ON WORDS:

AN ESSAY ON BELIEFS, BELIEF ATTRIBUTIONS AND THE ONTOLOGY OF LANGUAGE

By

DEVON BELCHER

B.A., REED COLLEGE 1989

M.A., UNIVERSITY OF COLORADO, 1999

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Philosophy

2005

This thesis entitled: On Words: An Essay on Belief, Belief Attributions and the Ontology of Language

Written by Devon Belcher

has been approved for the Department of Philosophy

Michael Tooley
Co-chair

Robert Hanna
First Reader

Date: July 13, 2005

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

HRC protocol# _____

Belcher, Devon (Ph.D., Philosophy)

On Words: An Essay on Belief, Belief Attributions and the Ontology of Language

Thesis directed by Professor George Bealer and Professor Michael Tooley

The received view of words holds that words are concrete, physical objects such as marks on paper or sounds (patterns of moving air), and that these physical objects have the meanings they do solely as a matter of contingent convention. So, for example, the word 'cat' is a pattern of ink spots on this page, and although it happens to refer to the species *Felis Domesticus*, the received view holds that it could just as easily have referred to the family *Homaridae* (the lobster). This work argues that the received view, compelling though it appears at first, is false. Certainly there are marks on paper, or sound waves, and we establish conventions that contingently associate these marks with various meanings. And sometimes we even call those marks and sounds ‘words’. I refer to such things as "thin words". But such thin words, I hold, are not the central objects involved in linguistic communication. Moreover, while such words enter into many of our thoughts, they are by no means the only words involved in our thoughts. Rather, I shall argue, languages are composed of what I call "thick words" - entities which *have their meanings essentially*. Since physical entities like sounds and marks are (intuitively) merely *contingently* associated with their meanings, it follows that thick words must be something more than such physical entities or physical entity-types. Thick words are instead, I argue, abstract objects with strong similarities to (albeit important

differences from) abstract objects like numbers or properties.

I offer a two-pronged argument for establishing this radical claim. First, I argue that several notable puzzles about belief attributions can be solved by referring to the words that are involved in the attributions. But this solution requires a theory of thick words. Second, I argue that our common-sense intuitions about the individuation of words actually supports a theory of thick words. The final section of this work outlines a positive theory of thick words as well as its application to several issues in the theory of attitude ascriptions.

Acknowledgments

My first and original debt is to Dr. Jerry DeWitt, my 10th grade Honors History teacher. I was somewhat of a smutty-minded clown, and one day Dr. DeWitt told me that I should try to use more colorful language. 'For instance', he said, 'instead of saying 'wet dream', say 'moist reverie'.' And thus was born a life-long fascination with words.

Special thanks are due to Sheralee Brindell, my "unindicted co-conspirator", who has stood by me all along, and over the course of 15 years matured both my writing and my philosophical position. I am afraid her attempt to mature *me* has been not so successful, however.

And although it has always seemed a bit maudlin to me to thank one's parents, I have to acknowledge a debt to both Don and Marie Belcher. I suppose it takes 30 or 40 years to realize how much one has been shaped by one's parents. Also, what little bits of humor have snuck past my advisors into this thesis, originated in Marie. And I've always kept them in mind as my audience - 'if Mom and Dad can't understand this, I better re-write it.'.

Thanks also to my friends who endured constant yammering about my work, and occasionally provided insight: Eric "Sparky" Anderson, Chad Carmichael, Brendan Jackson, Paul Studtmann and Simon Sparks.

Finally, thanks to my many advisors - George Bealer, Michael Tooley, Bob

Hanna, Graham Oddie and Daniel Stoljar. I wish you could all be here for this, but some of you have moved on.

Table of Contents

‘A man infallibly knows, as soon as he ever has them in his mind, that the ideas he calls *white* and *round* are the very ideas they are, and that they are not other ideas which he calls red or square. . .and if there ever happen to be any doubt about it, it will always be found to be about the names, and not the ideas themselves, whose identity and diversity will always be perceived as soon and as clearly as the ideas themselves are. . .’

- John Locke, An Essay Concerning Human Understanding, Book IV, Ch1, S4.

Introduction and Overview

Rational agents represent the world around them not only via internal states ("beliefs"), but also using external, shareable entities like words, sentences and other such signs - the elements of symbolic communication. These external entities that agents create (or exploit) are themselves part of the world and may enter into representations of the world, consciously or otherwise. Agents may uncontroversially have beliefs about words, such as my belief that the word ‘Words’ appears in the title of this work, and that it has 5 letters. Let us call these “metalinguistic” beliefs. But agents may also have beliefs involving words in other, more subtle ways that I will presently discuss.

It is undeniable that we have such metalinguistic thoughts, or thoughts about the words we use, just as it is undeniable that we have devices in natural language to represent words themselves rather than the things those words normally refer to. In the English language, for example, these devices include quotation marks, as in the expression

'cat'

as well as the use of italics as in

cat,

or the use of displaced text. But it is not obvious what sorts of entities words are (for

surely, when we think of ‘cat,’ we are thinking of a word). The received view of words holds that words are concrete, physical objects such as marks on paper or sounds / patterns of moving air (or abstract types of such physical obejects), and that these physical objects have their referents solely as a matter of contingent convention. So, for example, the word 'cat' is a pattern of ink spots on this page. Moreover, the word ‘cat’ refers to the species *Felis Domesticus*, but (so the received view holds) we could just as easily have decided to use that very same word to refer to the family *Homaridae* (the lobster), the number 3, or the Queen of Liechtenstein. Closely related to this view of words as mere physical tokens is a view of words as *types* of physical objects -that is, abstract entities tokened by phonological or orthographic objects; and the further Quinean view which reduces such types to the set of its tokens.

I believe that the received view, compelling though it appears at first, is false in important ways. Certainly there are marks on paper, or sound waves, and we establish conventions that contingently associate these marks with various referents. And sometimes we even call those marks and sounds ‘words’; I will refer to such things as "thin words.” But such thin words, I hold, are not the central objects involved in linguistic communication. Moreover, while such words enter into many of our thoughts, they are by no means the only words involved in our thoughts. In fact, they are rarely the words that enter into the sorts of thoughts and beliefs we will be interested in. Rather, I shall argue, languages are composed of what I call "thick

words" - entities which *have their referents necessarily*[1]. Since physical entities like sounds and marks are (intuitively) merely *contingently* associated with their referents, it follows that thick words must be something more than such physical entities or physical entity-types. Thick words are instead, I argue, abstract objects with strong similarities to (albeit important differences from) abstract objects like numbers or properties[2]. They are different from objects like numbers and properties, however, in that they are created objects, and hence they exist only contingently and mind-dependently.

This work is a two-pronged approach for establishing this thesis about words. The first prong involves a class of semantic puzzles about belief and belief attribution called the "substitutivity puzzles." These substitutivity puzzles appear to be neatly solved by holding, as I have suggested earlier, that our beliefs sometimes involve words in subtle and not-so-obvious ways. In fact, we often mistake beliefs that involve words in this way for beliefs that involve no words but instead involve only the things those words refer to. Let us refer to the former sorts of belief as *lexically augmented beliefs* and the latter sort as *purely objectual beliefs*. I argue that lexically

---

[1]Henceforth, I will use double quotes as in "cat" to indicate thick words, and I will leave single quotes 'cat' ambiguous as to whether they indicate thick or thin words. This ambiguity is deliberate, to mirror the ambiguity which I will argue is present in common usage. Of course, my account entails that there are actually *two* homonymous devices that look like single quotes ' ', and not a single symbol with two distinct usages.

[2]Just what sorts of abstract objects they are may, for the moment, be left unspecified. If the reader wishes, he may for the moment take them to be abstracted speaker intentions; or classes of conventions, or n-tuples (whether ordered or unordered) consisting of {phonotype, orthotype, referent}.

augmented beliefs must involve thick words if they are to solve the substitutivity puzzles[3]. So there is, I think, a theoretical need for thick words, in that they are required by the best solution to a vexing set of theoretical puzzles.

Nonetheless, one might think that thick words were just so counterintuitive and implausible as to rule out any solution involving them. But this turns out to be not so, and there is a considerable amount of evidence to support this conclusion. Thus, the second prong of this work is a direct discussion of words themselves. For the received view of words typically cites intuitions from what I call *folk linguistics* in its defense - where by *folk linguistics*, I mean the intuitions and beliefs about our practices and conventions surrounding words and our common usages of them. But these intuitions often cited in favor of thin words, I shall argue, are at best ambiguous. In fact, our ordinary folk-linguistic intuitions, surprisingly enough, provide compelling evidence that there is an important set of words used in everyday communication that are indeed thick.

Thus, this work neatly divides into two main parts. The first part is a discussion of the substitutivity puzzles and a critical review of some major theories of belief which respond to these puzzles. The central puzzles appear to show that, given otherwise unobjectionable claims about the nature of mental content, agents have beliefs that they (arguably) do not have. For example, consider the following puzzle (inspired by Benson Mates' famous puzzle about synonymy):

---

[3]I suggest this nomenclature because, as I will make clear, I do not think all such beliefs are metalinguistic. That is, although these beliefs involve words, they are not always *about* the words, and hence not always metalinguistic in the usual sense.

