subtle nuances of flavor that play little or no role in the judgment of a less experienced wine-taster — nuances which one must learn to appreciate, but which, with sufficient education and experience, practically everyone agrees are important to the evaluation of wine. In this case, isn't the more experienced taster's judgment in some sense *really better* than the less experienced taster's?

In considering this question, we must never lose sight of the particular sentences we are giving a semantic analysis of. I would suggest that the more experienced wine-taster's judgment *should* carry more weight in the assessment of sentences like *This is a high quality wine*, or *This is a good wine*. Certainly if I were asked, with regard to some particular wine, whether it were good, or of high quality, I would be content to defer in my judgment to someone who was more expert in such matters. But our examples so far have not involved predicates like *good* or *high quality*; they have involved *tasty* and *fun*. It seems to me that if I were asked whether some particular wine was tasty, I could simply taste the wine and give my judgment; I would be far less likely to defer to anyone else, and far less likely to regard the judgments of others as relevant to my own assessment. *Tasty* privileges the assessor more than *high quality* or *good* do.

But this does not mean that judgments of quality or goodness are purely objective, either. It seems quite possible that two equally sophisticated wine-tasters might differ in their assessments of some wines. Even if we privilege some perspectives over others in assigning truth values to

ascriptions of goodness or quality, some issues of goodness or quality may still ultimately come down to irresolvable differences of personal taste. In giving a semantic analysis, we should allow for the possibility that some disagreements over the applicability of a predicate are faultless, while other disagreements over the applicability of that very same predicate are non-faultless, with one of the parties making an objective error.

Therefore, let us suppose that perspectives may be partially ordered according to their objective superiority in assessing the truth of particular contents. We may write '$p_1 <_\Phi p_2$' to mean that perspective p_2 is superior to perspective p_1 in assessing the truth of Φ — for example, if Φ is the content of some use of *This wine is of high quality*, then '$\langle$John, t_1, $w_1\rangle <_\Phi \langle$Mary, t_2, $w_2\rangle$' means that the perspective of Mary at time t_2 in world w_2 provides superior judgments of that content than the perspective of John at time t_1 in world w_1. Of course we allow for the case where neither of two perspectives is superior to the other (that is, where for some perspectives p, q, neither $p < q$ nor $q < p$), even in cases where they do not provide the same judgments.

Now, suppose that at time t_1 in world w_1, John considers a certain wine to be of high quality. There may be some perspectives ranked higher than $\langle$John, t_1, $w_1\rangle$ which likewise count that wine as high quality; but as we ascend the hierarchy to better and better perspectives, we may reach a perspective p such that all perspectives superior to p count that wine as low quality. In such a case, we may say that John simply has inferior taste, and count the content of *This wine is of high quality* (used to refer to the wine in question) as false, even though John would assess it as true. But if, no matter how high we ascend the hierarchy, we find that some superior perspectives count the wine as being of high quality and others count it as being of low quality, then the issue of whether the wine is of high quality or not seems to come down purely to a matter of taste, and our semantics should count the sentence as neither true nor false *tout court*, but only as true relative certain perspectives and false relative to others.

Let us assume, then, that a sentence like *This wine is of high quality* or *This wine is good* sometimes expresses a content which is true or false "objectively" and sometimes does not, depending on the wine. If the content has its truth value objectively, we may say that it is true or false *simpliciter*, but if it does not, it will be true relative to some perspectives but not others.

This much is compatible with two different positions regarding relative truth for contents which are true or false *simpliciter*. If we retain our definition in (142), then any sentence which is true simpliciter must be true relative to all perspectives, and likewise for falsehood. This means that in the case where John has poor taste, so that the content of *This wine is of high quality* is simply false even though he assesses it to be true, we must analyze that same content as *false* relative to John's perspective, contrary to his own assessment. Alternatively, we could replace (142) with some other condition, which would allow the content of *This wine is of high quality* to be true *simpliciter* even while false relative to some (inferior) perspectives such as John's.

I think the latter alternative provides a better analysis. The primary reason for this is that we must ultimately account for sentences such as (252):

(252) To John, this is a good wine.

The intuitive meaning of this sentence is that the wine is good according to John's standards, with no implication that John's standards are themselves good or that the wine is good *tout court*.[178] We

[178] Quite the opposite — the sentence seems to imply (presumably for Gricean reasons) that John's standards are

may analyze the preposition *to* in such examples on the model of our earlier analysis of *for* in (145):[179]

(253) a. $to_{prep} \in$ (V/V)/N
b. If USE(α, to_{prep}) then $Lex(\alpha, u, w, \langle x, t, a\rangle) = \lambda z \in \mathbf{D}_{u,e}\, \lambda f \in \mathbf{D}_{u,\langle s,t\rangle}\, .\, f(w, \langle z, t, a\rangle)$

This gives the effect that a use of To_{prep} *John, this wine is good* denotes **truth** relative to $u,w,\langle x,t,a\rangle$ iff a corresponding use of *This wine is good* denotes **truth** relative to $u,w,\langle$John,$t,a\rangle$. But now, if we claim that a use of (252) can be true even if John's taste is poor and the wine is bad *tout court*, we must admit that sentences like *This wine is good* can express contents which are true relative to some perspectives while false *tout court*, and we must give up the definition of truth in (142).

What can we replace (142) with? I suggest (254), where p is an *actual* perspective iff for some x, t: $p = \langle x, t, w_@\rangle$:[180]

(254) a. Φ is true iff for all actual perspectives p there exists some actual perspective p' such that $p \leq_\Phi p'$ and for all actual perspectives p'', if $p' \leq_\Phi p''$ then $\Phi(w_@, p'')$ = **truth**;
b. Φ is false iff for all actual perspectives p there exists some actual perspective p' such that $p \leq_\Phi p'$ and for all actual perspectives p'', if $p' \leq_\Phi p''$ then $\Phi(w_@, p'')$ = **falsity**.

More informally: Φ is true iff, in the actual world, as one ascends to better and better perspectives for assessing Φ, one eventually will always reach a point at which the assessments stabilize to **truth**; and false if they stabilize to **falsity**.

This says nothing about which contents will stabilize in this way and which will not. For now we may simply note that if a content does not vary in truth value according to perspective — that is, if it expresses an objective truth or falsehood — then it will obviously have a stable truth value as one ascends the hierarchy of better and better perspectives. And if no perspective is better than any other for assessing a particular content, yet that content varies in truth value according to perspective, then the content is trivially unstable; we may say it deals with a purely subjective matter.

Between these two extremes we find several other possible patterns. Perspectives might be ordered in parallel chains which stabilize to different values, for example, so that for any p, as one ascends to better and better perspectives, one eventually reaches a stable value, but perhaps a different value if one starts from a perspective on one chain than if one starts from a perspective on a different chain. Or it could be that perspectives have a semilattice structure, so that for any p, q, there is guaranteed to be a perspective which is at least as good as both. In this case, if the truth value of a content stabilizes from one starting point, it will stabilize in the same way from all starting points, and we may say there is a single best answer as to its truth value, even though the truth value varies from perspective to perspective. Or, it could be that a content stabilizes from some starting points but not others. In any of these cases, it seems fair to say that the content is neither purely subjective nor purely objective, but something in between.

---

inferior and that the wine is not really good at all.

[179] We add the subscript *prep* to distinguish this use of *to* from the infinitival *to* defined in (203) above.

[180] Here, $\leq_\Phi$ is the reflexive closure of $<_\Phi$ (that is, the relation such that for all perspectives p, it holds that $p \leq_\Phi p$; and if $p \neq q$, then $p \leq_\Phi q$ iff $p <_\Phi q$).

9.2. Contingent futures

The idea of assessment-sensitivity was first developed by MacFarlane (2003), MacFarlane (2008b) for the semantics of the future auxiliary verb *will*: A contingent future sentence like Aristotle's *There will be a sea-battle tomorrow* is treated as truth-valueless relative to a context of assessment whose time falls before the day for which the sea-battle is predicted, but as having a definite truth value relative to contexts whose times fall after that day. This sensitivity to the context of assessment would seem to make the future tense a paradigm case of an expression calling for a relativist semantic analysis.

However, if we accept the view just argued in Section 9.1. that certain perspectives are objectively better than others and that the definition of truth is sensitive to this as in (254), it becomes apparent that future tense does not give as fully subjective an interpretation as personal taste predicates do, but instead occupies a middle ground between fully subjective and fully objective interpretation, similar in some respects to that which we have already seen for expressions like *good* or *high quality*.

It seems intuitively clear that if John says "There will be a sea-battle tomorrow" on October 28, 2014, then dates later than October 29, 2014 provide superior perspectives for evaluating the truth of John's claim than earlier dates do. On this assumption, a truth-definition like (254) will give the result that John's claim is true *simpliciter* or false *simpliciter*, even if it is truth-valueless relative to his own context at the time he makes the assertion.

To work out this idea, let us assign *will* to category (NOM\V)/(ACC\(V∩INF)). We may assign it the following semantics:

(255) If USE(α, *will*), then $Lex(\alpha, u, w, \langle x, t, a\rangle) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle e,\langle i,t\rangle\rangle\rangle} . \lambda y : y \in \mathbf{D}_e$. **truth**, if for all w' such that w' is identical to w up to t, there exists some time $t' \in Relevant_{u,\alpha}$ such that $time_{u,\alpha} < t'$ and $f(w')(y)(t') = 1$; **falsehood**, if for all w' such that w' is identical to w up to t, and all times t' such that $time_{u,\alpha} < t'$, $f(w')(y)(t') = 0$; undefined otherwise]

The idea here is that a use of the sentence *There will be a sea-battle* is true in a world w relative to a perspective p if in every world identical[181] to w up to the time of p, there is a sea-battle at some relevant time later than the time of use; false in w relative to p if in every such world there is no relevant time later than the time of use at which there is a sea-battle; and truth-valueless in w relative to p if in some such worlds there is a sea-battle at a relevant time later than the time of use, but in other such worlds there isn't.

Despite some differences, this repeats major features of the analysis of MacFarlane (2003).[182] Most importantly, it retains the idea that future contingent statements vary in truth value

[181] Some clarification may be called for as to what it means to say that two worlds w and w' are "identical" up to a time t. This notion can be worked out in a variety of different ways: If we take the word *identical* completely seriously, then we may understand the claim that w and w' are identical up to t as meaning that w and w' overlap mereologically, and that the portion of w which lies before t is literally the same thing as the portion of w' which lies before t. If we wish to maintain that worlds are mereologically disjoint, then we might take the claim that w and w' are identical up to t as meaning merely that the same event types occur at the same times in both w and w' up to t. Various other options are no doubt also available. The choice among these will not matter much for our purposes here.

[182] MacFarlane (2003) formalizes his analysis in a "branching times" model, rather than a "worlds-and-times" model

with the context of assessment. Suppose, for example, that in the actual world $w_@$, at 12:00 noon on October 28, 2014, John says "There will be a sea-battle tomorrow." Relative to a context of assessment whose judge is Mary, time is also 12:00 noon on October 28, 2014, and world is also $w_@$, John's statement is truth-valueless, because (let us suppose) there are worlds identical to $w_@$ up to 12:00 noon on October 28, 2014 in which a sea-battle occurs on October 29, 2014, and there are also worlds identical to $w_@$ up to 12:00 noon on October 28, 2014 in which no sea-battle occurs on that date. But relative to a context of assessment whose judge is Mary, whose time is 12:00 noon on October 30, 2014, and whose world is $w_@$, John's statement is false, because there was no sea-battle on October 29, 2014,[183] so there is no world identical with $w_@$ up to 12:00 noon on October 30, 2014 in which a sea-battle occurs on October 29, 2014.

If we were to retain our original definition of truth in (142), this would mean that John's use of the sentence simply has no unrelativized truth value — it is false relative to some perspectives, and truth-valueless relative to others, but not true or false *simpliciter* at all.[184] But let us give up (142) in favor of (254). Let us also assume that for all t later than October 29, 2014, $\langle$Mary, 12:00 noon on October 28, 2014, $w_@\rangle \leq_\Phi \langle$Mary, t, $w_@\rangle$, where Φ is the content of John's use of the sentence *There will be a sea-battle tomorrow*. Now we can derelativize and say that John's statement is false, *tout court* — even though it is not false relative to all perspectives.

MacFarlane, I think, would object to this last move. It gives us an unrelativized truth predicate for future contingent sentences, and does so by assuming a unique actual world. In effect, the analysis assumes that for any given moment m, there is a distinguished actual future from m — the future as it occurs in $w_@$.[185]

The worry behind this objection is that assuming a distinguished actual future seems to give up the assumption that the future is authentically open. Maintaining this assumption was one of the main motivations for a relativist semantics for future sentences in the first place. Whether this worry is well-placed will depend on what, precisely, we mean by an "open" future. We cannot consider in detail here exactly what it means to claim the future is open, but I would point out that the analysis presented here implies does not imply determinism about the future, nor does it imply that non-actual futures are merely epistemically possible, not objectively possible (*contra* MacFarlane (2003)).

On the contrary, (255) appeals directly to the idea that multiple worlds might be identical up to a certain point in time, yet continue differently after that time; and intuitively, this can only happen if events proceed in a *non*-deterministic fashion. If two worlds w, w' are identical up to t, but events proceed differently after t in w than they do in w', there can be nothing about w prior to t which determines that events after t proceed as they do in w, rather than as in w'. In that sense, we

---

as employed here. Points of evaluation are taken to be moment-history pairs. Sentences are assigned truth values relative to pairs of a context of utterance and a context of assessment (informally taken to be equivalent to assigning utterances truth values relative to contexts of assessment). A future tense sentence *Will* φ is true at a moment-history pair $\langle m, h\rangle$ iff there is some m' later than m on h and φ is true at $\langle m', h\rangle$. Any sentence ψ is true relative to a context of use u and context of assessment a iff ψ is true relative to $\langle m, h\rangle$, whenever m is the time of u, h passes through m, and if the time of a is later than m, h passes through the time of a. A non-relativized truth predicate is not defined.

[183] At least I have not been able to find any reports of such a battle.