(1) Belief (and belief attribution) relates agents to things believed ("propositions").
(2) The content of a proposition is a function of the content of its constituents - properties, relations and particulars.
(3) Sparky believes that chewing is chewing
(4) masticating = chewing[4]
(5) Thus, Sparky believes that masticating is chewing.

The puzzle is that (1-4) seem to entail (5), but surely Sparky might believe the triviality that chewing was chewing and yet (at least apparently) fail to believe that chewing was masticating. In the first part of this work, I will be concerned with a critical review of several leading philosophical theories of belief and belief attribution. These theories attempt to block the inference to (5), or to show that (5) is not in fact the falsehood that it seems to be. I think the best of these theories rest on the claim that these problematic beliefs involve thick words. Let me explain.

One common response we will see is to hold that in the problem cases, the contents of our beliefs are metalinguistic - that is, propositions about words and their relations to the things (objects, propositions, etc) that these words express. For example, one might hold that (5) is in fact true, and Sparky *does* in fact believe that masticating is chewing, but he does not believe the related metalinguistic proposition:

ML): the word 'masticating' and the word 'chewing' refer to the act of chewing.

I think that this type of solution, while flawed, nonetheless embodies a certain insight.

---

[4]Obviously, many of these premises may be challenged. For example, one might hold that (1) is false because belief relates agents to mental representations which in turn represent propositions; or one might hold that belief relates agents to sentences. Similarly, one might deny (2) and hold that the constituents of propositions are not properties, etc., but rather mentalese entities like functional states, or concepts, or words. The point is that the puzzle here provides motivation for each of these theoretical claims.

That is, sometimes our beliefs about the world come through an indirect source - our beliefs are only via the words used to tell us about the world. It seems possible to have all sorts of beliefs about the world solely in virtue of being given words. Words are ways that objects are presented to us, just as objects are presented via descriptive properties or concepts. But the metalinguistic solutions as they stand are flawed. In the first place, they do not even appear to have the correct form. The belief that seems to be mentioned in (5) - that masticating is chewing - simply involves two properties and the identity relation. But the metalinguistic analysis ML mentioned above involves one property, two words, and the reference relation. It does not seem like a good candidate for an analysis of a report of (4). For small children and other unsophisticated agents may surely believe that masticating is chewing without having such complex metalinguistic beliefs; and even highly competent speakers and thinkers may deny that (ML) is really what they are mean when they claim that masticating is chewing. More importantly, the metalinguistic account changes the modal value of one's beliefs. That is, if words are thin and have their referents contingently, as the received view would have it, then beliefs about necessarily true identity statements (that masticating is chewing) would be beliefs about *contingently true linguistic facts* (that two words happen to have the same referents). And such a change in the modal values of our beliefs has unacceptable consequences for one's theory of mental acts which depend on modal values - acts such as inferring and reasoning .

In the final chapter of part I, I sketch a theory of belief and attribution that

addresses these substitutivity puzzles. Sparky's beliefs are lexically augmented propositions, or propositions that have words themselves as constituents, in a novel fashion to be developed here. This claim, however, must be augmented with an account of words - an account that will preserve the modal value of beliefs once they have been analyzed in this way. The theory of thick words is this theory. Since thick words have their referents necessarily, both the beliefs

> O) that masticating is chewing, and
> ML) that the (thick) word "masticating" and the (thick) word "chewing" refer to the property (or act)[5] of chewing.

are necessarily true, since the thick words "masticating" and "chewing" both necessarily refer to the act of chewing.

Once we have seen the need for thick words, we can employ them in an alternative account of the (1) content, (2) modal value, and (3) form of such propositions involving words**.** On this account, the thick words "masticating" and "chewing" are constituents of the proposition that is believed when Sparky believes that masticating is chewing. But the proposition is not metalinguistic, in that it is not *about* the words and their relations to the properties (or actions) they refer to. These linguistic relations are not components of the proposition or its analysis at all. Rather, the proposition Sparky believes is a proposition relating the property or action of masticating to the property or action of chewing - by means of the words that present

[5]For the rest of this work, I will assume a simplified ontology of propositions, individuals, particulars and relations, equating entities such as *actions* with properties, and so forth. This is simply for simplicity's sake and should not affect any of the claims made here.

these properties or actions to us.

The second part of the dissertation will be a discussion of the ontology of words themselves. For if part one of this work establishes the need for thick words, it remains unclear just what sorts of entities words are - that is, it remains unclear what sorts of entities could have their referents necessarily. In part II, I will draw from "folk linguistics" - both to motivate the theory of thick words and to flesh out the details of such a theory. As I mentioned earlier, proponents of the received view of words typically cite intuitions from folk linguistics - such as the intuitions that we could have decided to use a word to mean something other than what it actually means, or that words can change their meanings over time, etc. But these are really intuitions about shapes and sounds, and to make the further claim that words are merely shapes and sounds simply begs the question. More importantly, folk linguistics itself supplies a preponderance of intuitions that support the thick words thesis. For example, we have a host of intuitions surrounding the way we individuate homonyms and cognates, the way we treat the phenomena of mispronunciation and malapropism, and the way we identify written and spoken presentations of the same word. I will demonstrate that these intuitions all suggest that words are thick.

Furthermore, folk linguistics provides a host of intuitions to guide us in the determining the structure of thick words. For example, words are created entities; words thus exist merely contingently; words have spellings and pronunciations as *presentations.* I will examine in depth two views of thick words - the sequence view, on which words are sequences of <orthotype, phonotype, referent> and the causal-

historical view, on which words are tree-like entities. In the end, I will defend a more thoroughgoing nonreductionist view of words. The concluding chapter of part II will then use the theory of thick words to give a more substantial answer to the substitutivity problems of part one, and finally, to suggest further work and areas of investigation for the theory of thick words.

Part I

Beliefs and Words

Beliefs and Words

Overview

The first half of this work (Part I) is a discussion of belief and belief attribution / belief reports. My aim here is to motivate three main theses, though not in the following order. First, belief is a simple, two - place relation that holds between a cognitive agent ("believer") and what is believed (a "proposition"). Thus the structure of belief mirrors the surface structure of an ordinary belief attribution such as

Sparky believes that Hesperus is bright

Correspondingly, belief attributions have no hidden elements such as "hidden indexicals," "unarticulated constituents," or covert references to further "modes of presentation" of propositions.

Second, the propositions believed are finely individuated (or what I shall call "fine-grained") entities. To begin with, two propositions may have the same truth conditions, yet be distinct entities. Furthermore, these propositions have their structures necessarily; but as we shall see, even structure cannot individuate propositions finely enough to account for our intuitions about beliefs and belief reports.

The third (and most central) thesis I intend to motivate is that these propositions sometimes involve words - specifically, they involve *thick words*, abstract lexical entities that have their referents necessarily. These thick words are, to borrow a common terminology, modes of presentation of an entity, just as ordinary descriptive

properties or concepts are modes of presentation of entities. It is these words that serve to individuate propositions finely enough to account for our intuitions about belief and belief reports. Moreover, they "appear" in propositions in much the way ordinary modes of presentation might "appear" in propositions. This insight, together with the first two claims, motivates a theory of belief and belief attribution sketched at the end of Part I.

In arguing for these three claims, I shall canvass a number of current theories of belief and belief attribution. These theories of belief are driven by intuitions about the *intensionality* of belief and the contents of belief. That is, we *seem* to be able to have beliefs about a an object (or property), without having corresponding beliefs about identical objects (or properties). So, for example, even though the property of masticating just *is* the property of chewing, it *seems* possible for an agent to believe that, for example, masticating in public is rude, while at the same time not believing that chewing in public is rude. And this needs to be reflected in an account of belief attribution. Even though we think that the truth (and content) of a belief report such as:

A believes that Fa

is a function of the content of its parts, we do not appear to be able to substitute synonymous predicates or names within the scope of the verb 'believes' (inside the *that*-clause)[6] while preserving truth. We shall refer to these sorts of puzzles as the

[6]And of course, if the names (or predicates) are synonymous, they have the same referent. Thus, if the truth of the report is a function of the referent of its parts, then the truth of the report containing one name should just be the same as the truth

substitutivity puzzles

Since at least the time of John Locke (as I indicated in the frontispiece of this work), philosophers have been attracted to the idea that the substitution-resisting beliefs and attributions involve linguistic elements somehow, whether in addition to the objects they represent, or instead of them (Locke's original claim is vague on the subject). This approach is made all the more attractive when one considers that sometimes the only knowledge we have of individuals is knowledge mediated by these linguistic elements. Most Americans, for example, know nothing of Cicero other than his name[7]. And in other cases - such as the masticating/chewing cases, or the infamous schoolyard taunt 'your epidermis is showing!,' the source of confusion seems obviously linguistic. And so I think that, in at least some of the cases of the substitutivity puzzles, the correct solution is one incorporating this insight - that the contents of our beliefs sometimes involve words without our realizing it.

Each of the current theories of belief that I shall discuss incorporates this insight in some way. These theories all offer different accounts of the nature of the contents of belief and the nature of the belief relation itself. And so I think it will be natural to proceed in a certain order, from the simplest theories to the most complex, in terms of

---

of the report which substitutes a synonymous name or predicate. But these puzzles appear to show that this is not so.

[7]Even amongst those who know some fact(s) about Cicero, few know more than that he was a Roman or an orator. But this might well be true of their beliefs involving Tully as well. Very few people have any descriptive content associated with Tully other than that he was an orator. So it seems that the only way most people differentiate their ideas involving Cicero from their ideas involving Tully is through the names.

both the contents of belief and the belief relation itself. I shall start with Robert Stalnaker, who treats belief as a simple two-place relation, and the contents of belief as unstructured propositions. We shall then move on to Stalnaker and Chalmers' elaboration of this claim, the so-called "two-dimensional proposition." The failure of these theories in the face of general arguments motivates my conclusion that the contents of belief - propositions - are essentially structured entities.

We then move on to a consideration of theories involving such structured entities. Even if one holds that the contents of belief are structured entities, substitutivity problems still arise. One approach to this is to somehow incorporate linguistic elements directly into the propositions, whether by treating names as disguised metalinguistic descriptions (such as *the thing known by the name 'Cicero'* or *the property referred to by the predicate 'masticating'*), or by simply "tacking the words onto" the propositions. So, for example, I shall look at Larsen and Ludlow's theory of Interpreted Logical Forms (ILFs), on which the content of belief is the syntactic tree corresponding to a certain sentence (the sentence given in the that-clause of a belief attribution), where each nodes of the tree is annotated with its semantic value. Most importantly, the terminal nodes of the tree - what we ordinarily think of as the constituent words of the sentence - are annotated with the objects and properties those words refer to. Such contents are fine-grained indeed; they are at least as finely individuated as the sentences we use to express them[8]. Nonetheless, they are too fine-

[8]Although Mark Richard's theory of "Russellian Annotated Matrices" or RAMs might naturally fit in here, I shall save a discussion of RAMs for the second part of this work. For RAMs invoke thick words in a way that makes Richard's

grained. There is no account of how speakers of different languages may share beliefs, and no account of how we may make true attributions of speakers of other languages.

Finally, I shall discuss theories involving more complex belief *relations*. On these accounts, belief itself is a more complex three-place (or greater!) relation, involving an agent, a proposition, and some further entity. Perhaps this further entity is another proposition. Perhaps it is a "mode of presentation" of the believed proposition, and perhaps it is a purely mental "idea" or "notion" which bears some relation to a proposition. Closely related to such approaches are those that treat belief attribution - although not belief itself - as a complex relation between an agent, a proposition, and something (such as a further proposition) conversationally implied by the attribution. And each of these approaches commonly invokes the claim that this further "mode of presentation" or "conversational implication" involves words in some way. But despite the added complexity and subtlety of these theories, they too suffer familiar problems. I suggest that at the root of many of these problems is the fact that the new entities involved - the "notions," "modes of presentation," or "conversational implicatures," are themselves inadequate. They are of the wrong "granularity" - they either individuate beliefs too coarsely or too finely. What is needed is a new sort of entity, one which is public and shareable (even shareable across linguistic contexts), but nonetheless very finely individuated. Thick words are these entities.

---

theory very similar to some of the theories of thick words discussed in the second part.

Although I believe that there are many ways that thick words might plausibly be involved in the contents of belief, the final chapter in this part of the work will sketch out what I think to be the most natural theory of beliefs and propositions involving thick words. In addition to addressing many of the vexatious substitution problems we have discussed, this theory has the advantage of simplicity and intuitiveness: for example, it invokes no unarticulated constituents or hidden structures, and requires no complex tacit intentions or implicit beliefs.

Without further ado, let us now investigate how words might enter into beliefs.

Chapter 1

Coarse-Grained Mental Content and the Inferring Argument

Overview of the Chapter

In this chapter, I investigate a theory of content - contents of sentences, contents of beliefs, etc - on which such contents or propositions are identified with sets of possible worlds. This identification is itself an implementation of some broader view about minds and beliefs; such as the view that beliefs are partitions of logical space, or that beliefs are extensional dispositions to interact with the world, or whatever. Such a view of propositions is commonly referred to as coarse-grained because it individuates propositions so coarsely. For example, on such an account the proposition *that Hesperus is Phosphorus* cannot be individuated from the proposition *that water is* H_2O. To deflate obvious objections, coarse-grained theorists often claim that what appears to be an ordinary belief about water is often a disguised belief about the word 'water' and what it refers to. We shall call this the metalinguistic rephrasal strategy.

The central argument of this chapter is that the coarse-grained metalinguistic strategy, while it ameliorates some of the obvious problems faced by coarse-grained theorists, nonetheless fails to accommodate many mental phenomena. In particular, the metalinguistic rephrasal strategy, at least as commonly understood, cannot account for the truism that rational agents make valid inferences amongst their beliefs. In this chapter, I will introduce what I call the "argument from inferring," and

apply it to the general metalinguistic rephrasal strategy. Then, I will show how the argument can be given against some currently popular, and theoretically sophisticated versions of that theory. The end result will be that we must abandon the coarse-grained theory of propositions; but the metalinguistic strategy may remain available to more fine-grained theories of propositions which we will investigate later in our inquiry.

A) Contents and Rephrasals

Our first question, then, is this: what is the content of a belief ? A family of theorists, including Carnap, Hintikka, Montague and (more recently) Robert Stalnaker, has long equated propositions with their truth conditions. These theories are often motivated by a position which holds that 'belief and other cognitive states are [best explained] as capacities and dispositions to interact with the extra-linguistic world.'[9] In cases where these extra-linguistic facts P and Q are necessarily equivalent (have the same truth conditions), a disposition to bring about P appears to be just a disposition to bring about Q. If this is so, then identifying the contents of beliefs with the truth conditions of those beliefs will ensure the identity of necessarily equivalent beliefs, as the dispositional account suggests. The formal device for representing this picture of belief is to treat the contents of the attitudes - propositions - as functions from possible worlds to truth values. Alternatively (and more or less equivalently), propositions are the sets of possible worlds which "make them true" . The proposition

[9]Stalnaker [87], p177

that the cat is on the mat, on this account, is the set of possible worlds in which the cat is on the mat. Let's call this a coarse-grained account of propositions.

The immediate consequence of a coarse-grained account is that all necessarily equivalent propositions are identical. The proposition that the cat is on the mat is the same proposition as the proposition that the mat is under the cat. On the supposition that belief is a dyadic relation between an agent and a proposition[10], this theory leads to some very counterintuitive results. Suppose Sparky believes that recursiveness is recursiveness. This entails that Sparky believes that recursiveness is effective calculability. But surely one may truly report the former while truthfully denying the latter; after all, this identity was taken to be one of the great discoveries of the foundations of mathematics. Similarly, suppose Sparky believes that -(A v B). It seems that Sparky might fail to believe the logically equivalent proposition that (-A & -B); however, on the coarse-grained account, he is related by belief to one and the same proposition.

Moreover, the coarse-grained account forces us to attribute incoherent beliefs to otherwise rational cognitive agents. Since all necessarily false beliefs are identical, to believe one is to believe them all. Yet when we make an attribution such as

Sparky believes that Hesperus is Mars,

'we do not seem to be attributing to [Sparky] the necessarily equivalent belief that (for example) Hesperus is non-self-identical.'[11] In order to avoid these results, many

---

[10]I restrict my arguments in this chapter to such a dyadic theory of the belief relation.

[11]Stalnaker, [87A] p179

coarse-grained theories utilize a rephrasal strategy. In attributions of (apparently) necessarily false beliefs, the beliefs attributed are not really necessarily false propositions, but rather, related and contingently false propositions. These beliefs we attribute to an agent are actually beliefs about the relation of sentences to the propositions they (contingently) express. Such a rephrasal takes one of two forms. First, on the sentential metalinguistic rephrasal, attributions of the form:

A believes the proposition that Fa (where Fa is necessarily false)

are rephrased as attributions of the form:

A believes the proposition that the sentence 'Fa' expresses a truth[12].

On the other hand, on the sub-sentential metalinguistic rephrasal, they are rephrased as being of the form:

A believes that the object referred to by ‘a’ is F.[13]

Let us call these the basic Metalinguistic Rephrasal Strategies[14]. Now the idea of rephrasing problematic beliefs simply to avoid problems might appear *ad hoc* to some, especially if there was no independent motivation for the rephrasal. But, as

---

[12]Alternatively, attributions of the form A believes that P are rephrased as being of the form

A believes that 'Fa' expresses the proposition that Fa.

[13]Again, an alternative is:

A believes that ‘a’ has the property expressed by the predicate ‘F,’

or even

A believes that the object referred to by ‘a’ has the property expressed by the predicate ‘F,’

depending on what one thinks the locus of A’s confusion is.