[184] Of course our monadic but relativistic predicate **false**, defined in (183), can apply to the content of John's assertion, but this is a separate matter.

[185] Following Belnap and Green (1994), MacFarlane characterizes this as the view that there is a "thin red line" marking out the actual future as distinct from other possible futures.

may say that at *t*, facts about the future in *w* have not yet been determined. This is, I think, just another way of saying that at *t* in *w*, the future is open in the relevant sense.

Assuming a distinguished actual future is also compatible with the view that non-actual futures are genuine, objective possibilities, and not "merely epistemic" possibilities. This is perhaps easiest to see if we consider an analogous claim about the present and past. The past and present are not open in the way the future is. We may assume that there is a unique actual past and present, without risking the same level of controversy as we do in assuming a unique actual future. Yet few of us would argue, I think, that there are no genuine, objective, past and present possibilities other than the actual past and present. Alternative pasts are possible in more than an epistemic sense: even things we know did not happen *could have* happened, and it was not necessary for events to unfold in precisely the way they did. I am in Urbana, Illinois as I write this; but this is a contingent matter, not a necessary truth. But if we are prepared to acknowledge genuine, not-merely-epistemic past and present possibilities, despite recognizing a unique actual past and present, then surely we can acknowledge genuine, not-merely-epistemic future possibilities, even if we consider there to be a unique actual future.

This does not put future sentences on a par with purely objective, perspective-independent statements about matters of fact, however. To say that a sentence content is purely objective is to say that it has the same truth value relative to all perspectives:

(256) Φ is objective iff for all w, p, p': $\Phi(w, p) = \Phi(w, p')$

The content of a future contingent sentence will not meet this requirement under the analysis suggested here, because its truth value (in *w*) relative to *p* may be undefined, but its truth value (in *w*) relative to *p′* defined, if the time of *p′* is later than the time of *p*. This variation in truth value from perspective to perspective obtains even though the content may have a definite truth value *tout court*.

As with personal taste sentences, this is not a case of a sentence expressing different contents relative to different perspectives, but rather one of a sentence expressing a content whose truth value varies from perspective to perspective. In this sense, the content is perspectival, not objective. But because it can also be assigned a truth value in the absolute (not just relative to perspectives), based on the superiority of later perspectives over earlier perspectives in assessing the truth of future claims, it is not as fully subjective as the content of a taste sentence. Future contingent claims, like claims of goodness or high quality, occupy a middle ground between purely factual claims and claims of tastiness or fun.

The analysis just sketched faces an apparent problem, once we consider it in the context of the theory of assertion, particularly if we maintain the view argued in Section 8.1. that the act of assertion is governed by a norm that the asserted content be **true**. If contingent future sentences express contents which are neither **true** nor **false** at the time of use, there is no way a speaker could conform to a norm that requires assertions to be **true**. Why, then, aren't all future contingent assertions by their nature abnormal and anomalous? Of course, this puzzle is not peculiar to our analysis, but is general to any analysis which leaves truth value gaps for future contingent sentences, but also regards assertion as governed by a norm of truthfulness.

The solution to this puzzle, I think, does not lie in giving up the idea of a norm of truthfulness, nor in disallowing the truth value gap for future contingents. Rather, we must

recognize that speakers frequently disregard certain possible worlds as pragmatically irrelevant for the purposes of their assertions.

For example, suppose John sees the two fleets making preparations for a battle. He knows they are mutually antagonistic and eager to fight, that the preparations will be complete sometime tomorrow, and that no negotiations are underway which might call a halt to the impending battle. He therefore says "There will be a sea-battle tomorrow." In making this assertion, he disregards various more-or-less remote possibilities which render the battle contingent rather than necessary: He sets aside the possibility that an unexpected storm will separate the fleets, or that the admirals on both sides will have a sudden conversion to pacifism, or that aliens will abduct one of the two fleets, etc. He regards his assertion as justified because he believes that in all the possible futures which are plausible enough to deserve serious consideration, the sea-battle occurs.

There are at least two different strategies we could adopt for working out this idea in more detail. Under one approach, we would revise (255) so that the quantification over worlds is restricted to some limited set of worlds which are relevant to α (the particular use of *will* being interpreted). Thus, a use of the sentence *There will be a sea-battle* will be true relative to a world *w* and perspective *p* iff in all worlds *relevant to this use* which are identical to *w* up to the time of *p* there is a sea-battle at some relevant time later than the time of use, and false iff in all such worlds there is no relevant time later than the time of use at which there is a sea-battle, and truth valueless iff — even limiting attention to the pragmatically relevant worlds — in some of them there is a sea-battle and in some there isn't.

One attractive feature of this approach is that it fits naturally with existing theories of the semantics of modal auxiliary verbs. Morphosyntactically, *will* is a modal auxiliary verb, not a tense marker like the past tense suffix *–ed*.[186] Existing analyses of modal auxiliary verbs frequently treat them as expressing restricted quantification over worlds, with the restriction sometimes expressed overtly (for example via an *if*-clause) and sometimes determined pragmatically. (See especially Kratzer (1977), Kratzer (1981) and related work.) We might therefore reasonably expect that *will* might also express pragmatically restricted quantification over worlds, in much the same way.

A second, alternative approach would not revise (255) at all. On this approach we would instead recall the point made in Section 8.1. above that the norm of truthfulness is not the only norm governing the act of assertion, and that violations of truthfulness are sometimes pragmatically licensed in order to avoid more serious violations of other norms. That is, the mere claim that assertion is governed by a norm of truthfulness does not by itself amount to a claim that untrue contents cannot be felicitously asserted. Certain kinds of falsehood are pragmatically irrelevant, and may be disregarded: Just as one may in some contexts say "John arrived at 3:00" even though he arrived at 3:02, one may also in some contexts say "There will be a sea-battle tomorrow" even though there is a remote possibility that a freak storm will separate the fleets and prevent the battle from happening. In this case the assertion is neither **true** nor **false** at the time of use, but comes "close enough to **true** for its context" (to adapt the terminology of Lasersohn (1999)).

The two approaches differ in whether the disregard of pragmatically irrelevant worlds affects the truth value of future contingent assertions. On the first approach it does, and an

[186] English is not unusual in this regard. It is quite common cross-linguistically for future markers to pattern morphosyntactically more like mood or modality markers than like past or present tense markers.

assertion that there will be a sea-battle tomorrow can be **true** at the time of utterance even if there exists a remote, pragmatically irrelevant possibility of the sea-battle not occurring. Even though we allow future contingent sentences in principle to express contents which are neither **true** nor **false** at the time of use, restrictions on quantification eliminate the truth value gap much of the time. On the second approach, disregarding irrelevant worlds does not affect truth conditions, and an assertion that there will be a sea-battle tomorrow is neither **true** nor **false** at the time of use, but is still pragmatically licensed, because the possibilities which prevent it from being **true** are pragmatically ignorable.

The choice between these two approaches does not matter much for our present purposes. (My own suspicions favor the second approach, but I do not have compelling arguments either way.) What matters more is that a semantic analysis of the future auxiliary *will* along the lines of (255) is not incompatible with the view that assertion is governed by a norm that the asserted content be **true**.

9.3. Epistemic modality

Epistemic readings of modal auxiliary verbs such as *may* and *must* present a similar case to contingent future *will*. Certain considerations suggest that sentences containing epistemic modals express contents whose truth values vary from perspective to perspective; yet some perspectives seem inherently better than others for evaluating the truth of such contents. We may therefore assign such contents "absolute" truth values in addition to relativized truth values, despite the perspectival variation.

The need for a relativist analyses of epistemic modals has been argued in Egan (2007), Egan, Hawthorne, and Weatherson (2005), MacFarlane (2008a) and elsewhere. The primary motivation for a relativist analysis is that claims made with epistemic modals intuitively seem appropriately assessed as true or false according to the knowledge state of the assessor.

For example, the relevant reading of (257) is roughly paraphrased as "It is consistent with what is known that John is in New Orleans":

(257) John may be in New Orleans.

But if we accept that the sentence has such a reading, in which its truth value depends in part on "what is known," we immediately face the question, *known by whom?*

It seems problematic to claim that the truth of the statement depends specifically on what is known by the *speaker*. This would render the sentence equivalent to something like (258):

(258) It is consistent with what I know that John is in New Orleans.

But if Mary asserts (257) and I happen to know that John is not, in fact, in New Orleans, I can disagree with her by asserting (259):

(259) No, John is in Baton Rouge.

However, the claim that John is in Baton Rouge is perfectly compatible with the claim that it is consistent with what the speaker of (258) knows that John is in New Orleans. That is, (259)

conflicts with (257) but not (258), so (258) cannot be an accurate paraphrase of (257).

We might attempt to remedy this problem by claiming that it is not specifically the *speaker's* knowledge which matters to the truth value of (257), but the pooled knowledge of all the participants in the conversation in which (257) occurs. Then if I am a participant of a conversation in which Mary asserts (257), but I happen to know that John is in Baton Rouge, it is not consistent with the pooled knowledge of all the participants in the conversation that John is in New Orleans, and I would be contradicting Mary by asserting (259).

Even this revision fails, however, once we recognize that even people who are not participants in the conversation where (257) occurs may disagree with it or contradict it, if they know that John is not in New Orleans (at the time of use of (257)). An eavesdropper on the conversation, for example, even while not a participant in the usual sense, may disagree with Mary as she asserts (257), as may a person listening to a recording of the conversation later.

It appears, in fact, that anyone who knows that John is not in New Orleans is in disagreement with (257), regardless of whether they are connected to the conversation or the context of use in any way. It seems appropriate, therefore, to analyze (257) as expressing a content which has different truth values relative to different contexts of assessment, depending on the knowledge state of the judge — that is, to give epistemic *may* a relativist analysis.

We may adapt this idea to our current grammar by assigning *may* (like *will*) to category (NOM\V)/(ACC\(V∩INF)) and defining it as follows, where $\mathbf{K}_{x,t,w}$ is the set of worlds compatible with what *x* knows at time *t* in world *w*:

(260) If USE(α, *may*), then $Lex(\alpha, u, w, \langle x, t, a\rangle) = [\lambda f : f \in \mathbf{D}_{\langle s,\langle e,\langle i,t\rangle\rangle\rangle} . \lambda y : y \in \mathbf{D}_e . \exists w' \in \mathbf{K}_{x,t,a} f(w')(y)(time_{u,\alpha}) = \mathbf{truth}]$

This far, the analysis is much like that of predicates of personal taste. A use of *John may be in New Orleans* is true relative to a context of assessment in which I am the judge, the time is 5:00 p.m. on February 3, 2015, and the world is w_1 iff it is consistent with what I know at that time in that world that John is in New Orleans at the time of use; it is true relative to a context of assessment in which you are the judge, the time is 2:30 a.m. on March 1, 2014 and the world is w_2 iff it is consistent with what you know at that time in that world that John is in New Orleans at the time of use. The content is the same in either case. There is no world-perspective pair at which both this use and a use of the syntactic negation of this sentence (with the values of the indexicals set identically) are true, so these uses contradict each other.

However, it seems intuitively clear that if I know John's location and you do not, then I am in a better position to assess the truth of *John may be in New Orleans*. More precisely, if two perspectives have the same world, but at the time of the first perspective, its judge knows John's location at the time of use, and at the time of the second perspective, its judge does not, then the first perspective is superior for the assessment of the content of that use than the second perspective: Where Φ is the content of a use of *John may be in New Orleans*, $\langle x, t, w\rangle <_\Phi \langle x', t', w\rangle$ if in *w* at *t′*, *x′* knows John's location at time of use, but in *w* at *t*, *x* does not.

But now it must be the case that the truth value of this content stabilizes as one ascends from worse to better perspectives.[187] If John is in New Orleans (at the time of use in $w_@$), Φ is true

[187] We must make a couple obvious additional assumptions to guarantee this stability: that it is a matter of objective fact, not taste, whether John is in New Orleans, and that false contents cannot be known.

relative to the actual perspectives of all individuals who know his location; if he isn't, it is false relative to all such perspectives. Under (254), then, we may say that Φ is true or false *simpliciter*, according as John is in New Orleans at the time of use or not.

If we restrict attention to "absolute" truth and falsity, then, *John may be in New Orleans* seems to have the same truth conditions as *John is in New Orleans*. It should be carefully noted, however, that these two sentences do *not* express identical contents under the analysis suggested here. The difference isn't in the conditions under which they are true or false *tout court*, but in the conditions under which they are true or false relative to perspectives: the content expressed by uses of *John is in New Orleans* does not vary in truth value from perspective to perspective, while the content expressed by uses of *John may be in New Orleans* does. This is, I maintain, a difference in truth conditions; but it is a difference which is only apparent in a theory which appeals to a relativized assignment of truth values, and not (just) an "absolute" assignment.

Because the norm for truthfulness of assertion requires merely that the asserted content be **true**, not that it be true, a speaker may conform to this norm (as assessed from his or her own perspective), even while asserting a content which is false *simpliciter*. This seems correct; a speaker who asserts that John *may* be in New Orleans cannot be held responsible if it is discovered later that John was not in New Orleans. In that sense, an assertion that John may be in New Orleans may be "faultless" even if it turns out that John is not in New Orleans, so that the assertion is false. If someone disagrees with such an assertion by pointing out that John is not in New Orleans, the disagreement may therefore also be seen as "faultless" in some sense. But this is noticeably different from the case of disagreement over matters of taste. Even though epistemic modal sentences express contents which vary in truth value from perspective to perspective, some of these perspectives may be seen as objectively superior to others for the purposes of assessing those contents. Epistemic modal sentences, like future contingents, therefore occupy a middle ground, between sentences like those expressing personal taste, for which no unrelativized truth values can be assigned, and purely objective sentences dealing with matters of fact, which do not show any variation in truth value according to perspective.

9.4. Non-taste candidates for relativism: scalar cut-offs, sufficiency, and derogation

I have suggested that aesthetic terms like *good* and *high quality*, contingent futures, and epistemic modals all occupy a middle ground between relativistic terms like predicates of personal taste, and fully objective expressions which express contents whose denotations do not vary with perspective. We may well ask whether personal taste expressions are unique in their degree of relativism, or if there are other kinds of expressions which give rise to similar effects.

One class of examples which appears to me to be a good candidate for a fully relativistic analysis are *scalar predicates*. By this I mean predicates such as *tall* or *rich*, which in some way relate their arguments to degrees on a scale: John is tall (or rich) to a certain degree, Mary to a different degree, and so on. The semantic effect of scalarity is analogous in some ways to that of sensitivity to personal taste, but also shows some important differences. Many predicates — including our paradigmatic examples of predicates of personal taste, *fun* and *tasty* — are *both* scalar *and* predicates of taste.

We cannot give a full-scale investigation of scalarity here, but a brief and simplified sketch is in order. Sometimes, scalar expressions appear with a *degree phrase*, as in (261)a., where *tall*

appears with *six feet*; and sometimes they appear without any degree phrase, as in (261)b.:[188]

(261) a. John is six feet tall.
b. John is tall.

The relevant cases for our purposes here will be those with no degree phrase, such as (261)b.

If a sentence like (261)b. is used, at least two contextual parameters associated with *tall* may need to be fixed before we can assign a definite truth value. The first of these parameters is a *comparison class*. Sentence (261)b. may say something true in a context where we are talking about jockeys, but false in a context where we are talking about basketball players, even if John's height is the same in both contexts. The second parameter is a *cut-off* for the relevant comparison class, such that members of that class whose height is above the cut-off are considered tall, and those whose height falls below are considered not tall.

The situation is complicated by the fact that sentences like (261)b. are often used (and interpreted, and assessed for truth and falsity) without the speaker (or interpreter, or assessor) ever precisely identifying an exact cut-off *point*. Indeed, the intuition is strong that no precise boundary exists between the tall and the not-tall, even after all aspects of the contexts of use and assessment for the word *tall* are taken fully into account. This is the classic problem of "vagueness," and it is well beyond the scope of this book to try to solve it (or attendant problems such as the sorites paradox).

Nonetheless, certain aspects of this parameterization are very relevant to our present concerns. Even if no definite cut-off point is established, the cut-off is certainly *constrained* contextually, and there is an interesting contrast between the way context constrains the cut-off and the way it constrains the choice of comparison class. The comparison class appears to be constrained only by the context of use,[189] but the cut-off is constrained at least in part by the context of assessment. The result is that faultless disagreement may arise as a result of differences as to the location of the cut-off.

The possibility of faultless disagreement as a result of differences as to the location of the cut-off for a scalar predicate was pointed out (though with a somewhat different terminology and theoretical development) by Richard (2004), Richard (2008). As Richard notes, two people may disagree as to the truth of (262), even if both are fully informed how much money Mary has (a million dollars, say), simply because they have different views about just how much money someone must have in order to qualify as rich:

(262) Mary is rich.

The possibility of disagreement persists, even if we hold the comparison class constant. Both parties to the disagreement may take the comparison class to be the set of New Yorkers, for

[188] Scalar expressions vary in the kinds of degree phrases they accept, sometimes for apparently arbitrary reasons. For example, *?John is 150 pounds heavy* is much less natural than *John is six feet tall*, even though heaviness is measured in pounds. However, *heavy* does combine naturally with degree words such as *too*, *enough*, *very*, etc.

[189] If a speaker calls John tall in the context of a conversation about jockeys, and another calls him not tall in the context of a conversation about basketball players, we need not take the two speakers to be disagreeing with each other; the content of *tall* is different in the two contexts.

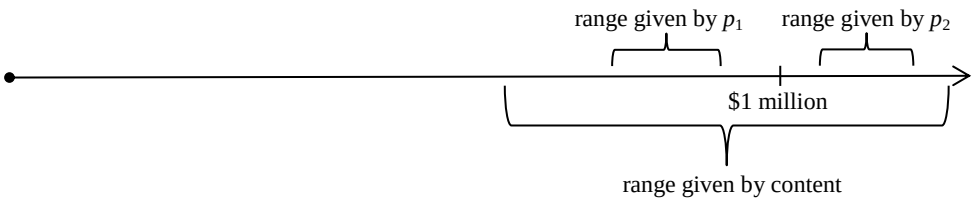
example, yet still differ as to how much a New Yorker must have in order to qualify as rich.[190] We may even suppose that both parties are fully informed as to how much money the average New Yorker has — or the median, etc. — in addition to how much money Mary has; the possibility of disagreement still persists. And as with matters of taste, disagreement is possible regardless of whether the parties to the disagreement are engaged in conversation with one another, or share a common language, or are even aware of each other's existence.

These considerations suggest that the location of the cut-off between the rich and the not-rich (in a given comparison class) is constrained by the context of assessment. Different assessors may constrain it to fall in different locations on the scale, giving rise to perspectival variations in truth value for the same content.

Moreover, it seems intuitively that two perspectives may differ from one another as to where they constrain the cut-off to fall, without either perspective having any legitimate claim to objective superiority over the other. In such a case, disagreements which turn on this difference in where to place the cut-off are fully "faultless," much like disagreements which turn on differences in personal taste.

Therefore, let us assume that the content of *rich* (in a given context of use) leaves open a range of possible cut-off locations, but rules others out. A given perspective likewise leaves open a range of possible cut-off locations and rules others out. We assume that the range given by a perspective must be included in the range left open by the content. An example is illustrated in (263): a person whose perspective is p_1 disagrees faultlessly with a person whose perspective is p_2 over the truth of *Mary is rich*, where Mary has one million dollars:

(263)



We count a use of *Mary is rich* true relative to a given perspective *p* if Mary's wealth falls above the range of possible cut-offs given by *p*, and false relative to *p* if it falls below this range. (We do not consider here what its truth value should be if it falls within the range.[191])

Formalizing this idea would require a more detailed position on the analysis of vagueness than is possible here, but the fundamental point of this section, that sentences whose truth value depends on the placement of scalar cut-offs provide plausible candidates for as fully relativistic a semantic analysis as predicates of personal taste, seems compatible with a variety of more detailed possible developments of the analysis.

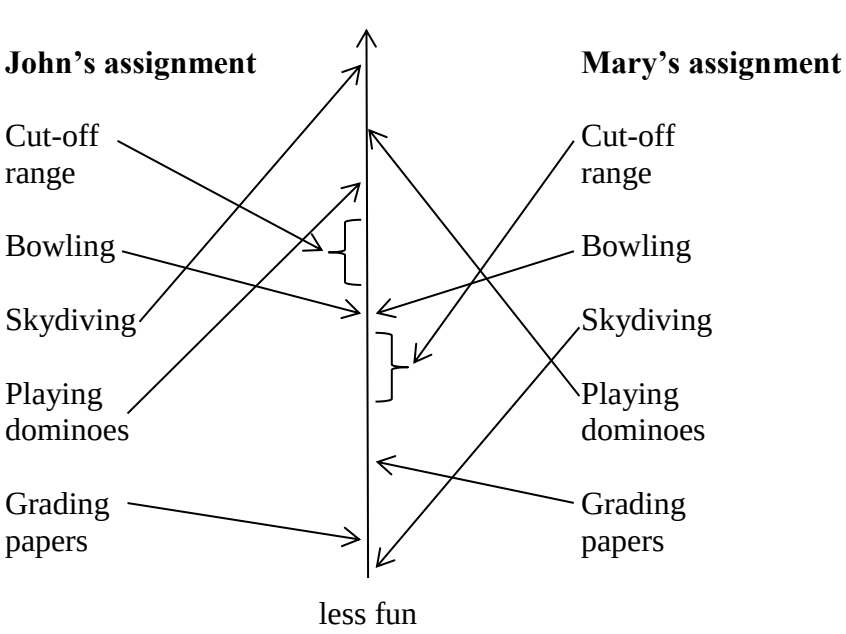
Even though scalarity and sensitivity to taste can both give rise to faultless disagreement, they do so for different reasons. Faultless disagreement due to scalarity arises because of personal differences in where to place the cut-off. What seems crucial for disagreements over taste

[190] As Richard points out, we may also hold constant the practical purpose for which the classification of New Yorkers as rich or not rich is made.

[191] Here we might add another layer of relativization, to "precisifications" as in supervaluationist theories of vagueness, or any other technique which appears well-motivated and compatible with the rest of our analysis.

(involving scalar predicates) is not the location of the cut-off, but the initial assignment of degrees. Different people may assign markedly different degrees of fun or tastiness to the same items, and may differ radically in the relative order of these items on the fun or tastiness scale; but no objective "matter of fact" would seem to select any one of these assignments or orderings as the correct one. John and Mary may disagree whether skydiving is fun, not because they both realize it is fun to degree *d* and differ as to whether *d* is sufficiently high to count as fun, but because John (who enjoys a good thrill) assigns it a high degree of fun, while Mary (who is terrified of falling) does the opposite. This is illustrated in (264).[192] Here, John and Mary agree that playing dominoes is fun and grading papers is not fun; they disagree about whether skydiving is fun and about whether bowling is fun — but for different reasons.

(264)



The degree of fun or tastiness which one assigns to an item seems tightly bound up with affect and experience. It is presumably for that reason that we have a kind of epistemic privilege in making taste claims, as discussed above in Section 2.3. But the placement of the cut-off seems less so; experiencing a game of bowling is enough to tell you *how* fun that game was, but deciding whether that degree of fun is enough to qualify as (just plain) fun involves an additional step, and does not follow as an automatic consequence of that experience. Because we need not fix an exact degree as the cut-off in order to use the language appropriately, we may feel hesitation in asserting or assessing sentences which ascribe a scalar predicate to an object which falls somewhere in the range where the cut-off may be located (according to our own perspectives). Scalarity and sensitivity to taste thus give rise to different patterns of confidence and hesitation in assessment and assertion, despite both calling for a relativist analysis.

A related class of examples which seem to be good candidates for a relativist analysis

[192] The diagram in (264) is adapted from one given in Lasersohn (2008).

involve a certain reading of degree words expressing sufficiency or excess, such as *enough* or *too*, as in (265):

(265) a. This food is warm enough.
b. This food is too warm.

The relevant reading of *enough* and *too* must be carefully distinguished from other readings which do not require a relativist analysis. For example, *enough* and *too* often combine with an infinitival clause or some other phrase indicating a goal or purpose:

(266) a. This food is warm enough to serve without risk from bacterial contamination.
b. This food is too warm to eat without burning my mouth.

Examples like these do not call for a relativist treatment; they express contents which are true or false absolutely, with no perspectival variation in truth value.

In some cases, *enough* or *too* appears without a purpose phrase, but the context of use makes a particular goal or purpose salient:

(267) *John:* Can we serve this food without risk from bacterial contamination?
Mary: Yes, it is warm enough.

Mary's assertion in (267) is equivalent to (266)a., and therefore should not receive a relativist analysis, despite the absence of an overt purpose phrase.

But sentences like those in (265) can also be used in contexts where the salient goal or purpose is not one for which it can be determined objectively whether that goal has been met. For example, if John prefers piping hot food and Mary prefers cooler food, they may disagree whether some particular portion of food is warm enough to eat. Mary asserts *The food is warm enough* (or *The food is warm enough to eat*), and John asserts *The food is not warm enough* (or *The food is not warm enough to eat*). The intuition is strong that they contradict each other, but there seems to be no factual basis for deciding who is right. They are not disagreeing about whether the food is warm enough for *John* to eat, or for *Mary* to eat, or for people in general to eat, just whether it is warm enough to eat — and this seems to come down to a matter of personal taste.

It appears, then, that expressions like *too* or *enough* (or at least an implicit goal phrase with which these words are associated) can introduce a sensitivity to taste. Although it is not *prima facie* obvious that these are "expressions of personal taste" the way it is with *tasty* and *fun*, neither is it clear that they represent a completely separate case, providing independent evidence for a relativist semantic theory.

Perhaps the same thing can be said of derogatory epithets such as *jerk* or *asshole*. What are the truth conditions of a sentence like *John is an asshole*? They seem roughly the same as those of *John is obnoxious*, though the former sentence packs a stronger emotional punch, and also exhibits some obvious stylistic differences. But there seems to be no objective way to classify people as obnoxious, let alone assholes. Some of us consider certain behaviors to be asinine which others find unobjectionable or even admire. Some people view even the nicest, most considerate people as assholes, precisely because they are *too* nice — excessive niceness is objectionable by some

lights. If a person with one set of tastes says that John is an asshole, and a person with conflicting tastes says he is not, we take them to be disagreeing. But ultimately, who counts as an asshole seems to depend on our tastes in human behavior. Such epithets are in some sense, expressions of personal taste.[193]

9.5. Conclusion

In this chapter we examined a variety of examples which arguably call for a relativist analysis. However, it was suggested that some of these — involving predicates of aesthetic quality, contingent futures, or epistemic modals — fall short of the degree of relativism exhibited by predicates of personal taste. Instead, they occupy a kind of middle ground, allowing for objective truth values with no relativization, alongside a relativized assignment of truth values that varies according to perspective. Other examples — involving scalar cut-offs, *too*, *enough*, and derogatory epithets — seem like better candidates for a fully relativist analysis. But several of these, on closer inspection, seem to involve a sensitivity to taste, rather than representing a completely separate phenomenon.

Given the brevity of this chapter and relative informality of the discussion, any conclusions must be regarded as highly tentative. But the initial indications are that a fully relativist analysis, in which truth values vary by perspective and no unrelativized truth values are assigned, should be limited to sentences dealing with matters of taste or scalarity.

[193] I would *not* extend this claim to derogatory epithets which are reserved for people belonging to particular racial or ethnic groups, such as *kike* or *nigger*. Any perspective which yields a truth value for assertions that someone belongs to one of these categories must be regarded as objectively inferior for the assessment of such assertions than perspectives which do not.