[14]See, for example, Stalnaker 1987: "If I do not recognize some complicated truth-functional compound to be a tautology, and so doubt whether what it says is true, this can naturally be explained as doubt or error about what the sentence says." (emphasis mine)

Stalnaker has pointed out, we do often have independent motivation for rephrasal. For example, we can appeal to the principle of charity, which tells us that we should always interpret a speaker on the assumption that he is being rational. Clearly, (in most cases) to attribute a belief that Hesperus is not Hesperus to a speaker is to treat that speaker as being irrational. The principle of charity suggests that we rephrase the attribution in such a way as to attribute a coherent belief to the agent. Of course, this new belief must be such that it is relevantly similar to the belief that the original attribution seems to imply; and we should have a method for determining from a given utterance exactly what metalinguistic belief is *really* being expressed. The rephrasal strategies we are considering do just that.

Before we consider objections, there is an ontological point to make about each strategy. That is, each strategy discussed here is committed to the existence of heavily abstract objects, that is, propositions. These propositions involve purely linguistic entities, that is, words or sentences. In this way, the strategies are unlike positivistic reductions such as (for example) Carnap's,[15] which reduces even propositions to purely physical objects, namely, sentences (or sentence-types). On Carnap's account, belief is simply a relation (such as *a disposition to accept*) which an agent bears to a sentence. On the metalinguistic accounts we are currently considering, however, belief is irreducibly a relation to a proposition, and of course, in some cases (but not all, and in fact a distinct minority), those propositions involve sentences. Sometimes we believe that Hesperus is bright; sometimes we have beliefs about the ways that

[15]Carnap [1947]

people represent their beliefs that Hesperus is bright. The former belief is represented by a set of possible worlds in which Hesperus is bright; the latter belief is represented by a set of worlds in which people use a certain string of words (rather than another string of words) to represent the fact that Hesperus is bright.

Let us now consider the sub-sentential metalinguistic strategy. One can quickly see that changing the modal value of the proposition believed will create problems. Consider the attribution,

Kripke believes that water is H_2O.

On the sub-sentential rephrasal strategy, this becomes something like:

> Kripke believes that the substance referred to by the term 'water' is identical to the substance referred to by the term 'H_2O.'[16]

So far so good - on this account, Kripke believes a merely contingent fact about language (and if it is false, it is merely contingently false). Moreover, the set of worlds in which 'water' refers to the same stuff as 'H_2O,' is distinct from (for example) the set of worlds at which 'Hesperus' refers to the same thing as 'Phosphorus.' This gives the coarse-grained theory the intuitive result that one can believe that water is H_2O without, say, also believing that Hesperus is Phosphorus.

Unfortunately, the sub-sentential strategy gives incorrect analyses when dealing with propositions involving embedded modal terms. Consider an example similar to

---

[16]Clearly, *the substance referred to by the term 'water'* must be read *de dicto* and not *de re*, or the content of belief is still a necessary truth, and rephrasal has gained us nothing.

the one just given. Kripke famously believed that all identities are necessary, and so it is also true that:

> Kripke believes that necessarily water is H_2O.

After all, this is a famous thesis in the philosophy of language, contrasted with earlier beliefs that water is only contingently H_2O. But the sub-sentential rephrasal yields:

> Kripke believes that necessarily, the substance referred to by the term 'water' is identical to the substance referred to by the term 'H2O.' [17]

This is surely a false attribution. As far as I am aware, very few people believe that words have their referents necessarily; Kripke is not one of them. At any rate, one might surely have beliefs about the necessary identity of water and H_2O without believing such radical claims about words having their referents necessarily.

The sentential rephrasal, on the other hand, does not suffer this problem, for it yields simply:

> Kripke believes that the sentence 'Necessarily, water is H_2O' expresses a truth.

Finally, the sub-sentential rephrasal strategy is unable to account for a large class of (seemingly) necessarily false beliefs. For consider an agent who understands quite well what the symbols '&,' 'v,' and so on, express (we may imagine that this agent

---

[17]One might suppose that the correct rephrasal was really:
Kripke believes that necessarily, the substance referred to by the term 'water' bears the relation expressed by the predicate 'is necessarily identical to' to the substance referred to by the term 'H_2O.'

This added complexity seems unwarranted, however; for Kripke appears to understand perfectly the concept of necessary identity. The locus of his confusion is with water and H2O, not the relation they bear to one another.

has correctly answered questions about the associated truth tables, has given correct answers about the truth values of simple sentences involving these symbols, etc.). Such an agent might nonetheless erroneously claim that, for example, that the distribution principle:

A & (B v C) <-> (A& B) v (A & C)

is not a tautology. Intuitively, however, this cannot be explained as some error about what the individual words express, but rather about what the sentence as a whole expresses, in virtue of its structural/ syntactic properties.

With these objections in mind, I hold that if the metalinguistic rephrasal strategy is to be adequate at all, we must use the sentential version. Now that we have outlined the basics of the coarse-grained theory of propositional content, and shown how it may defend itself against some obvious objections, I would like to take a brief detour and discuss some of the work that a theory of propositions is supposed to do. After we consider the role that propositions must play in any theory of our mental lives, we shall return to the coarse-grained theory of propositions and examine whether such a theory can be consistent with such an account of our mental lives. Not surprisingly, the answer is 'no'- although the discussion of coarse grained propositions will provide us some general insights that we will apply over the course of this inquiry.

B) Propositions and inferring

One of the central tasks of philosophical psychology is to provide an account of the way agents revise and extend their stores of beliefs. Rational agents extend their

stocks of beliefs not only through experience, but through logical inference as well. We reason from the beliefs we have to add new beliefs logically implied by them. Any theory that fails to accommodate at least the appearance of this ubiquitous phenomenon is seriously flawed.

To see what it is we are interested in about inferring, consider Sparky, an ordinary person. Sparky is a rational agent and as such is constantly changing his mental states in light of new information. For our purposes, we may suppose that the relations between these changing states are causal relations (though perhaps not ordinary causal or quasi-causal relations). Corresponding to many of these changing states are changing particular contents - or, as I will assume, propositions. And logical relations between contents are not plausibly causal relations. The proposition that the cat is on the mat does not cause the proposition (e.g.) that something is under the cat. Rather, these propositions are analytically connected. In certain distinctive cases, when there is an appropriate connection between these states, we say that Sparky has made an inference. All too often, Sparky (like the rest of us) does not change belief states in a way warranted by the logical or analytic connections amongst propositions. Sparky may change from a state of believing that the cat is on the mat to a state of believing that the cat is sleeping. In this case, Sparky made an erronious inference. That is, he has not made a logically valid inference. The inference is invalid because the contents of the belief states - the propositions Sparky believes -fail to stand in the appropriate relation(s) to each other.

This suggests a necessary condition for validly inferring something from one's

beliefs:

> Agent A validly infers Q from his belief that P *only if*:
> 1) A believes P
> 2) A believes Q [18]
> 3) P entails Q (i.e., P could not be true without Q being true).

Some comments are in order. First, note that this is not a sufficient condition. To be a sufficient condition, there would have to be an appropriate psychological relation between the states of believing P and Q - A's belief that Q would have to be brought about in some appropriate way by his belief that P. To hit Sparky on the head with a hammer on some occasion when he believes P, and thus bring about his belief that Q, would not be a case of Sparky's inferring Q from P.

Second, notice that the above condition holds only for inferring from one's beliefs. It is of course also possible to infer validly from information that is not actually believed. For example, one might *adopt as a premise* or *suppose* the proposition that P without actually believing that P, and subsequently infer from P the conclusion that Q. In so doing, one would come to have an associated propositional attitude towards the proposition that Q (e.g., accepting Q conditional on P). Evidently, a general necessary condition on valid acts of inferring which involve such attitudes would look much like the condition given. However, for simplicity's sake, I will restrict the arguments in this paper to cases of inferring from beliefs, and not acts of inference involving contents of other attitudes.

---

[18]Need (1) and (2) occur at the same time? I leave this open. Perhaps it is possible to infer P from Q in an instant; perhaps the process takes a short amount of time.

Third, notice that a similar constraint may hold for an attitude such as assertion. That is, it is plausible that agents sometimes (verbally) give arguments. This process may be thought of as much like an act of inference. The agent asserts (or supposes) a premise, and concludes some proposition. If this is so, then a natural condition for giving a valid argument would be an adaptation of the above condition: for example, one would require that to give a valid argument one must assert (or suppose) the proposition which is the premise, assert the proposition which is the conclusion, and that the propositions asserted actually stand in the appropriate entailment relations.

Finally, notice that this condition implicitly treats belief as a dyadic relation. Nonetheless, there appears to be a analogous necessary condition for a valid act of inference even if belief were a triadic relation (such as a relation between an agent, a sentence, and the content of that sentence). If Sparky is related via sentence S to proposition P, and via sentence S' to proposition Q, then surely P must entail Q in order for any reasoning from P to Q (via S and S') to be considered valid.

C) Inferring and the Basic Metalinguistic Rephrasal Strategy

The basic MLS cannot accommodate many ordinary acts of valid inferring. Suppose that one has correctly reported the following act of inference:

> Sparky validly infers that Phosphorus is dimmer than Hesperus from his belief that Hesperus is brighter than Phosphorus .