Chapter 10 : Reliability, Imagination, and the Functional Motivation for Relativism

I have argued that personal taste sentences express contents whose truth values vary with "perspectives" in addition to worlds, and that when such a content is assessed for truth and falsity, it is the context of assessment, not the context in which the sentence was used, which provides the relevant perspective for assigning a truth value. This kind of variation in truth value contrasts with that produced by indexical expressions, which involves a variation in content, not just in the assignment of truth values to contents, with the relevant values are fixed by the context of use, not the context of assessment.

One may well wonder *why* a language would be structured this way, with two distinct kinds of contextual variability. It is not hard to imagine reasons why a language might employ expressions whose interpretation depends on context — for example, to increase "efficiency" as Barwise and Perry (1983) have emphasized. But why would it employ (in some cases but not others) a notion of relative truth that is ineliminable in terms of absolute truth? Isn't a semantic system which uses a relativistic notion of truth in some sense more "marked" or complicated? Isn't it strange and unexpected — something that requires special explanation, in a way that ordinary unrelativized truth does not?

As theorists, certainly we should require compelling arguments before giving up a theory that has served us well — hence before abandoning more conventional semantic theories to adopt ones incorporating irreducibly relative truth. I think the recent literature on relativist semantics has made a good start in developing such arguments, and hope this book as strengthened the case. But even if *we* as theorists are starting from a social or historical position where unrelativized truth can be assumed as a default in linguistic analysis — so that *we* need special justification before moving to a system with relativized truth in developing such an analysis — this does not mean that unrelativized truth is *linguistically* or *cognitively more primitive*, or that a language comes "by default" with a non-relativized notion of truth, and needs some kind of special "push" in order to adopt a relativized notion of truth. In other words, we should not assume that linguistic evolution or markedness relations recapitulate the historical progression of semantic theorizing.

In this chapter I will argue that a relativized notion of truth is expected, if we think of truth as relating to a more primitive concept of *reliability*, and if we consider how reliability behaves under imagined permutations of location. Consideration of reliability under such permutations also gives us clues as to which kinds of contextually-sensitive expression will require a relativist analysis, and which require an indexical analysis.

In order to explore these ideas, we will take a highly speculative look at language evolution. We will also develop a notion of "logical space," broadly in the tradition of Wittgenstein (1961), van Fraasen (1967), Thomason (1979), and set up formalism for representing permutations of position in locative space and in logical space.

10.1. An evolutionary fable

As a preliminary to considering why a language might develop a notion of relative truth, it is worth considering how and why *any* notion of truth might develop. That is, for what kinds of representational or communicative systems is a notion of *truth* appropriate, and what motivates the

use of such systems, as opposed to other kinds of representational or communicative systems which do not make use of a notion of truth? I will approach this question in evolutionary terms, considering the first the needs of cognitively simple organisms with simple representational systems, and considering in turn various kinds of alteration to such simple systems that come with increasing cognitive sophistication. I do not intend this as a serious scientific account of how language or other representational systems evolved, and caution readers not to interpret it that way. The point, rather, is simply to consider at an abstract conceptual level certain ways in which languages differ from other possible representational systems, in order to highlight a reason why languages might use a relativized assignment of truth to contents for certain expressions, but a relativized assignment of contents to expressions for others.

Consider first an organism, cognitively much more primitive than ourselves, which must traverse its environment and interact with the things it encounters. Presumably, there would be an advantage to the organism to develop some system of internal representations of this environment, to store information, aid in decision-making, etc. This system need not be very language-like; it might be more map-like or image-like, for example. In this case it would not be appropriate to consider whether representations in this system are "true" or not.

However, it might still be the case that the organism sometimes misrepresents its environment to itself, and therefore "makes mistakes" in its interaction with the environment — mistakes which reduce its chances of survival or reproduction. It would therefore be to the organism's advantage to develop a system for revising its own internal representations.

Such a system must include a way of identifying certain representations (or parts of representations) as "unreliable" and in need of revision, and others as (presumed to be) reliable, and not in need of revision. *Reliability* is not the same as truth, or even the same as accuracy. One can imagine representations which misrepresent the environment in ways which are not harmful to the organism — in this case the organism could *rely* on the representation, despite the inaccuracy.

We as theorists can distinguish between *egocentric* representations, as when the organism represents to itself that a wolf is nearby, and *geocentric* representations, as when it represents to itself that a wolf is near the river. But the notion of *reliability* does not appeal to the distinction between egocentric and non-egocentric representations, or require an organism to determine whether a representation is egocentric or not in order to mark it as unreliable or reliable. There is no reason to think that in evaluating the reliability of an egocentric representation, the organism must mentally fix a contextual parameter to itself. This would only be necessary if more than one option were available for how to fix the parameter — but *reliable* here just means something like "works for me" or "nothing goes wrong if I use this." Even if an organism develops both egocentric and geocentric representations, this does not, by itself, mean that the organism must "keep track" of which representations are egocentric and which are not, or use a different procedure for evaluating them for reliability.

Now consider how such a representational system might develop further — in particular how it might provide the basis for a *communication system*. It would be to the advantage of the organism to be able to communicate with others, and share information about food, predators, etc.

We should note that behavior which is an automatic response to environmental stimulus can convey information to other organisms which observe it, and is therefore communicative, but may not require the mediation of internal representations. However, more complex and

sophisticated forms of communication presumably do express (or correspond to) internal representations in some way. Communication which is not an automatic response, or which is about absent objects, for example, would seem to work this way.

Consider communication from the standpoint of an organism which *receives* a signal from another organism. The recipient might make use of the communicated information by responding to it in a purely automatic fashion, but there would be an advantage to developing some system for classifying communications as reliable or unreliable, just as with internal representations. Here, "reliable" means that the *recipient* can rely on the communication.

At this point, there is still nothing which requires the organism to have any notion of perspective, or to do any mental "filling in" of contextual parameters in interpreting or evaluating either its own representations, or the communications of others. However, it is natural to suppose that an organism might do a better job of deciding whether the communications of others can be relied on if it developed the ability to represent their perspectives to itself, and recognize that a representation which is reliable from one individual's perspective need not be reliable from another's. For example, a signal used when a wolf is nearby allows a different response when received from someone distant than it does when received from someone nearby. At a more sophisticated level, communications about what is to the right or to the left call for different responses when received from someone facing you than when received from someone facing the same direction as you.

How could the ability to assess the reliability of communications relative to others develop? Presumably, by building on the pre-existing capacity to assess reliability in general, including reliability for egocentric representations. I would suggest, in fact that the natural first step in developing such an ability is *to imagine oneself in the place of the individual producing the communication, and assess it relative to oneself, in this imagined circumstance*. This style of assessment, which we may term *semi-exocentric*, is in some sense intermediate between fully autocentric and fully exocentric assessment.

What is important to note for our current purposes is that *different perspective-dependent communications behave differently when assessed in this way*. As one imagines oneself in different positions, one's judgments about the reliability of communications of spatial distance or orientation shift to match the judgments of individuals who are actually in those positions. But judgments about the reliability of communications of personal taste will not shift in this same way.

For example, suppose some individual B spots a wolf, and issues a signal, received by a different individual A, that a wolf is nearby. If A imagines herself in B's position, identifies the mental representation expressed by the signal, and assesses it for reliability in this imagined circumstance, she should come to the same assessment as B. The same effect occurs with communications about one object being to the left of another, etc.: If you issue a signal associated with the sort of mental representation you have when something is on your left, and I receive the signal, imagine myself in your position, and assess for reliability, I should reach the same assessment as you.

However, if you are eating something, and issue a signal associated with the kind of representation you have when what you are eating is tasty, and I receive this signal and imagine myself in your position, I might well not reach the same assessment as you. Particularly if I already have experience with the kind of thing you are eating, and know that it is not palatable to me, I might judge your signal to be unreliable, even in the imagined circumstance where I am in your

position.

I suggest that this difference in how communications are assessed for reliability when one imagines oneself in the place of another leads naturally to differences in truth conditions and truth-conditional content, in a system where these concepts are applicable, because *truth* is an idealization of *reliability*.[194]

What do we mean by imagining oneself in a different "position"? This cannot just mean imagining oneself in a different spatial location, or it would be trivial and unsurprising that non-spatial expressions like *tasty* do not vary in reliability with shifts in position. Nor could it mean imagining that one *is* another person, complete with that person's tastes, dispositions, knowledge, etc. — or then *tasty* would vary in reliability with shifts in position just like *nearby* or *left*.[195]

I would suggest that for our purposes, imagining oneself in the position of someone who issues a signal means imagining that one has all the same "outward" or "extrinsic" properties that led the signaler to issue the signal — spatial location, engagement in activities (such as tasting food), etc., but retaining (to the extent compatible with this) all of one's "intrinsic" properties — knowledge, dispositions, preferences, etc. Retaining these is what makes it *imagining oneself in someone else's position*, rather than *imagining that one is someone else*. We should have no illusions about the difficulties involved in making this distinction between "outward" and "intrinsic" properties precise, but this does not mean that there is no distinction to be drawn. Most of what follows is compatible with various different ways one might draw this distinction, and with the idea that it may be subject to pragmatic and/or personal variation.

I suggest that *patterns of reliability assessment are "echoed" in patterns of truth conditions*. This is expected if we regard truth as an idealization of reliability. Idealizing reliability to develop a notion of truth involves at least three shifts: First, whether a representation is *reliable* for an individual *x* is strictly a matter of what the practical consequences would be for *x* of using that representation. But beneficial consequences would normally be maximized by ensuring a "tight," systematic correlation between representations and types of circumstance in which the individual might find itself. Truth differs from reliability in abstracting away from practical consequence of representation use, and depending instead directly on whether this systematic correlation obtains. Second, reliability is scalar — a representation can be more reliable or less — but truth is (I assume) bivalent. Third, a representation can be reliable for certain purposes, but not others; but truth is (I assume) not purpose-sensitive.[196] There may be additional shifts as one passes from reliability to truth — I do not claim this to be a comprehensive list. But aside from

[194] A similar appeal to differences in assessment "as if to oneself" is central to the semantic analysis of personal taste sentences given by Moltmann (2010), but she develops the idea rather differently, and draws very different conclusions about the contents of such sentences.

[195] We would also get the incorrect result that someone who knew there was no wolf near the signaler, but who realized that he or she would also be fooled by the evidence available to the signaler, would judge the signal to be reliable.

[196] By this I mean only that semantic theory should not count a given sentence-content as true for some purposes but false for others. This is compatible with the view that a given sentence may express different contents on different occasions, according to the purpose for which the assertion is made, and with the view that a sentence-content may reference the interests, purposes and intentions of the individuals it is about.

differences forced by such shifts, judgments of reliability should "persevere" as judgments of truth.

In particular, if one's assessment of an expression as reliable or unreliable shifts as one imagines oneself in another's place, this will translate into an assignment of truth conditions which depends on context. For example, a sentence like *A wolf is nearby* should be assigned truth conditions which render it true relative to one context but false relative to another; likewise for sentences placing objects to the left or right.

Moreover, this variation in truth value according to context should correspond to variation in content according to context. There is an obvious advantage to keeping one's own judgments consistent with one another. If you judge *A wolf is nearby* to be false in your current context, but imagine yourself in someone else's place and judge it to be true in this imagined context, a way must be found of reconciling these judgments. This can be done by assigning different contents to the sentence in the two contexts — contents which do not conflict with one another. In this way, an indexical interpretation naturally arises from the fact that one's own truth-value judgments vary under imagined changes of position.

But if your judgment of whether *Spinach is tasty* is reliable or true does not shift as you imagine yourself in someone else's place, this poses no threat to the consistency of your judgments, and does not provide the same pressure to assign different contents relative to these contexts. So, sentences like this do not vary in content from context to context.

Completely objective sentences like *Spinach is a plant* will likewise not shift in reliability under imagined changes of position; so far, our fable does not distinguish between such sentences and personal taste sentences like *Spinach is tasty*. Drawing this distinction requires one more step, the development of fully exocentric assessment.

This last step has been taken when one recognizes that the claim that spinach is tasty might be reliable for someone else — in the sense that that other individual suffers no unpleasant consequences by acting on it — even if it is completely unreliable for oneself, even when one imagines oneself in the other's place. This pattern of reliability can persevere into a pattern of truth conditions, so that *Spinach is tasty* is true relative to one person and false relative to another. Recognizing this pattern of reliability requires a capacity for assessing a representation relative to someone else that goes beyond simply imagining oneself in that person's place, and then assessing relative to oneself; in other words, a capacity for fully exocentric assessment, as opposed to the intermediate style of assessment discussed so far, in which one imagines oneself in someone else's place, then assesses autocentrically. Because this strategy does not carry implications for how *you* would assess a representation, if *you* were in another's place, the fact that it yields different truth values relative to different individuals poses no threat to the consistency of your judgments,[197] so it still does not require that content vary with context.

The result of this last step is a semantic system in which statements of personal taste vary in truth value from person to person, without a corresponding contextual variation in content, and in which statements of spatial distance or orientation do vary in content from context to context. More generally, a relativistic interpretation emerges for expressions whose "reliability" varies across authentically exocentric assessments, but remains constant across autocentric assessments in imagined exchanges of place; but a contextualist interpretation emerges for expressions whose

[197] Consistency is maintained by compartmentalizing into "*p*-bodies" as in Section 7.1. above.

reliability varies even with autocentric assessment in imagined exchanges of place.

10.2. Space and logical space

Our next step is to work out these ideas more formally, with the goal of motivating a truth predicate which behaves in the way just outlined, given a system for rating representations as reliable or unreliable, and a distinction between extrinsic and intrinsic properties.

To capitalize on the idea that relativistically interpreted expressions are those whose reliability varies across individuals, but not across imagined changes in position, it will be useful to employ a formalism in which predication is represented as positioning in *logical space*.[198]

By a *space*, let us mean a set S of "points," paired with a set D of "dimensions," where each element of D is a partial ordering on S.[199] For example, we may pair $space_{x,t,w}$ (the set of stationary points fixed by the figure-ground perception of x at t in w, as discussed in Section 3.5. above) with three partial ordering relations $<_{higher\text{-}than}$, $<_{further\text{-}forward\text{-}than}$ and $<_{further\text{-}left\text{-}than}$, relating to the normal orientations of the intrinsic front and top of x at t in w in the intuitive way. This yields a 3-dimensional space, which (with suitable additional constraints) we may regard as model of locations in the ordinary intuitive sense. If we let a set S consist of pairs of a time and an element of $space_{x,t,w}$, and let the corresponding D include an $<_{earlier\text{-}than}$ relation on such pairs in addition to the three dimensions just mentioned, we obtain a 4-dimensional system of space-time. Alternatively, we could represent space-time as a set of primitive points, rather than as a set of pairs, so long as D contains four ordering relations with the right properties.