Such a true report is certainly possible, although the actual content of this report is subject to debate. However, we may use Stalnaker's own metalinguistic theory to find the content of the above report, and in conjunction with the necessary condition for

valid inferring, demonstrate that this report, upon rephrasal, must result in an absurdity.

First, the report of the inference commits the reporter to holding that the necessary condition for such an inference has been met. That is, the reporter is committed to holding that Sparky has certain beliefs; specifically, the beliefs that would be accurately reported by the attributions:

Sparky believes that Hesperus is brighter than Phosphorus (henceforth H>P).

and

Sparky believes that Phosphorus is dimmer than Hesperus (henceforth P<H).

Moreover, since this is a report of a valid inference, the reporter is committed to holding that the beliefs involved in the reported inference do, in fact, entail each other.

However, the beliefs ascribed to Sparky by these reports are necessarily false. Since Stalnaker holds that under ordinary circumstances (and this certainly appears to be such an ordinary circumstance) rational agents do not believe necessary falsehoods, these true reports must attribute some beliefs to Sparky other than the beliefs they appear to attribute {for if the reporter is attributing a valid inference to Sparky, then he must be committed to holding that it is the beliefs which Sparky does in fact have which stand in the appropriate entailment relations to each other}. The reports should be rephrased according to the MLS, using the very idiom of the original inference report. Thus, the rephrased beliefs actually involved in the report of inference are:

> Sparky believes that the sentence 'H>P' expresses (in his native language) a true proposition[19].

and

> Sparky believes that the sentence 'P<H' expresses (in his native language) a true proposition.

Correspondingly, the original report must be rephrased to account for these beliefs involved in the reported inference. The rephrasal of the reported inference is:

> Sparky validly infers that the sentence ‘P<H’ (in his native language) expresses a truth, from his belief that the sentence ‘H>P’ (in his native language) expresses a truth.

But this is absurd. Our report of the act of valid inferring imputes an invalid inference to Sparky. Since, by hypothesis, we have correctly attributed inferences and beliefs to Sparky, it follows that Sparky has contingently false metalinguistic beliefs. But if this is so, Sparky will have made an invalid inference. We can see this by noticing that Sparky’s act of inference has failed clause 3 of our necessary condition for acts of valid inference: that the proposition Sparky believes as his “premise” must entail the proposition Sparky concludes. This is not the case after rephrasal. Consider that in one world, '<' may not be the antonym of '>.’ For, on the usual understanding of symbols such as ‘>’ and ‘<‘, it is a contingent fact that '>' expresses the brighter-than relation in the language Sparky happens to speak, and that '<' expresses the dimmer-than relation in this language[20]. In this world, 'P > H' may express a truth yet

---

[19]Alternatively, the attribution might be that Sparky believes that the sentence 'H > P' expresses the proposition that H > Mars. The subsequent arguments go through mutatis mutandis in this case.

[20]As I have hinted, I think that the “usual understanding” is wrong here, and that ‘>’ and ‘<‘ are in fact necessarily antonyms. Nonetheless, I do not think this is

'P<H' express a falsity. In this case, it is not true that the concluded proposition is true in all the worlds where the premise proposition is true. Thus, Sparky's act of inference would be, *contra* our hypothesis, invalid.

D) Some possible responses

What responses are available? One might attempt to rephrase slightly differently so as to avoid this difficulty. For example, one might rephrase the original attribution of Sparky's belief in terms of language-relative sentences. For example, the attribution might be that Sparky believes that the sentence 'H>P'-in-English expresses the proposition that Hesperus is brighter than Mars, where 'H>P'-in-English means something like 'the sentence 'H>P' in the language having the linguistic properties English actually has, where '>' expresses the brighter-than relation, and has the syntactical rules it *actually* has, etc..' If the conclusion is similarly rephrased, then the premise initially believed would entail the premise concluded. This is so since, in all the worlds where 'H>P' expresses a truth, and '>' has the semantic properties it does (e.g., '>' expresses the contrary of '<'), 'P<H' will express a truth. But the content of Sparky's beliefs would then be necessarily false: there are no worlds (in which English words express the relations they do, etc.) where 'H>P'-in-English expresses a true proposition. Thus, the content of Sparky s metalinguistic belief - according to Stalnaker - requires rephrasal into some contingently false proposition, and the

---

really a position that is available to coarse-grained theorists, for reasons I will discuss in the conclusion of this chapter.

original problem arises again.

Another possible response is to treat the act of inferring as something like an enthymeme. On this response, an act of inferring like the above is an act with a suppressed premise. The suppressed premise would be that Sparky believes (perhaps tacitly) some semantic information required for the act of inference to be valid. In our above example, we might suppose that we have suppressed the premise:

> Sparky believes that the symbol '>' expresses the brighter-than relation; its contrary is expressed by the symbol '<';

With this premise taken into account, the act of inference is once again valid. In all conditions where 'H>P' expresses the proposition that H>P, and '>' expresses the contrary of '<' (and so forth), it will be the case that 'P<H' expresses the proposition that P<H.

However, this approach suffers from a problem similar to the first response. This is that in many cases, the added premise is incompatible with the initial rephrased premise. For example, suppose we report that

> Sparky validly infers ∃y(x)(DimmerThan x,y)from his belief that ∃x(y)(BrighterThan x,y)

Again, this is an intuitively valid act of inference (although the propositions are, of course, necessarily false, since they entail that something is brighter/dimmer than itself). Since the propositions attributed to Sparky are necessarily false, however, the attributions must be rephrased as:

> Sparky believes that the sentence '∃x(y)(BrighterThan x,y)' expresses a truth

and:

> Sparky believes that the sentence '∃y(x)(DimmerThan x,y)' expresses a truth.

As in the original example, the act of inference is no longer valid upon rephrasal. But according to the suggested response, it is no longer valid, because a vital premise has been suppressed. This premise might be (e.g.):

> Sparky believes that the expressions '∃x,' and '(x)' express the existential quantifier and universal quantifier, '<' and '>' express contraries, and those symbols have the syntactic properties they actually do.

In this case, the conclusion will be true in all the worlds where the premises are true. However, the premises are now incompatible. If Sparky believes that the connective terms behave the way they actually do, then the conjunction of the two premises is necessarily false. There are no worlds in which the logical symbols express the connectives and quantifiers they actually do, and the embedded sentence of the belief attributed to Sparky (as his premise) expresses a truth (Since the expression '∃y' indicates *some y*, and the expression '(x)' indicates *every x (including the y just mentioned),* and there is no world in which there exists some y that everything - *including that very y* - is dimmer than).

A final possibility is to apply the rephrasal strategy to the report of the inference itself, as well as to the beliefs involved in the inference. Such a rephrasal could take many forms. One possibility is to rephrase the original inference as:

> Sparky validly infers* that Phosphorus is dimmer than Hesperus from his belief that Hesperus is brighter than Phosphorus .

-where inferring* is a relation with the following properties:

A validly infers* some belief Q from a belief P only if:
1) A believes P (where P is suitably rephrased)
2) A believes Q (where Q is suitably rephrased)
3) The proposition expressed by the embedded sentential complement of P entails the proposition expressed by the embedded sentential complement of Q.[21]

But this rephrasal appears completely unmotivated. Recall that the original rephrasal strategy was independently motivated by the principle of charity - that we should attempt to interpret our interlocutors in such a way as is consistent with their rationality. In this case, we have no such motivation. We have no reason to rephrase our interlocutors in such a way as all their inferences are valid. If a logic student, for example, inferred from A v B, that A, we would not be inclined to say he really inferred* A, where inferring* was valid when the proposition expressed by one embedded clause of a sentence (A, in the sentence 'A v B,' in this case) entailed the other. Moreover, although we can make valid inferences* between necessarily false propositions on this account, the fact remains that we cannot make genuine valid inferences. We are presumably interested in the relations that hold between the contents of our beliefs when we speak of inferring, and not relations that hold between parts of our belief.

E) Conclusion

Doubtless, some of our logical errors are errors about what propositions are expressed by certain strings of symbols. But if all such logical errors are

[21]Alternatively, the sentence formed by taking the embedded sentential clause of P, followed by the symbol '->' (or |-), followed by the embedded sentential complement of Q, expresses a truth.

metalinguistic in *this* sense, then many of our intuitively valid acts of reasoning are not valid at all. I conclude that the basic MLRS is inadequate and cannot account for many of our (apparently) necessarily false beliefs or assertions. We can see from the examples in this chapter that any analysis of content which rephrases necessary truths as contingent ones will risk failure to preserve the relations amongst contents which allow for acts of valid inferring.

There is, however, one possible means of escaping this problem. If one supposes that words and sentences have their referents necessarily, then the modal values of the rephrased necessarily true/false beliefs are not shifted from necessary to contingent. And if the modal values remain the same, then phenomena that depend on modality - such as inferring - are not affected. Nonetheless, this approach does not appear available to Stalnaker for two reasons. First, suppose that words have their referents necessarily. Then rephrasing a belief that H>P to become a belief that

'H > P' expresses a truth

leaves the modal value of the sentence untouched; it is still necessarily false, and hence the same belief (on Stalnaker's account) as the belief that (e.g.)

Cats are really lobsters.

Second, suppose we can ignore the modal issues. Nonetheless, we have allowed certain entities into our ontology that are structured and abstract - namely, sentences. For, as I shall argue in the second section of this work, no concrete entities could have their referents necessarily. So a sentence must be some sort of abstract entity;

and such sentences appear to be at least partially individuated by their structure.[22] But now our ontology admits structured, abstract objects which can be processed by agents to yield the truth values of the propositions they express. Are these not more or less structured propositions, as the classical theorist has envisioned them? What remaining reason is there to hold that propositions - qua objects of thought - have no structure?

We will return to a consideration of the thesis that words have their referents necessarily in the conclusion to this section of the dissertation. Let us turn now to a consideration of a more sophisticated metalinguistic rephrasal strategy for maintaining a coarse-grained theory of belief. This is Stalnaker's theory of "Propositional Concepts."

[22]Stalnaker appears to suggest such a theory of sentences in 'Mental Content and Linguistic Form,' where he claims: "Beliefs about semantics can be about specific expressions, speech acts or representational tokens, but they can also be about more abstract structures shared by certain tokens." (from Context and Content, p 236-237).

Chapter 2

Two-dimensionalism

Introduction and overview

In the last chapter, I gave a basic general strategy for analyzing assertions and belief ascriptions in terms of metalinguistic descriptions. The version of this strategy I discussed had two central parts: first, the analyzed belief ascription should ascribe a contingent belief, not a necessary one, and second, the modal value of the ascribed proposition should be shifted in this way by incorporating linguistic or metalinguistic information[23]. Of course, analyzing belief ascriptions in terms of metalinguistic descriptions is not the only way (simply the most obvious) to implement these two claims.

In this chapter, I discuss a more sophisticated metalinguistic account of content. This is "two-dimensionalism," an account developed by Stalnaker, and more recently espoused by Chalmers and Jackson. In Inquiry and 'Assertion,' Robert Stalnaker introduces the *propositional concept*, a device for using metalinguistic information to determine (non-metalinguistic) semantic content of a sentence. A propositional

[23]Of course, this strategy is not limited to providing analyses of ascriptions involving necessarily true beliefs. It is obviously intended to generalize to an analysis of necessarily equivalent beliefs (which are nonetheless intuitively distinct), such as the beliefs *that Hesperus is bright* and *that Phosphorus is bright*. In this case, although the modal status of the beliefs is not shifted, the arguments from the last chapter still apply, for the modal relations between the beliefs is shifted.

concept tells us not simply what a given sentence means, but what it *would* mean if uttered in different circumstances[24]. On Stalnaker's account, it is a function which assigns a sentence a proposition in each different world in which that sentence may be uttered. Technically speaking, it is a function which takes sentences and worlds (or perhaps "contexts of utterance") as arguments, and yields further functions (functions from worlds to truth values, i.e., propositions) as values. Finally, the propositional concept associated with a sentence can be used to select alternative propositions actually expressed by that sentence.

In this chapter, I will detail both Stalnaker's and Chalmers' account of such two-dimensional propositions and how they may be used to resolve the substitutivity problems discussed earlier[25]. Chalmers' and Stalnaker's accounts differ primarily with respect to the sorts of worlds at which propositional concepts evaluate a sentence. On Stalnaker's account (as well as Jackson's), a propositional concept evaluates sentences at relevant metaphysically possible worlds "considered as actual," and on Chalmers' account, a propositional concept evaluates sentences at relevant epistemically possible worlds. Unfortunately, each account suffers from the same general flaw. For each flavor of two-dimensionalism changes the modal value of the proposition expressed by an assertion (or attributed in a belief ascription). And

---

[24]See Chalmers, 'The Foundations of Two-Dimensional Semantics' for a comprehensive bibliography and survey of the various approaches to two-dimensionalism.

[25]Jackson [98] offers an account of two-dimensionalism very much along the lines of Chalmers. Although I shall not detail the account here, I believe the objections I raise (especially the general objection on p 58) are easily extended to this account.

I shall argue here that each account thus falls prey once more to the inferring argument of the last chapter.

A) Two-dimensional functions in general

Before we examine Stalnaker's and Chalmers' accounts of two dimensional propositions (and how they resolve the substitutivity problems), I would like to examine the general form of such two-dimensional functions. The device needed is one which assigns content to assertions or belief attributions based on contexts, and such devices have many general properties.

Let us introduce some notation here. Let us use a sentence letter followed by a subscript, such as $S_{p,}$ to indicate the sentence S that ordinarily expresses some proposition P (or 'in the actual world, the sentence S expresses the proposition that P') . We will use w_n to refer to possible worlds, and, for ease of reference, we will refer to the contexts of utterance relevant to S as $c_1...c_n$, (assuming that a context is a world). The use of 'c' and 'w' is simply to highlight the different roles a world plays - the role of a world as a context in which the sentences are being uttered, and the role of a world as a world in which a truth value is being evaluated. Let us finally use PC_S to indicate the propositional concept associated with S.

Notice that the elements of the propositional concept fit neatly into Stalnaker's ontology. First, there are sentences, which are simply concrete particulars. Second, there are the propositions expressed by those sentences, which are simply functions from worlds to truth values. Finally, there are contexts, which, for our purposes, will

be defined as possible worlds[26]. Intuitively, these are the worlds which are compatible with the speaker's presuppositions. I intend to leave this last claim rather vague right now; in a few sections I will examine some methods by which the relevant worlds might be specified.

Once a propositional concept for a sentence is defined, we can determine what I will call the "default context" and the "default intension." The default context is the actual world, and the default intension is the proposition expressed by the sentence in the actual world[27]. Generally, I shall refer to the proposition expressed by the sentence at a world w_n the horizontal proposition of S at w_n. Chalmers refers to this as the "secondary intension" of the sentence at that world. Thus, a propositional concept can be thought of as a matrix of horizontal propositions or secondary intensions. For example, the propositional concept associated with the sentence 'Hesperus is bright' would be:

'Hesperus is bright'	w_1	w_2	w_3...
c_a	T	F	F
c_b	F	T	T
c_c	T	F	T

Row 2 is the default intension of S. It is the horizontal proposition of S at the

[26]Although I treat contexts as worlds for the sake of simplification, many uses of such propositional concepts might treat contexts as ordered tuples of worlds (and perhaps "partial worlds"), and further entities such as audiences, speakers, times, and so on.

[27]On Chalmers account, it is the proposition *ordinarily thought of as being expressed* by the sentence at that world. This is so because Chalmers, unlike Stalnaker, holds that a sentence generally does not *really* express a secondary intension after all, but rather it expresses some closely related intension he calls the "primary intension"

actual world (c_a), or in other words, the proposition usually expressed by an actual utterance of 'Hesperus is bright.' This proposition is true; in some other worlds w_2 and w_3 where Venus is far from earth, and thus dim, it is false; and so on. But the propositional concept tells us what the sentence would express if the actual world had turned out differently from the way it did. For example, suppose that in the actual world, the planet Neptune had in fact risen in the spot where Venus actually appears Let us call this world c_b. In c_b, we might have used 'Hesperus' to refer to the planet Neptune, and not to the planet Venus. In w_1, the actual world, Neptune is dim, and hence 'Hesperus is bright' expresses a proposition that is false at w_1. Moreover, In w_2, unlike w_1, the planet Neptune (not Venus!) is close to the earth, and hence is bright. So the proposition expressed by an utterance of 'Hesperus is bright' at c_b is true in w_2. Thus, at c_b the proposition 'Hesperus is bright' expresses a different proposition than the one it expresses at c_a. This is the horizontal proposition of 'Hesperus is bright' at c_b, or what Chalmers calls the secondary intension of the sentence at c_b. And we can imagine that still some further world c_c had been actual, and that world is just like c_b except it was a wholly different planet, planet q, that rose where Venus actually rises in the morning. Thus if c_c had been actual, we might have used 'Hesperus' to refer to the planet Mars, instead of Neptune (as in c_b) or Venus (as in c_a). Since Mars is close to the earth in w_1, 'Hesperus is bright' expresses a proposition that is true at w_1. And let us suppose that Mars is far from the Earth in w_2, but near to the earth in w_3. Thus while Venus, Neptune and Q are bright in the actual world, at worlds w_2, w_3, etc., one planet might be bright and the others not dim,

since they are not the same planet. Thus 'Hesperus is bright' determines a different function from worlds to truth-values in each context c_n. In other words, 'Hesperus is bright' expresses a different horizontal proposition (or secondary intension) in each context. Finally, let us add a column to the propositional concept which gives us the truth-value of a sentence S uttered in some context cn is at that same world (let us call this w_n). In other words, this column tells us if S would have expressed a sentence that was true in the world at which it was uttered. In our case, S uttered in c_a would express a truth at w_a (since c_a is simply w1), S uttered at c_b would express a truth at w_b, and S uttered at c_c would express a truth at w_c.