The elements of D may correspond not just to spatial and temporal dimensions, but also to the range of possible values for various kinds of measurement. For example, ordering along the "hue" dimension corresponds to wavelength of reflected light; ordering along the "weight" dimension corresponds to the force of gravity on an object, etc. By locating an object in a space with dimensions such as these, we specify its hue and weight, not just its location and time.

Ordinary binary categories can also be regarded as corresponding to dimensions, though in this case the partial ordering is almost trivial. For example, we may assume a dimension $<_{asleep}$, which orders points in our space into an "asleep" region and a "not asleep" region.[200] In fact, we may regard every predicate in our language as corresponding in some way to a dimension (or some combination of dimensions) in a multidimensional "logical space" $L = \langle S, D \rangle$.

Each individual occupies a region of logical space, according to its properties. But for any individual x, in addition to the region of logical space it occupies, there are various other regions it *may* occupy (or may have occupied); in this sense, logical space is modal. Different assignments of regions in logical space to individuals correspond to different possible worlds. In fact, rather than taking possible worlds as primitive, we may define them in terms of logical space: a *possible world*

---

[198] Our approach will be closer to the conception of logical space in Thomason (1979) than to the classical conception developed in Wittgenstein (1961); but I make no attempt to remain completely true to either.

[199] For convenience, let us adopt a "strong" definition of partial orderings as irreflexive, asymmetric, transitive relations. We could easily reformulate the system in terms of "weak" partial orderings (reflexive, antisymmetric, transitive relations) if desired.

[200] That is, for any points a, b in our space S, let $a <_{asleep} b$ iff a is in the "not asleep" region and b is in the "asleep" region. It is easy to see that $<_{asleep}$ is irreflexive, asymmetric and (trivially) transitive.

is a partial function mapping elements of a set U (of "possible individuals") onto subsets of S, constrained in certain ways to be discussed shortly.

As a concrete illustration of this idea, consider a "toy" system consisting of a set $S = \{a_1,\ldots,a_{32}\}$ ordered by five dimensions: $<_{higher\text{-}than}$, $<_{further\text{-}forward\text{-}than}$, $<_{further\text{-}left\text{-}than}$, $<_{earlier\text{-}than}$ and $<_{asleep}$, where:

(268) a. $a_1,\ldots,a_{16} <_{higher\text{-}than} a_{17},\ldots,a_{32}$

b. $a_1,\ldots,a_8, a_{17},\ldots,a_{24} <_{further\text{-}forward\text{-}than} a_9,\ldots,a_{16}, a_{25},\ldots,a_{32}$

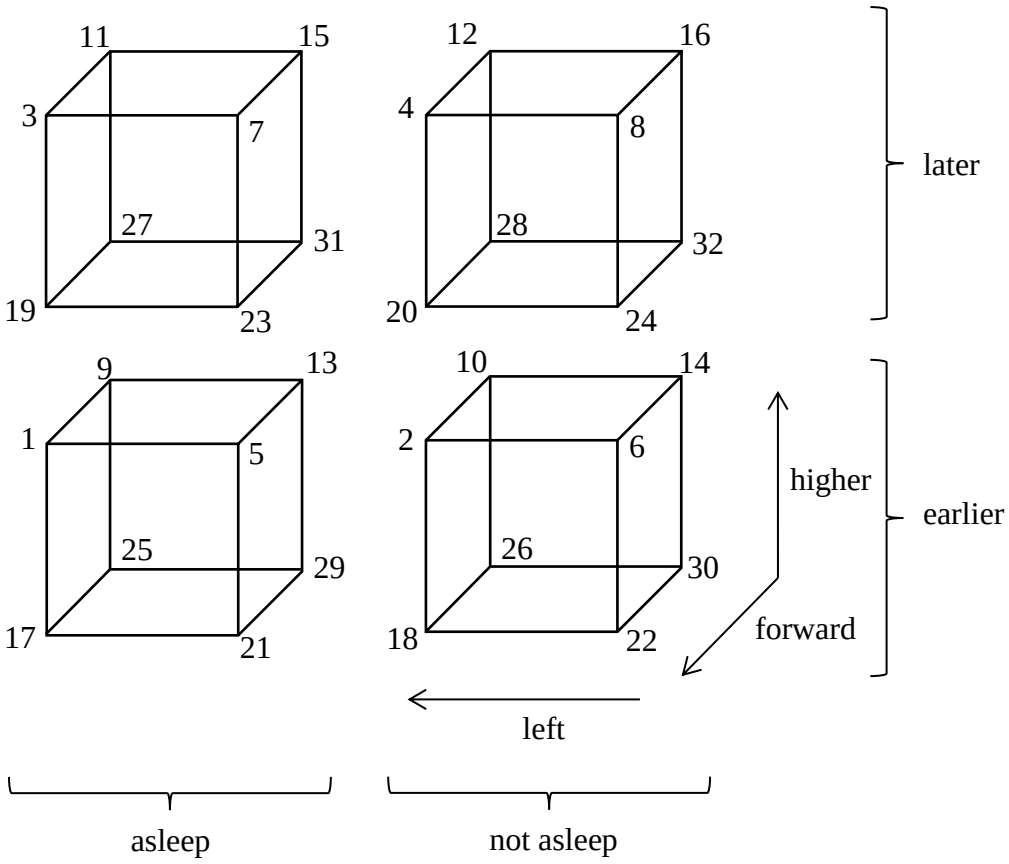
c. $a_1,\ldots,a_4, a_9,\ldots,a_{12}, a_{17},\ldots,a_{20}, a_{25},\ldots,a_{28} <_{further\text{-}left\text{-}than} a_5,\ldots,a_8, a_{13},\ldots,a_{16}, a_{21},\ldots,a_{24}, a_{29},\ldots,a_{32}$

d. $a_1, a_2, a_5, a_6, a_9, a_{10}, a_{13}, a_{14}, a_{17}, a_{18}, a_{21}, a_{22}, a_{25}, a_{26}, a_{29}, a_{30} <_{earlier\text{-}than} a_3, a_4, a_7, a_8, a_{11}, a_{12}, a_{15}, a_{16}, a_{19}, a_{20}, a_{23}, a_{24}, a_{24}, a_{27}, a_{28}, a_{31}, a_{32}$

e. $a_1, a_3, a_5, a_7, a_9, a_{11}, a_{13}, a_{15}, a_{17}, a_{19}, a_{21}, a_{23}, a_{25}, a_{27}, a_{29}, a_{31} <_{asleep} a_2, a_4, a_6, a_8, a_{10}, a_{12}, a_{14}, a_{16}, a_{18}, a_{20}, a_{22}, a_{24}, a_{26}, a_{28}, a_{30}, a_{32}$

This can be visualized as in (269):

(269)



Suppose U = {John, Mary}. Let w_1(John) = $\{a_1, a_4\}$ and w_1(Mary) = $\{a_5, a_7\}$. Then w_1 is a world where John is at first asleep and to the left of Mary, who is also asleep; John later wakes up, but remains to the left of Mary, who is still asleep. In contrast, if w_2(John) = $\{a_1, a_3\}$ and w_2(Mary) = $\{a_5, a_8\}$, it is Mary who wakes up, while John remains asleep.

None of this implies any change to our semantic rules; we may continue to assign denotations relative to worlds and perspectives, using our existing type system. All we have done is to elaborate a particular conception of what a "world" is — namely, a configuration of individuals in logical space. But if desired, we could relate our semantic theory more explicitly to the idea of logical space by reformulating our lexical entries for predicates to refer to locations in logical space, rather than using ordinary non-technical English vocabulary. For example, our current lexical entry for $sleep_{inf}$, in (270)a. (repeated from (122)) could be reformulated as in (270)b.:

(270) a. If USE(α, $sleep_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda y : y \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . y$ sleeps at t in $w]$

b. If USE(α, $sleep_{inf}$) then $Lex(\alpha, u, w, p) = [\lambda y : y \in \mathbf{D}_e . \lambda t \in \mathbf{D}_i . \exists a,b[a \in w(y) \ \& \ b <_{asleep} a \ \& \ \mathbf{time}(a) = t]$

(Here, **time**(*a*) is the time of *a*; we would require that for any *a*, *b*: $a <_{earlier\text{-}than} b$ iff **time**(*a*) < **time**(*b*).)

Note that the typing, argument structure, content, etc., all remain unaltered by this reformulation; it is simply a rewording of the same stipulation in terms of a particular conception of logical space and its relation to possible worlds.

Depending on what kinds of dimensions we consider logical space to be structured by, we will need to constrain the set of functions from individuals to regions of logical space which we regard as possible worlds. If, for example, our space includes a dimension $<_{=j}$, sorting points into an "equal to John" region and a "not equal to John" region (in much the same way as $<_{asleep}$ sorts points into an "asleep" region and a "not asleep" region), we should regard a function as a possible world only if it maps John, and nobody else, onto a location in the "equal to John" region. If John has any other properties essentially, we should likewise regard as worlds only those functions which map him onto the regions in logical space corresponding to those properties. If one property relates logically to another, only those functions respecting this relation may be regarded as worlds: For example a world can map John onto a location in the "loves Mary" region only if it maps Mary onto a location in the "loved by John" region. These and many other constraints would have to be thoroughly explored and formalized in a full-scale development of the idea of logical space. For our purposes here, only a brief sketch is necessary; so we forego a fuller development and simply note that such constraints are necessary.

10.3. Reliability and traversable dimensions

Imagining oneself in a different position in ordinary or logical space may be conceived as "traversing" one or more dimensions in one's imagination. Imagining yourself further to the left than you are now involves a shift in your position along the $<_{further\text{-}left\text{-}than}$ dimension; imagining yourself asleep when you are awake involves a shift along the $<_{asleep}$ dimension.

But constraints like those discussed at the end of the last section render certain dimensions untraversable. For example, no one may "move" along the $<_{=j}$ dimension; if $a <_{=j} b$, then there are

no worlds w, w' and no individual x such that $a \in w(x)$ and $b \in w'(x)$.[201] No one can be John but John. In contrast, one may easily traverse the $<_{asleep}$ dimension: It could easily happen that $a \in w(x)$ and $b \in w'(x)$ even if $a <_{asleep} b$, because someone who is awake may be asleep in another world, or in the same world at a different time.

If, as suggested in Section 10.1. imagining yourself in someone else's position involves retaining one's "intrinsic" properties and altering only one's "outward" properties, then we must regard the dimensions corresponding to intrinsic properties as untraversable (at least for the purposes of this imaginative task), even though they may not be *essential* in the usual sense. For example, imagining yourself in the place of someone eating chili may involve a shift along the $<_{eating\text{-}chili}$ dimension, as well as along one or more ordinary spatial dimensions like $<_{further\text{-}left\text{-}than}$ or $<_{further\text{-}forward\text{-}than}$, but not along dimensions corresponding to one's personal tastes and preferences (such as, say, $<_{enjoys\text{-}spicy\text{-}food}$); this is what makes it imagining yourself in the other person's position, rather than imagining that you are the other person. This is so despite the fact that the degree to which one enjoys spicy food can vary from world to world and from time to time.

The distinction between traversable and untraversable dimensions allows us to draw an analogous distinction between *adoptable* and *unadoptable* perspectives. Recall that a perspective is simply a triple $p = \langle x, t, w \rangle$, where x is an individual, t is a time, and w is a world. Let $loc(p)$ be x's location at t in w — that is, $\{a \in S \mid a \in w(x)$ & $\mathbf{time}(a) = t\}$. We say that p is *adoptable* from p' iff every point a in $loc(p)$ corresponds to a point b in $loc(p')$, such that there is a path between a and b along traversable dimensions only. More precisely:

(271) $p = \langle x, t, w \rangle$ is *adoptable* from $p' = \langle x', t', w' \rangle$ iff there is a function f from $loc(p')$ to $loc(p)$ such that for all dimensions $<_d$ and all $a \in loc(p')$: If $a <_d f(a)$ or $f(a) <_d a$, then $<_d$ is traversable.

Intuitively, $\langle x, t, w \rangle$ is adoptable from $\langle x', t', w' \rangle$ if at t' in w', x' can imagine himself or herself at $loc(x, t, w)$, with no alteration to intrinsic properties.

It should be noted here that because dimensions like $<_{=x}$ are untraversable, $\langle x, t, w \rangle$ will be adoptable from $\langle x', t', w' \rangle$ only if $x = x'$. In other words, the only perspectives one may "adopt" in this sense are one's own perspectives at other worlds or times, not perspectives which properly belong to another individual.

Now let us consider reliability of representations. As discussed in Section 10.1. , a representation may be reliable to varying degrees. An individual organism x may need to assess the reliability of multiple representations, so let us assume that (in any given world w and time t), x has some assignment R of degrees of reliability to representations. We write '$R(\boldsymbol{\rho}) = d$' to mean that (at t in w), x regards representation $\boldsymbol{\rho}$ is reliable to degree d. We may represent degrees of reliability as real numbers between 0 and 1 inclusive. '$R(\boldsymbol{\rho}) = 0$' means that $\boldsymbol{\rho}$ is regarded as completely unreliable, '$R(\boldsymbol{\rho}) = 1$' means that $\boldsymbol{\rho}$ is regarded as perfectly reliable, and intermediate values mean that $\boldsymbol{\rho}$ is regarded as somewhat reliable, but imperfectly.