According to Stalnaker, in ordinary speaking circumstances (circumstances where we believe the speaker to be using language the way we do, where the speaker is not contradicting himself, and so on), we may easily determine the content of a speaker's assertions and beliefs. The content of an assertion is simply the default intension (or secondary intension at c_a) of the sentence used[28]. However, where pragmatic presuppositions are violated, we may look for different contexts, or, more appropriately, try to determine *the way the world would have to be in order for the sentence to make sense or otherwise preserve pragmatic presuppositions.* In this case, we should analyze the content of an assertion or ascribed belief in terms of what we Stalnaker calls the diagonal proposition.[29]

---

[28]And, of course, similarly for the content of a reported belief.

[29]Again, there appears to be no such pragmatic subtlety in Chalmers' account. According to Chalmers, a sentence simply expresses its diagonal proposition (or what he calls its primary intension).

B) Two-dimensional functions and the "diagonal proposition"

Now let us consider how this would apply to the problem at hand - how to assign contents to assertions and belief reports. I will focus here on assertions of (seemingly) necessarily false propositions and reports of necessarily false beliefs (although, of course, the treatment here will extend to differentiating the contents of seemingly necessarily equivalent assertions and belief reports.)

Consider again our old friend Sparky, who utters S, 'Hesperus is not Phosphorus.' On the presupposition that the speaker is rational and does not believe contradictions, we might assume that the speaker believes himself to be in a world where astronomical facts (and thus the resulting linguistic practices) are slightly different. Perhaps he believes himself to be in a world where the planet Mars, and not Venus, appears as the brightest star in the morning. Thus, in this world, 'Hesperus' refers to Mars (let this be c_2). Or perhaps believes himself to be in a similar world where 'Hesperus' refers to Mercury (let this be c_3). Thus, we might associate Sparky's utterance of S with the following propositional concept:

S	w_a	c_a	w_2	c_b	w_3	c_c
c_a	F	F	F	F[30]	F	F
c_b	T	T	T	T	T	T
c_c	T	T	T	T	T	T

[30]It is not clear what truth value S has at these worlds, for we have conflicting requirements on S. Since S is uttered at c_a, 'Hesperus' and 'Phosphorus' have their actual referents (Venus), but then how do we evaluate its truth in a world where - by stipulation - 'Hesperus' refers to Mars? I will assume that even in THAT world, the use of 'Hesperus' at c_a refers nonetheless to Venus, and hence S is false at c_b and c_c.

Given the world(s) Sparky believes himself to be in, Sparky intends to express not the default proposition for S (i.e., the proposition corresponding to the horizontal row c_a) but rather the something much more like proposition expressed by rows c_b or c_c.

But this is still not the proposition Sparky actually expresses. First, Sparky is just wrong about what world he is in. Given this, his belief that Hesperus and Phosphorus are identical is simply a false belief. Yet the beliefs expressed by S in c_b and c_c are true. The horizontal propositions identified with rows c_b and c_c do not account for what is false about Sparky's belief.

Second, the propositions corresponding to rows c_b and c_c are necessarily true. But, the theory holds that necessarily true statements are to be analyzed with merely contingently true contents. Thus, we are back where we started. The problem becomes more acute when we consider rephrasing necessarily true contents of belief ascriptions. In cases of, for example, true identity statements, the non-trivial horizontal propositions are all necessarily false. This is so, because in every world where (say) ‘Hesperus’ and ‘Phosphorus’ refer to distinct objects, ‘Hesperus is Phosphorus’ will express a horizontal proposition which is false at every world.

What we want to assign to Sparky is a proposition that is actually false, but *would* have been true if Sparky had been in c_2 or c_3. This is so because what Sparky said was false in the world it was uttered in; it depended on a false belief about which world Sparky was in. If Sparky’ belief that he was in c_2 *were* true, then Sparky *would* have uttered a truth.

There is a proposition (set of worlds) represented in the matrix which is exactly the proposition we want - a proposition that is actually false but would have been true had linguistic circumstances been different. This is the diagonal proposition[31]: the function from a context of utterance c_n to the truth-value of S in that very context. In other words, the diagonal proposition is the function that takes us from c_a to the truth value of S in c_a, from c_b to the truth-value of S in c_b, and so on. This gives us the proposition that is actually false, that is true in c_a, and true in c_b. If Sparky had uttered S in c_a or c_b, he would have uttered a necessary truth; and the diagonal proposition captures this. As it stands, he has uttered a contingent falsity. Let us abbreviate the diagonal proposition of a sentence **S** as **†S**, where † is the diagonalization function.

We are ready now to initially consider the application of the propositional concept to our problems. Recall that the problem is about assertion and belief attribution. As we discussed in chapter I, Stalnaker's ontology of propositions entails that cognitive agents believe all necessary truths, and no necessary falsehoods. Moreover, for any believed proposition P, they believe all propositions necessarily equivalent to P. However, this requires a sophisticated belief attribution strategy. Speakers habitually dissent from sentences which (seemingly) express necessary truths, and speakers appear to be in error about the propositions logically equivalent to propositions they in fact hold. The diagonal proposition allows us to account for this quite nicely. First, for assertions and belief reports S *appearing* to express necessary falsehoods, we should analyze agents as instead asserting or believing the **†S,** which are merely

[31]Chalmers refers to this as the “primary intension” of S.

contingent falsehoods. Similarly for agents who believe P and appea to disbelieve some equivalent proposition(s) P'. Suppose Sparky asserts that ‘Hesperus is bright’. When Sparky then reports that he does not believe that Phosphorus is bright, we should rephrase one or more of his assertions. Although he presumably believes *that Phosphorus is bright* (since he believes *that Hesperus is bright*), perhaps does not believe †'Phosphorus is bright.’ The diagonal proposition determined by a propositional concept thus allows us to rephrase the contents of assertions and belief attributions to avoid many of the counterintuitive results of a coarse-grained theory of content.

In a moment, I will examine how the propositional concept fares against the arguments I offered in chapter I. First, however, I would like to return to the issue of how the contexts c_n are specified in a propositional concept.

C) From contexts to worlds

At this point, we should be very explicit about what contexts are involved in any given propositional concept. Since a propositional concept for S is a function from contexts to the proposition expressed by S in that context, the exact contexts involved will determine both the identity of the propositional concept and the diagonal proposition **†S**. Thus, the identity of the contexts relevant to S will determine the proposition conveyed by the sentence where that sentence appears to express some necessary truth or necessary falsehood (since the contexts will determine the diagonal proposition **†S)**. Thus it will be of interest of us to have some general theory of just

what contexts are relevant to an utterance of S.

Let us put the issue somewhat differently. A context of utterance can be itself modeled as a (set of) possible worlds. So a propositional concept PC_S for a sentence S is a function from possible worlds $c_1....c_n$, to some (further) functions $p_1.....p_n$. Thus the identity of the $c_1....c_n$ will partly determine the identity of PC_S, and thus the identity of **†S**. If $c_1....c_3$ are relevant to an utterance of S_1 but not S_2, and $c_4....c_6$ are relevant to an utterance of S_2 but not S_1, $\mathbf{\dagger S}_1$ and $\mathbf{\dagger S}_2$ may very well be distinct. So we would like to know how to specify which worlds $c_1...c_n$ are relevant to an utterance of S.

Stalnaker himself appears to offer at least two methods for specifying these worlds, and one can easily see a variation on these methods which might also be used. First, the relevant worlds may be those worlds compatible with the agent's background beliefs. As Stalnaker says,

> When someone makes an assertion, his words determine not only a proposition, but a propositional concept, relative to the possible worlds compatible with the speaker's presumed background information[32]

The natural way to read this is to take the propositional concept of a sentence as giving the propositions expressed by S in those worlds compatible with the non-metalinguistic) propositions the agent believes. For example, a propositional concept will show what proposition is expressed in those worlds in which the agent's beliefs about the linguistic facts are true (as our above exegesis has described). Let us call this the compatible-belief method of specification.

---

[32] Stalnaker [87A] p182

However, there is a slightly different means for specifying relevant contexts. As Stalnaker writes in 'Assertion,'

> ...the question, "what is O'Leary saying"? not by asking about the semantical rules for the sentence O'Leary is using, but instead by asking what the world would have to be like if what O'Leary seems intuitively to be saying were true...[33]

On this account, the contexts specified are not specified directly in terms of propositions believed. Rather the relevant contexts are *worlds which would make true the utterer's sentences*. These, after all, are the worlds that the agent believes he is in. This method, in fact, is reminiscent of a suggestion of Kripke's in Naming and Necessity:

> What does the intuition that [this] table might have been made of ice, that it might even have turned out not to have been made of molecules, amount to? I think that it means simply that there might have been a table looking and feeling just like this one and placed in this very position...which was made of ice. I could have been in qualitatively the same epistemic situation that in fact obtains, I could have the same sensory evidence that I in fact have, about a table that was made of ice.[34]

In short, we are interested in contexts epistemically identical to the actual context, but perhaps involving distinct objects (planets, substances, etc). Specifically, in the case of utterances which appear to express necessarily false propositions, we are interested in epistemically identical contexts which make the agent's utterances true. For example, in our case of the utterance 'Hesperus is not Phosphorus' we are interested in those worlds where an utterance of 'Hesperus is not Phosphorus' would have expressed some truth. Perhaps one of these worlds is one where 'Hesperus'

---

[33] Stalnaker [87A] p183

[34] Kripke [82] P142

refers to some entity other than (the actual) Hesperus (and likewise for 'Phosphorus'). Call these planets H_2 and P_2. In this world - something like the world our utterer takes himself to be in - 'Hesperus is not Phosphorus' expresses a necessary truth - namely, *that H_2 is not P_2*. This method is also the one suggested by Chalmers [2000]. According to Chalmers, the diagonal proposition (or what he calls the "primary content") expressed by a sentence is the function from an epistemically possible world to the truth value the sentence would have expressed in that world[35].