A very simple representation system, as with our primitive organism traversing its environment, or relying on the communications of others without the use of imaginative

[201] Note that there is no assumption here that $w \neq w'$.

assessment in their positions, will perhaps need no more of an interpretative system than this. All assessment is autocentric. The organism assigns a degree of reliability to a representation **ρ**; if it does so well, this degree will match its actual degree of reliability closely and the organism will successfully interact with its environment.

As long as it is representations rather than particular *uses* of representations which are assessed,[202] and all reliability assessment is fully autocentric, this sort of system will not allow an individual to assess the reliability of a signal differently depending on the position of its source. For example, it will not allow a "Wolf is near" signal to be assessed differently depending on how distant the individual issuing the signal is. A natural later development, therefore — particularly for representations in a communication system — is to begin assigning degrees of reliability to representation-uses, and not just representations as reusable objects. We therefore extend our notation and write '$R(\rho) = d$' to mean that ρ is regarded as reliable to degree d where ρ is a representation-use.

Assessing uses rather than representations for reliability allows an individual to treat a "Wolf is near" signal received from a distant source as less deserving of attention than a "Wolf is near" signal from a closer source. But it would be useful not simply to reduce the attention paid to distant signals of this kind, but to gather from them the information that a wolf is near the signaler — that the signaler stands in the same spatial relation to the wolf as the assessor does when a wolf is nearby. This takes an additional step, the development of semi-exocentric assessment.

Semi-exocentric assessment of reliability allows an organism to "track" not just a single degree of reliability for a given representation-use, but an assignment of degrees to adoptable perspectives. That is, rather than just employing a simple assignment R from representation-uses to degrees, an individual x (at a time t in a world w) must use a more complex assignment $\boldsymbol{R}$ from *pairs* of a representation-use and an adoptable perspective to degrees. We now write '$\boldsymbol{R}(\rho, p) = d$' to mean that ρ is regarded as reliable to degree d from perspective p, where p is adoptable from $\langle x, t, w\rangle$.

Semi-exocentric assessment involves "simulating" another individual's perspective by imagining oneself in that other individual's position. Where a and b are points in logical space, let us write '$a \approx b$' to mean that a and b differ only on untraversable dimensions. Now we can define simulation as follows:

(272) *p simulates p′* iff there is a function f from $loc(p)$ to $loc(p')$ such that for all $a \in loc(p)$: $a \approx f(a)$.

Intuitively, one perspective simulates another if it is just like it, except for those differences forced by the untraversability of certain dimensions.

Suppose some individual y issues a signal ρ at time t in world w. If some other individual x (at time t' in world w') calculates $\boldsymbol{R}(\rho, p)$, for all those perspectives p which simulate $\langle y, t, w\rangle$ and are adoptable from $\langle x, t', w'\rangle$, this amounts to x assessing ρ relative to all possible circumstances where x is in y's position, but which are otherwise as close to y's original circumstance as possible.

In this situation, on the assumption that y is sincere in issuing ρ, it must be the case that y assigns ρ a high degree of reliability. If $\boldsymbol{R}(\rho, p)$ is low, however, then x assesses ρ as *un*reliable in the imagined circumstance where x is in y's position. Because x assesses ρ differently than y, in the

[202] And the assessment is "absolute" rather than relative to contexts of use.

imagined circumstance where *x* is in *y*'s position, it is apparent that *x* disagrees with *y*. In contrast, if $\boldsymbol{R}(\rho, p)$ is high, then it is clear that (in imagining himself or herself in *y*'s position) *x* reaches approximately the same judgment about the reliability of ρ as *y* does; so *x* and *y* are in at least rough agreement.

Assuming spatial dimensions are traversable, we get the desired result that if ρ is the sort of signal that one issues when a wolf is near, and *x* sees that a wolf is near *y*, then *x* agrees with *y*; but if *x* knows that there is actually no wolf near *y*, then $\boldsymbol{R}(\rho, p)$ will be low, and *x* disagrees with *y* — even if *x* assigns a high degree of reliability to ρ relative to *x*'s own actual position.

In contrast, assuming dimensions like $<_{enjoys\text{-}spinach}$ are untraversable, if ρ is the sort of signal that one issues when one is eating tasty food, and *x* sees that *y* is eating spinach, then $\boldsymbol{R}(\rho, p)$ will always be equal to $\boldsymbol{R}(\rho, \langle x, t', w' \rangle)$. If *x* does not enjoy eating spinach, then $\boldsymbol{R}(\rho, p)$ will be low; *x* assesses ρ differently than *y*, even in the imagined circumstance where *x* is in *y*'s position, and *x* disagrees with *y*. Here we see already that the our system draws a distinction between indexical signals (like "wolf is near" signals) on the one hand, and relativistic signals (like "tasty food" signals) on the other, even though as yet we have not defined any notion of truth or truth-conditional content for it.

If ρ is the sort of signal one issues, based not on the peculiarities of one's circumstance or intrinsic properties, but purely on the basis of some eternal, context-independent, objective fact — if 2+2=4, for example — the pattern is the same as for signals like the one for eating tasty food: $\boldsymbol{R}(\rho, p)$ will always be equal to $\boldsymbol{R}(\rho, \langle x, t', w' \rangle)$, with no shift as *x* imagines himself or herself in *y*'s position.[203] In the assignment of degrees of reliability relative to simulations of the perspective of the issuer of a signal, we can see the beginnings of a system in which representations may express different contents relative to different contexts of use; but where taste-sensitivity patterns with objectivity in *not* motivating such variation.

Recognizing that a representation-use might be reliable for someone else — that is, from an *un*adoptable perspective — even if unreliable for oneself, involves yet another step: the development of fully exocentric assessment. It is only at this step that the user of a representation system needs to assess relative to the perspectives of others. This means that the user *x* must employ (at *t* in *w*) a reliability assignment $\boldsymbol{R}$ such that $\boldsymbol{R}(\rho, p)$ is defined even when *p* is not adoptable from $\langle x, t, w \rangle$.

Suppose again that *y* produces a representation-use ρ (at some time *t* in some world *w*), of the sort one issues when one is eating tasty food. If *x* receives the signal and assumes *y* is sincere, then the reliability assignment $\boldsymbol{R}$ used by *x* will be such that $\boldsymbol{R}(\rho, \langle y, t, w \rangle)$ is high. But assuming dimensions like $<_{enjoys\text{-}spinach}$ are not traversable, if *x* sees that *y* is eating spinach, and *x* does not enjoy spinach, then all simulations *p* of $\langle y, t, w \rangle$ adoptable from $\langle x, t, w \rangle$ will be such that $\boldsymbol{R}(\rho, p)$ is low. Thus *x* recognizes that *x* and *y* have different assessments about the reliability of ρ, even in when *x* imagines himself of herself in *y*'s position. In this case it is clear that *x* disagrees with *y*.

[203] It is hard to imagine what use an organism which is cognitively much simpler than ourselves could have for such signals — a fact which does not undermine the theoretical point, but which should give pause to anyone who thinks of context-independence and objectivity as simple defaults from which indexically and relativistically interpreted expressions must be seen as marked departures requiring special explanation.

10.4. Reliability-based content

So far, our discussion has been at the level of representations and representation-uses, not with the "contents" of those representations and uses. But it would be useful to a communicating organism to have a notion of content — and particularly of sameness and difference of content — because certain representation-uses are systematically equivalent to one another in their reliability, while others differ. The "content" of a representation-use is what it has in common with other uses with which it is equivalent in some semantic or quasi-semantic property such as truth or reliability.

For a very simple system used by an organism that engages only in autocentric assessment of reliability, we may identify the content of a representation-use ρ with its reliability profile relative to perspectives which are adoptable by that individual:

(273) If *x* uses ***R*** at *t* in *w*, the content of ρ induced by ***R*** for *x* at *t* in *w* is the function [λ*p* : *p* is adoptable from ⟨*x*, *t*, *w*⟩ . ***R***(ρ, *p*)]

The idea here is that the content of ρ is something which determines how reliable ρ is for *x* in the various possible circumstances *x* might find himself or herself in.

As soon as we look at a system involving semi-exocentric assessment, however, things become more complicated. I suggested in Section 10.1. that an individual will differentiate contents as needed in order to keep his or her own judgments consistent with one another. For example, if *x* receives a signal ρ from *y*, and judges ρ to be unreliable from *x*'s own perspective, but reliable from some other perspective which is adoptable by *x* and simulates *y*'s perspective, then it is natural for *x* to construe ρ as actually having a different content than it would have if issued from *x*'s position. The difference in reliability is resolved by construing ρ as expressing different contents in the two cases. We can achieve this effect by modeling the content of ρ as a function which assigns to each adoptable perspective *p*, not the reliability of ρ at *p*, but the reliability of ρ at those perspectives which are adoptable from *p*, and simulate the perspective of the "speaker" of ρ at the time of ρ in the world of *p*:

(274) If *x* uses ***R*** at *t* in *w*, the content of ρ induced by ***R*** for *x* at *t* in *w* is the function [λ*p* : *p* is adoptable from ⟨*x*, *t*, *w*⟩ and for all *p′* adoptable from *p* and simulating ⟨$speaker_{w,\rho}$, $time_{w,\rho}$, *w*⟩, ***R***(ρ, *p′*) = *n* . *n*]

For example, suppose *x* receives a signal ρ of the sort that one issues when a wolf is nearby. There are various times and worlds which where *x* might be, without alteration to intrinsic properties — that is, various perspectives adoptable by *x* — and as in (273), the content of ρ will be an assignment of reliability rating for ρ relative to these perspectives. The times and worlds of these perspectives vary as to who the "speaker" of ρ is and as to whether there is a wolf near ρ. But for each such perspective *p*, if *x* imagines himself or herself in the position of the speaker of ρ at the time of ρ and autocentrically assesses from this imagined position — that is, if *x* assesses from all those perspectives adoptable from *p* which simulate the perspective of the speaker — the result will be the reliability rating for the *content* of ρ relative to *p*. Thus the content of ρ will give, for any perspective *p* adoptable by *x*, a value corresponding to the degree to which *x* can rely on the information that a wolf is near the speaker of ρ at the time of ρ, in the world of *p*.

Under this way of assigning contents, two distinct signals, both of the sort issued when a

wolf is nearby but with different speakers, will receive different contents; the signals are indexical. But two signals of the sort issued when one is eating tasty food, likewise with different speakers, will receive the same content, on the assumption that the dimension corresponding to what one is eating is traversable, but the dimension corresponding to what kinds of foods one enjoys is not.

This is still not a very language-like system, however. Interpretation is based on reliability rather than truth, and all assessment is relative to adoptable perspectives. According to (274), the content which ***R*** induces for *x* at *t* in *w* is a function whose domain is limited to perspectives adoptable from ⟨*x*, *t*, *w*⟩. Because different perspectives are adoptable by different individuals, content will vary from individual to individual: Even though an individual *x* will assign the same content to a "tasty food" signal no matter who issues it, another individual *y* will assign a different content, based on which perspectives are adoptable by *y*.

A more "interpersonal" notion of content becomes useful as the users of a representation system begin to engage in fully exocentric assessment. Assessing a content relative to someone other than oneself requires that content to be applicable to perspectives other than just those which one can adopt as one's own. That is, (274) must be modified so that the content induced by ***R*** has a wider domain, not restricted to those perspectives adoptable from ⟨*x*, *t*, *w*⟩. Dropping this restriction and keeping everything else the same, we obtain:

(275) If *x* uses ***R*** at *t* in *w*, the content of ρ induced by ***R*** for *x* at *t* in *w* is the function [λ*p* : for all *p′* adoptable from *p* and simulating ⟨$speaker_{w,\rho}$, $time_{w,\rho}$, *w*⟩, ***R***(ρ, *p′*) = *n* . *n*]

This definition still appeals to the idea of adoptable perspectives simulating the perspective of the "speaker" of a signal, but the content itself is now a function which applies to any perspective, not just those adoptable by *x*. ***R*** induces the same content for ρ for all individuals — and, for that matter, all times. We may simplify (275) as follows:

(276) The content of ρ induced by ***R*** in *w* is the function [λ*p* : for all *p′* adoptable from *p* and simulating ⟨$speaker_{w,\rho}$, $time_{w,\rho}$, *w*⟩, ***R***(ρ, *p′*) = *n* . *n*]

From here, it is a short step to adapt the notation of previous chapters:

(277) $\text{⟨}\rho\text{⟩}^{u}$ = [λ*p* : for all *p′* adoptable from *p* and simulating ⟨$speaker_{u,\rho}$, $time_{u,\rho}$, *u*⟩, ***R***(ρ, *p′*) = *n* . *n*]

If contents are assigned this way, we obtain a distinction between relativistically and objectively interpreted representations, which was absent at the previous stage. If ρ is a signal of the sort one issues when eating tasty food, for example, it might well be the case that for every perspective adoptable from ⟨*x*, *t*, *w*⟩ simulating ⟨$speaker_{w,\rho}$, $time_{w,\rho}$, *w*⟩, ***R*** assigns a low degree of reliability, but for every perspective adoptable from ⟨*y*, *t*, *w*⟩ simulating ⟨$speaker_{w,\rho}$, $time_{w,\rho}$, *w*⟩, ***R*** assigns a high degree of reliability. The content induced by ***R*** for *x* at *t* in *w* would therefore map *x*'s own current actual perspective onto a low number, but *y*'s current actual perspective onto a high number. So *x* can autocentrically judge whatever the speaker is eating not to be tasty, but exocentrically judge it to be tasty relative to *y*. A fully objective signal, however, like our earlier

example of a signal one issues if 2+2=4, will yield the same value for all perspectives; whether *x* simulates the speaker's perspective or *y* simulates the speaker's perspective, the signal is equally reliable.

10.5. From reliability to truth

We have reached a system which shows significant parallels with the relativist semantics for natural languages argued for in earlier chapters: Indexical expressions vary in content, according to their context of use (here represented as the speaker and time of use); non-indexical expressions do not. Relativistically interpreted expressions have contents whose values vary according to perspectives; objectively interpreted expressions do not. However, the system does not yet explicitly involve *truth*, or any kind of semantic compositionality. My suggestion is that truth is an idealization of reliability, and that therefore an assignment of truth values will retain certain features of reliability assignment. We may regard a relativist assignment of truth conditions as a reflection of the patterns of reliability we have already seen. Our next task, therefore, is to explore in a little more detail how the system changes if we define content in terms of truth rather than reliability.

In Section 10.1. I suggested there were at least three differences between reliability and truth: First, reliability is strictly a matter of the practical consequences of using a representation, while truth abstracts away from these consequences and depends instead on a tight, systematic correlation between representations and the environment of the users of the representation-system. Second, reliability is scalar, but truth is bivalent. Third, a representation can be reliable for certain purposes but not others, but truth is purpose-insensitive.

In our formal development, we have simplified by ignoring the purpose-sensitivity of reliability, so in making the transition from reliability-based content to truth-based content, we will not have to make any adjustments on that account.

Developing a bivalent notion of truth from our scalar notion of reliability will require some adjustment, however. Perhaps the simplest, most natural way of deriving a bivalent notion of truth from our scalar notion of reliability would be to regard a content as true iff it is perfectly reliable, and false iff it is less than perfectly reliable. In this case we could adapt (277) to define truth-based content as in (278):

(278) $\text{ɕ}\rho\text{ɔ}^{u} = [\lambda p$. for all p' adoptable from p and simulating $\langle speaker_{u,\rho}, time_{u,\rho}, u\rangle$, $\boldsymbol{R}(\rho, p') = 1]$

But this does not yet take into account the third suggested difference between truth and reliability, namely the dependence of truth on systematic correlation between the representations in a system and the environments in which the users of that system may find themselves, as opposed to direct dependence on the practical consequences of using those representations. This systematic correlation may be identified with a *denotation* relation.

The notion of denotation, as we have assumed it since the beginning of this book,[204] is inextricably linked to the notion of truth: The denotation of a use α of an expression **α** is whatever **α** is employed (in α) to "talk about," in the sense of *talk about* where the truth value of what is said in using a sentence depends on what the parts of the sentence are used to talk about. This idea

[204] Recall the discussion in Section 1.5.

motivates the substitution principle in (8), and from there, the idea that interpretation is assigned homomorphically, both at the level of denotation and at the level of content.

Any useful notion of denotation must be *systematic*, not only in the sense that the denotations of syntactically complex expressions are determined from the denotations of their parts, but also in the sense that the denotations of the lexical items do not vary randomly from use to use or index to index, but follow patterns fixed by their lexical entries. In certain cases, the systematicity of the denotation relation motivates a separation of the idea of *truth* from that of perfect reliability, in the sense developed above.

This can be seen by comparing sentences containing first person pronouns, for example. Suppose you receive a signal which is similar in relevant respects to the sentence *I am John* — the signal includes a part corresponding to the pronoun *I*, and another part corresponding to the predicate *am John*. Dimensions like $<_{=j}$ are untraversable, so (unless you are John) there will be no perspective which is adoptable from your perspective which makes the signal reliable. If truth always coincided with perfect reliability, we would count this use of the sentence as *false* relative to your perspective. In fact, we would count *all* uses of this sentence as false relative to your perspective, but true relative to John's perspective; the sentence would receive a relativist semantics. Given that *am John* holds John and no one else, we could derive this reading compositionally by lexically stipulating that the pronoun *I* (or whatever the corresponding part of the signal might be) denotes the judge of its context of assessment:

(279) If USE(α, *I*) then $Lex(\alpha, u, w, \langle x, t, a\rangle) = x$.

But now consider a signal like *I am near the river*. Dimensions like $<_{near\text{-}the\text{-}river}$, unlike $<_{=j}$, *are* traversable, so relative to any perspective which is adoptable from your perspective and which simulates the perspective of the speaker, the signal will be just as reliable as it is relative to the speaker's perspective. If truth is just perfect reliability, this renders any use of this sentence true iff the speaker of that use is near the river. We will not derive this reading compositionally if we lexically stipulate that the pronoun *I* denotes the judge of its context of assessment, as the previous example seemed to suggest. Instead, we must stipulate that it denotes the speaker:

(280) If USE(α, *I*) then $Lex(\alpha, u, w, \langle x, t, a\rangle) = speaker_{u,\alpha}$.

These two kinds of example impose conflicting constraints on the semantics of first person pronouns — at least if we assume that truth is the same as perfect reliability, and also that the semantics of lexical items conform to systematic stipulations like (279) or (280). Our choices, therefore, are either to give up the identification of truth with perfect reliability (at least in our technical sense of *reliability*), or to adopt much more complex and irregular lexical stipulations, for example by letting *I* sometimes denote the speaker and sometimes denote the judge, depending on what other expressions are in its syntactic context.

The semantics of the English pronoun *I* conforms to (280), not (279), so in that respect English has "chosen" the first of these two options, separating the notion of truth from that of reliability, in the interest of maintaining a regular, systematic denotation assignment.[205] This is not

[205] One may ask further why English opted for (280) rather than (279) as the general semantics for *I*. I suggest that (280) is the more natural choice because (279) is motivated only by examples in which the predicate represents an

something unusual or idiosyncratic about the English pronoun *I*, but rather seems typical. I suggest that in general, languages are structured so that the systematicity of lexical denotation assignment is maintained, even at the expense of identifying truth with perfect reliability; but except as this principle requires, truth-based content will reflect reliability-based content as in (278).

In this connection we may note that the idea of truth as perfect reliability does not seem to conflict with the systematicity of denotation assignment as these principles apply to expressions like *fun* or *tasty*, in contrast to *I*. Uses of sentences like *The chili is tasty* or *The roller coaster is fun* will be true relative to a perspective $p = \langle x, t, a \rangle$ iff for all p' adoptable from p and simulating the perspective of the speaker, those uses are perfectly reliable relative to p'. On the assumption that the dimensions corresponding to taste preferences are untraversable, this has the effect that it is the tastes of x, not the tastes of the speaker, on which the truth value relative to $\langle x, t, a \rangle$ is based. In other words, a relativist interpretation for such examples falls out from our assumptions about which dimensions are traversable, and how truth relates to reliability.

We have not yet motivated a notion of content exactly like that adopted in the main part of the book. Content as defined in (278), adjusted to reflect the dependence of truth on systematic denotation assignment, is a function from perspectives to truth values. But our grammar treats sentence-contents as functions from *world-perspective pairs* to truth values. Each perspective, of course, includes a world as a component, but the notion of content we are ultimately aiming for involves a second world parameter, in addition to the one which comes as part of the perspective parameter. How is that parameter motivated?

That world parameter plays a role only in the semantics of constructions containing modal operators, attitude verbs, or other expressions which create intensional contexts. A language without such expressions would not need a notion of content which parameterized truth to worlds, independently of perspectives. Unlike the perspective parameter, the world index is not motivated by general considerations about the imaginative basis of assessment and the relation between truth and reliability, but only by the inclusion in the language of intensional vocabulary. A detailed account of how a language might come to employ such an index is therefore an independent topic from our concerns in this section, and will not be attempted here.

Before concluding this section, it is worth reminding ourselves that the account presented here will make specific predictions about which expressions receive a relativist interpretation and which receive an indexical interpretation only *given* a specification of which dimensions are traversable and which are not. I have suggested that this distinction corresponds to a distinction between "intrinsic" and "extrinsic" properties, but have not presented any definite criteria for identifying which properties are intrinsic and which are extrinsic. It is natural to assume that any property which an individual has *essentially* (that is, in all possible worlds), it will also have intrinsically; but I suggested that an individual's knowledge, dispositions, and preferences should be considered intrinsic, even though such things can vary from time to time and world to world.

Given the difficulties involved in making this distinction precise, it would not be very surprising if our informal, pretheoretic intuitions on the matter were also somewhat variable and

---

untraversable dimension, such as $<_{=j}$. Such predicates represent a special case, with traversability being more typical. Either choice forces a separation between truth and reliability, (280) forces a separation in a smaller range of cases, and therefore better maintains the idea of truth as idealized reliability.

obscure. If so, we might expect that our intuitions about which kinds of linguistic examples call for a relativist analysis and which call for an indexical analysis will be similarly variable and obscure. If, for example, the reliability of some expression varies along a dimension which some people regard as traversable and others do not, we may expect that those who regard the dimension as traversable will interpret the expression as indexical, while those who regard it as untraversable will assign it a relativist interpretation. Interpersonal variation in where we draw the line between intrinsic and extrinsic properties may thus be a source for idiolectal variation in linguistic meaning.[206] If the classification of properties as intrinsic or extrinsic varies with the purpose for which the distinction is drawn, we might also expect that certain uses of an expression might call for an indexical analysis and other uses for a relativist analysis, according to the purpose for which the expression is used. These, and many related topics, suggest themselves as interesting avenues for further exploration in the framework developed here; but will have to left to another time.

[206] I have occasionally encountered individuals who claimed not to share the intuition of faultless disagreement for pairs like *Licorice is tasty/Licorice is not tasty* or *Roller coaster are fun/Roller coaster are not fun*. Perhaps such individuals simply assume that logical space is more fully traversable than other people do — perhaps even that all dimensions are traversable, hence that all perspectives are adoptable.

References

Bach, Emmon (1979) Control in Montague Grammar, *Linguistic Inquiry* 10.4.515–531.

Banfield, Ann (1982) *Unspeakable Sentences: Narration and Representation in the Language of Fiction*. Routledge & Kegan Paul.

Barker, Chris (2002) The dynamics of vagueness, *Linguistics and Philosophy* 25.1.1–36.

Barker, Chris (2013) Negotiating taste, *Inquiry* 56.2-3.240–257.

Barwise, Jon and Cooper, Robin (1981) Generalized quantifiers and natural language, *Linguistics and Philosophy* 4.1.159–219.

Barwise, Jon and Perry, John (1983) *Situations and Attitudes*. MIT press.

Beaver, David I. (1997) Presupposition, *Handbook of Logic and Language*, 939–1008, ed. by van Benthem, Johan and Alice ter Meulen, MIT Press.

Belnap, Nuel and Green, Mitchell (1994) Indeterminism and the thin red line, *Philosophical Perspectives* 8.365–388.

Bhatt, Rajesh and Izvorski, Roumyana (1995), Genericity, implicit arguments and control.

Bittner, Thomas (1997) A qualitative coordinate language of location of figures within the ground, *Spatial Information Theory A Theoretical Basis for GIS*, 223–240, ed. by Hirtle, Stephen C. and Andrew U. Frank, Springer.

Bittner, Thomas (1998), Toward a model-theory for figure ground location. *Proceedings of the 6th Symposium on Mathematics and AI*.

Boolos, George (1998) *Logic, Logic, and Logic*. Harvard University Press.

Bylinina, Lisa (2013) *The Grammar of Standards: Judge-Dependence, Purpose-Relativity, and Comparison Classes in Degree Constructions*, University of Utrecht dissertation.

Cappelen, Herman and Hawthorne, John (2009) *Relativism and Monadic Truth*. Oxford University Press.

Chierchia, Gennaro (1984) *Topics in the Syntax and Semantics of Infinitives and Gerunds*, University of Massachusetts dissertation.

Chierchia, Gennaro (1989) Anaphora and attitudes *de se*, *Semantics and Contextual Expression*, 1–31, ed. by Bartsch, Renate, Johan van Benthem, and P. van Emde Boas, Foris Publications.

Chomsky, Noam (1977) On Wh-movement, *Formal Syntax*, 71–132, ed. by Culicover, Peter, Thomas Wasow, and Adrian Akmajian, Academic Press.

Chomsky, Noam (1980) On binding, *Linguistic Inquiry* 11.1.1–46.

Chomsky, Noam (1981) *Lectures on Government and Binding*. Foris Publications.

Chung, Sandra, Ladusaw, William A., and McCloskey, James (1995) Sluicing and logical form, *Natural Language Semantics* 3.3.239–282.

Cohen, Jonathan (2013) Indexicality and the puzzle of the answering machine, *Journal of Philosophy* 110.1.5–32.

Collins, John (2013) The syntax of personal taste, *Philosophical Perspectives* 27.51–103.

Cresswell, M. J. (1990) *Entities and Indices*, 41 ed. Kluwer Academic Publishers.

Curme, George (1931) *A Grammar of the English Language* vol. II: Syntax. D.C. Heath & Co.

Davidson, Donald (1967) The logical form of action sentences, *The Logic of Decision and Action*, 81–95, ed. by Rescher, Nicholas, University of Pittsburgh Press.

Decker, Nan (2010) Machine translation: the challenge of ambiguity, *Language in the Real World:*

an Introduction to Linguistics, 107–122, ed. by Behrens, Susan J. and Judith A. Parker, Routledge.
Dekker, Paul (1993) Existential disclosure, *Linguistics and Philosophy* 16.6.561–587.
Doron, Edit (1991), Point of view as a factor of content. *Proceedings From Semantics and Linguistic Theory I*, 51–64, ed. by Moore, Steven and Wyner, Zachary, DMLL, Cornell University.
Dowty, David (1985) On recent analyses of the semantics of control, *Linguistics and Philosophy* 8.3.291–331.
Egan, Andy (2007) Epistemic modals, relativism and assertion, *Philosophical Studies* 133.1–22.
Egan, Andy, Hawthorne, John, and Weatherson, Brian (2005) Epistemic modals in context, *Contextualism in Philosophy: Knowledge, Meaning and Truth*, 131–168, ed. by Preyer, Gerhard and Georg Peter, Oxford University Press.
Enç, Mürvet (1981) *Tense Without Scope: An Analysis of Nouns As Indexicals*, University of Wisconsin – Madison dissertation.
Epstein, Samuel David (1984) Quantifier-pro and the LF representation of PRO_{ARB}, *Linguistic Inquiry* 15.3.499–505.
Evans, Gareth (1977) Pronouns, quantifiers, and relative clauses (I), *Canadian Journal of Philosophy* 7.3.467–536.
Evans, Gareth (1985) Does tense logic rest upon a mistake?, *Collected Papers*, 343–363, Oxford University Press.
Fillmore, Charles (1986), Pragmatically controlled zero anaphora. *Proceedings From the 12th Annual Meeting of the Berkeley Linguistics Society*, 95–107, ed. by Nikiforidou, V., Van Clay, M., Niepokuj, M., and Feder, D., Berkeley Linguistics Society.
Fillmore, Charles (1997) *Lectures on Deixis*. CSLI Publications.
Fillmore, Charles, Kay, Paul, and O'Connor, Mary Catherine (1988) Regularity and idiomaticity in grammatical constructions: the case of *let alone*, *Language* 64.3.501–538.
Frege, Gottlob (1892) On sense and reference, *Translations From the Philosophical Writings of Gottlob Frege*, 2nd ed., ed. by Geach, Peter and Max Black, Basil Blackwell.
Frege, Gottlob (1918) Thought, *The Frege Reader*, 325–345, ed. by Beaney, Michael, Blackwell.
Gamut, L. T. F. (1991) *Logic, Language and Meaning, Volume 2: Intensional Logic and Logical Grammar*. University of Chicago Press.
Geach, P. T. (1965) Assertion, *The Philosophical Review* 74.4.449–465.
Glanzberg, Michael (2007) Context, content, and relativism, *Philosophical Studies* 136.1–29.
Goldberg, Adele (1995) *Constructions: a Construction Grammar Approach to Argument Structure*. University of Chicago Press.
Grice, H. P. (1975) Logic and conversation, *Speech Acts*, 41–58, ed. by Cole, Peter and Jerry L. Morgan, Academic Press.
Groenendijk, J. A. G. and Stokhof, M. B. J. (1984) *Studies on the Semantics of Questions and the Pragmatics of Answers*, University of Amsterdam dissertation.
Hamblin, C. L. (1958) Questions, *Australasian Journal of Philosophy* 36.3.159–168.
Hankamer, Jorge and Sag, Ivan (1976) Deep and surface anaphora, *Linguistic Inquiry* 7.3.391–428.
Heim, Irene and Kratzer, Angelika (1998) *Semantics in Generative Grammar*. Blackwell.
Heim, Irene Roswitha (1982) *The Semantics of Definite and Indefinite Noun Phrases*. Graduate Linguistics Students Association, University of Massachusetts.

Higginbotham, James (2003) Remembering, imagining, and the first person, *Epistemology of Language*, 496–533, ed. by Barber, Alan, Oxford University Press.
Higginbotham, James and May, Robert (1981) Questions, quantifiers and crossing, *Linguistic Review* 1.1.41–80.
Hintikka, Jaakko (1969) Semantics for propositional attitudes, *Philosophical Logic*, 21–45, ed. by Davis, J. W., D. J. Hockney, and W. K. Wilson, Reidel.
Horn, Laurence R. (1985) Metalinguistic negation and pragmatic ambiguity, *Language* 61.1.121–174.
Kamp, Hans (1971) Formal properties of 'now', *Theoria* 37.227–273.
Kamp, Hans (1981) A theory of truth and semantic representation, *Formal Methods in the Study of Language*, 277–322, ed. by Groenendijk, J. A. G., T. M. V. Janssen, and M. B. J. Stokhof, Mathematisch Centrum.
Kaplan, David (1989) Demonstratives, *Themes From Kaplan*, 481–563, ed. by Almog, Joseph, John Perry, and Howard K. Wettstein, Oxford University Press.
Karttunen, Lauri (1973) Presuppositions of compound sentences, *Linguistic Inquiry* 4.2.169–193.
Kiparsky, Paul and Kiparsky, Carol (1970) Fact, *Progress in Linguistics*, 143–173, ed. by Bierwisch, Manfred and Karl Erich Heidolph, Mouton.
Kölbel, Max (2002) *Truth Without Objectivity*. Routledge.
Kölbel, Max (2004) Faultless disagreement, *Proceedings of the Aristotelian Society* 104.53–73.
Koster, Jan and May, Robert (1982) On the constituency of infinitives, *Language* 58.1.116–143.
Kratzer, Angelika (1977) What 'must' and 'can' must and can mean, *Linguistics and Philosophy* 1.337–355.
Kratzer, Angelika (1981) The notional category of modality, *Words, Worlds and Contexts: New Approaches to Word Semantics*, 38–74, ed. by Eikmeyer, H.-J. and H. Rieser, Walter DeGruyter.
Kratzer, Angelika (2009) Making a pronoun: fake indexicals as windows into the properties of pronouns, *Linguistic Inquiry* 40.2.187–237.
Landau, Idan (2000) *Elements of Control: Structure and Meaning in Infinitival Constructions*. Dordrecht.
Lasersohn, Peter (1999) Pragmatic halos, *Language* 75.3.522–551.
Lasersohn, Peter (2005) Context dependence, disagreement, and predicates of personal taste, *Linguistics and Philosophy* 28.6.643–686.
Lasersohn, Peter (2007) Expressives, perspective and presupposition, *Theoretical Linguistics* 33.2.223–230.
Lasersohn, Peter (2008) Quantification and perspective in relativist semantics, *Philosophical Perspectives* 22.1.305–337.
Lasersohn, Peter (2009a) Compositional interpretation in which the meanings of complex expressions are not computable from the meanings of their parts, *Theory and Evidence in Semantics*, 133–158, ed. by Hinrichs, Erhard and John Nerbonne, CSLI Publications.
Lasersohn, Peter (2009b) Relative truth, speaker commitment, and control of implicit arguments, *Synthese* 166.2.359–374.
Lasersohn, Peter (2011) Context, Relevant Parts and (Lack of) Disagreement over Taste, *Philosophical Perspectives* 156.3.433–439.
Lasersohn, Peter (2012) Compositionality and contextualism, *Linguistics and Philosophy* 35.2.171–189.

Lasersohn, Peter (2013) Non-World Indices and Assessment-Sensitivity, *Inquiry* 56.2-3.122–148.
Lewis, David (1972) General semantics, *Semantics of Natural Language*, 169–218, ed. by Davidson, Donald and Gilbert H. Harman, D. Reidel Publishing Co.
Lewis, David K. (1975) Adverbs of quantification, *Formal Semantics of Natural Language*, 3–15, ed. by Keenan, Edward L., Cambridge University Press.
Lewis, David K. (1979) Attitudes de dicto and de se, *The Philosophical Review* 88.513–543.
Lewis, David K. (1980) Index, context and content, *Philosophy and Grammar*, 79–101, ed. by Kanger, Stig and Sven Öhman, D. Reidel Publishing Co.
Ludlow, Peter (1999) *Semantics, Tense and Time: An Essay in the Metaphysics of Natural Language*. MIT Press.
Ludlow, Peter (2006) Tense, *The Oxford Handbook of Philosophy of Language*, 689–715, ed. by Lepore, Ernest and Barry C. Smith, Oxford University Press.
MacFarlane, John (2003) Future contingents and relative truth, *Philosophical Quarterly* 53.212.321–336.
MacFarlane, John (2005) Making sense of relative truth, *Proceedings of the Aristotelian Society* 104.53–73.
MacFarlane, John (2007) Relativism and disagreement, *Philosophical Studies* 132.1.17–31.
MacFarlane, John (2008a) Epistemic modals are assessment-sensitive, *Epistemic Modality*, ed. by Weatherson, Brian and Andy Egan, Oxford University Press.
MacFarlane, John (2008b) Truth in the garden of forking paths, *Relative Truth*, 81–102, ed. by García-Carpintero, Manuel and Max Kölbel, Oxford University Press.
MacFarlane, John (2009) Non-indexical contextualism, *Synthese* 166.2.231–250.
Moltmann, Frederike (2010) Relative truth and the first person, *Philosophical Studies* 150.2.187–220.
Montague, Richard (1968) Pragmatics, *Contemporary Philosophy: A Survey*, 102–122, ed. by Klibansky, Raymond, La Nuova Italia Editrice.
Montague, Richard (1970) Universal grammar, *Theoria* 36.373–398.
Montague, Richard (1973) The proper treatment of quantification in ordinary English, *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, 221–242, ed. by Hintikka, K. J. J., J. M. E. Moravcsik, and P. Suppes, D. Reidel Publishing Company.
Morgan, Jerry L. (1970), On the criterion of identity for noun phrase deletion. *Papers From the Sixth Regional Meeting of the Chicago Linguistic Society*, 380–389, Chicago Linguistic Society.
Mount, Allyson (2008) The impurity of "pure" indexicals, *Philosophical Studies* 138.2.193–209.
Neale, Stephen (1990) *Descriptions*. MIT Press.
Neale, Stephen (2004) This, that and the other, *Descriptions and Beyond*, 68–182, ed. by Reimer, Marga and Anne Bezuidenhout, Oxford University Press.
Nunberg, Geoffrey (1993) Indexicality and deixis, *Linguistics and Philosophy* 16.1–43.
Parsons, Josh (2011) Assessment-Contextual Indexicals, *Australasian Journal of Philosophy* 89.1.1–17.
Parsons, Terence (1980) Modifiers and quantifiers in natural language, *Canadian Journal of Philosophy* Supplementary Volume VI.29–60.
Parsons, Terence (1990) *Events in the Semantics of English: A Study in Subatomic Semantics*, 19 ed. MIT Press.

Partee, Barbara H. (1973) Some structural analogies between tenses and pronouns in English, *Journal of Philosophy* 70.18.601–609.
Perry, John (1977) Frege on demonstratives, *The Philosophical Review* 86.4.474–497.
Perry, John (1986) Thought without representation, *Supplementary Proceedings of the Aristotelian Society* 60.137–152.
Perry, John (2001) *Reference and Reflexivity*. CSLI Publications.
Predelli, Stefano (1998) I am not here now, *Analysis* 58.107–115.
Predelli, Stefano (2005) *Contexts: Meaning, Truth and the Use of Language*. Clarendon Press.
Prior, A. N. (1957) *Time and Modality*. Clarendon Press.
Prior, A. N. (1959) Thank goodness that's over, *Philosophy* 34.128.12–17.
Putnam, Hilary (1975) The meaning of 'meaning', *Language, Mind and Knowledge*, 131–193, ed. by Gunderson, Keith, University of Minnesota Press.
Richard, Mark (2004) Contextualism and relativism, *Philosophical Studies* 119.215–242.
Richard, Mark (2008) *When Truth Gives Out*. Oxford University Press.
Schaffer, Jonathan (2009) Perspective in taste predicates and epistemic modals, *Epistemic Modality*, 179–226, ed. by Egan, Andy and Brian Weatherson, Oxford University Press.
Schein, Barry (1993) *Plurals and Events*. MIT Press.
Schlenker, Philippe (2004) Context of thought and context of utterance: A note on free indirect discourse and the historical present, *Mind and Language* 19.3.279–304.
Sidelle, Allan (1991) The answering machine paradox, *Canadian Journal of Philosophy* 21.4.525–539.
Snyder, Eric (2013) Binding, Genericity, and Predicates of Personal Taste, *Inquiry* 56.2-3.278–306.
Stalnaker, Robert (1970) Pragmatics, *Synthese* 22.1-2.272–289.
Stalnaker, Robert (1979) Assertion, *Pragmatics*, 315–332, ed. by Cole, Peter, Academic Press.
Stalnaker, Robert (1999) *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford University Press.
Stanley, Jason (2000) Context and logical form, *Linguistics and Philosophy* 23.4.391–434.
Stanley, Jason (2011) *Know How*. Oxford University Press.
Stanley, Jason and Szabó, Zoltán Gendler (2000) On quantifier domain restriction, *Mind and Language* 15.219–261.
Stephenson, Tamina (2007a) Judge dependence, epistemic modals, and predicates of personal taste, *Linguistics and Philosophy* 30.4.487–525.
Stephenson, Tamina (2007b) *Towards a Theory of Subjective Meaning*, MIT dissertation.
Stephenson, Tamina (2008), Relativism and the *de se* interpretation of PRO. *Proceedings of the Thirty-Seventh Annual Meeting of the North East Linguistic Society*, vol. 2, 387–400, ed. by Schardl, Anisa, Walkow, Martin, and Abdurrahman, Muhammad, Graduate Student Linguistic Association.
Stephenson, Tamina (2010) Control in centered worlds, *Journal of Semantics* 27.4.409–436.
Takahashi, Daiko (2008) Noun Phrase Ellipsis, *The Oxford Handbook of Japanese Linguistics*, 394–422, ed. by Miyagawa, Shigeru and Mamoru Saito, Oxford University Press.
Talmy, Leonard (2000) *Toward a Cognitive Semantics. Vol.1: Concept Structuring Systems*. MIT Press.
Tarski, Alfred (1944) The semantic conception of truth and the foundations of semantics, *Philosophy and Phenomenological Research* 4.341–375.

Thomason, Richmond H. (1979) Home is where the heart is, *Contemporary Perspectives in the Philosophy of Language*, 209–219, ed. by French, Peter A., Theodore E. Uehling, Jr., and Howard K. Wettstein, University of Minnesota Press.
van Fraasen, Bas (1967) Meaning relations among predicates, *Noûs* 1.2.161–179.
von Fintel, Kai (1994) *Restrictions on Quantifier Domains*, University of Massachusetts dissertation.
Weatherson, Brian (2002) Misleading indexicals, *Analysis* 62.4.308–310.
Weatherson, Brian (2009) Conditionals and indexical relativism, *Synthese* 166.2.333–357.
Weir, Andrew (2012) Left-edge deletion in English and subject omission in diaries, *English Language and Linguistics* 16.1.105–129.
Williams, Edwin (1980) Predication, *Linguistic Inquiry* 11.1.203–238.
Williamson, Timothy (1994) *Vagueness*. Routledge.
Wittgenstein, Ludwig (1961) *Tractatus Logico-Philosophicus*. Routledge & Kegan Paul.
Wurmbrand, Susi (2014) Tense and aspect in English infinitive, *Linguistic Inquiry* 45.3.403–447.
Zagona, Karen (2002) *The Syntax of Spanish*. Cambridge University Press.
Zwicky, Arnold M. and Pullum, Geoffrey K. (1983) Deleting named morphemes, *Lingua* 59.2-3.155–175.