Let us call this the compatible-world method of specification. Neither means will allow us to escape the original inferring argument, as I will now show.

D) Inferring and the diagonal proposition

However one specifies the relevant contexts for an utterance of S, we will still have the original problem of the inferring puzzle. For the relevant contexts in which utterances S_1 and S_2 express distinct horizontal propositions will not guarantee the modal connections between S_1 and S_2 (and hence the modal connections between **†S**$_1$ and **†S**$_2$), even in cases where the contents of S_1 and S_2 intuitively entail each other.

---

[35]It is worth noting that Chalmers places some restrictions on these epistemically possible worlds. First, the function are not total functions; e.g., we need not ask what 'water is yummy' expresses in a world where the stuff that people drink is composed out of cheese molecules and has wildly different phenomenal qualities. Moreover, by 'epistemically possible,' Chalmers is referring to what he calls *deep epistemic possibility*, or compatible what an ideal reasoner would believe. Thus, the proposition *that water = xyz* is deeply epistemically possible, but the propositions *that not all who masticate, chew* or *that something is brighter than everything* are not deeply epistemically possible. This will be important for our treatment of predicate terms.

That is, even in cases where the content of S_1 intuitively entails the content of S_2, there will be relevant contexts where an utterance of S_1 expresses a horizontal proposition that is true in that context, but the same utterance of S_2 in that context expresses a proposition that is false in that world. In these cases, there will be worlds where is **†S**$_1$ true but is **†S**$_2$ false, contra the hypothesis that the content of S1 entails the content of S2. In what follows, I shall detail a case in which the propositional concept account appears to accord with our intuitions, and then I will look at a few cases where the propositional concept account falls prey to the inferring problem.

Consider again the case in which the hapless Sparky appears to believe that Hesperus is brighter than Phosphorus, and (appears to) infer from this that Phosphorus is dimmer than Hesperus. Again, this is surely a valid inference. And again, since the proposition that H>P is necessarily false, the coarse-grained theory calls for that proposition (as well as the proposition that P<H) to be rephrased, in this case as † 'H>P' and † 'P<H.'

Now let us consider those worlds compatible with Sparky's beliefs. Perhaps there is a world c_b in which there are 2 distinct planets, H_2 and P_2, and these are respectively called 'Hesperus' and 'Phosphorus.' So we might define a propositional concept for 'Hesperus is brighter than Phosphorus,' such as:

'H>P'	w_a	w_2	c_b	w_3

c_a	F	F	F[36]	F
c_b	F	F	T	F

Is c_b compatible with Sparky's beliefs? According to Stalnaker and Chalmers (and any other advocate of coarse-grained, unstructured propositions), the answer must be, no. For Sparky believes that 'Hesperus' refers to Hesperus, and hence he believes that 'Hesperus' refers to Phosphorus[37]. In c_b, however, he does not believe that 'Hesperus' refers to Phosphorus. Still, it is possible that we might restrict the $c_1...c_n$ to those worlds compatible with the agent's *non-metalinguistc* beliefs. Alternatively, we could consider the $c_1...c_n$ that represent the ways the world would have to be in order for the claim 'Hesperus is Phosphorus' to be true, as the compatible-world method of context specification suggests. Nonetheless, this cannot be a world where any old planets H_2 and P_2 rise in the position where Venus actually rises. Intuitively, they must be worlds in which H_2 and P_2 are respectively known as 'Hesperus' and 'Phosphorus.'

It is now easy to see that the account fails to accommodate our necessary condition on valid inferring. There are many worlds compatible with Sparky's non-metalinguistic beliefs, relevant to his assertion that 'Hesperus is brighter than Phosphorus.' And there are many worlds which would make true an assertion of 'Hesperus is brighter than Phosphorus'. None of these are worlds in which English is used as it actually is, since in these worlds the English names 'Hesperus' and

---

[36]Again, I am unsure what the truth value should be here. For Hesperus and Phosphorus do not exist in c_b. I assume that 'H>P' as uttered in the actual world would have been false at c_b. Similar considerations apply to row c_b / column c_a. The issue is merely academic, since the only relevant boxes will be c_a/w_a and c_b/c_b.

[37]For further discussion of this point, see Soames [87]

'Phosphorus' are used differently than they actually are[38]. Consider again c_b, where 'Hesperus' and 'Phosphorus' refer to the planets H_2 and P_2, and H_2 really is brighter than P_2. This too is a world in which English is used non-standardly; it is a world compatible with Sparky's non-metalinguistic beliefs, and a world in which the belief Sparky apparently expresses with *'Hesperus is brighter than Phosphorus'* would be true. And finally consider c_c, in which 'Hesperus' and 'Phosphorus' both refer to Venus (like ca) but 'Brighter' and 'Dimmer' are not antonyms (perhaps they refer to the *at-least-as-bright-as* relation and the *dimmer-than* relation)

With this in mind, let us consider the propositional concepts associated with 'H>P' and 'P<H':

'H>P'	w_a	w_2	c_b	c_c
c_a	F	F	F	F
c_b	F	F	T	F
c_c	T	T	?	T

We will suppose that in w_2 and w_3, Venus exists, but H_2 and P_2 do not. Thus, in the context of utterance c_a, 'H>P' a necessary falsehood. If c_b had been the context of utterance, 'H>P' would have been (presumably) false (or at least indeterminate) in w_a, w_2, and w_3, since the planets it refers to do not exist in those worlds. But it would have been true at c_b, since it would have (intuitively) expressed the proposition that

[38]One might object that proper names are not really part of English. Thus c2 is nonetheless a world in which the same English as our own is spoken. We will substitute examples such as 'masticating is ruder than chewing' if this is the case.

H_2 is brighter than P_2, which is true in c_b. If the context of utterance had been c_c, it would have expressed the proposition *that Venus is at least as bright as Venus*, which is a necessary truth, which would have been true in c_b[39]. We similarly begin to define the propositional concept for 'P<H' as follows:

'P<H'	w_a	w_2	c_b	w_3	c_c
c_a	F	F	F	F	F
c_b	F	F	T	F	F
c_c	F	F	F	F	F

The only difference here is that in the context c_c, 'P<H' intuitively expresses the proposition that Venus is dimmer than Venus, which is necessarily false. The diagonal propositions will be formed by taking the values at c_a/w_a, c_b/c_b, and c_c/c_c. Thus, the proposition †'H>P' is true in c_c, yet the proposition †'P<H' is false in that world. This is possible since in c_d, '>' and '<' are not antonyms, and hence even the coarse grained propositions expressed by 'H>P' and 'P<H' will have different truth conditions. In fact, since '>' and '<' do not express antonyms, there is guaranteed to be a world in which the sentences 'H>P' and 'P<H' (as uttered in c_d) differ in truth value. The conclusion, of course, is that †'H>P' and †'P<H' do not entail each other, since there are worlds (namely c_c) in which the former proposition is true but the latter proposition is false. Thus, they cannot *both* be the contents of Sparky's

[39]Of course, since the referents of 'H' and 'P' are distinct in c_b, 'H>P' uttered in c_b will be true in some further worlds containing both H_2 and P_2, and false at others, depending on the relative brightnesses of H_2 and P_2 in those worlds.

respective beliefs, since Sparky intuitively made a valid inference.

If the diagonal proposition is to serve as the content of Sparky's beliefs, then worlds like c_d cannot be considered when we identify the propositional concepts relevant to Sparky's' beliefs. So, perhaps the answer is to limit the propositional concept for some sentence S to only the closest possible contexts compatible with the agents' (non-metalinguistic) beliefs, or to only the closest possible worlds (specifically, the closest possible worlds with distinct metalinguistic facts) which would make the agent's belief true. It is difficult to see exactly how such a "closeness" relation would rule out worlds like c_d while retaining worlds like c_c. Perhaps the idea is that c_c and c_b differ from c_a in only one fact, whereas in c_3 not only are the names non-standard, but the relation terms '>' and '<' are nonstandard. Or perhaps the right response is to hold, as Chalmers does, that the contexts are selected not on the basis of any metalinguistic facts, but that they are selected on the basis of being qualitatively identical.

Even if these sorts of restrictions on relevant contexts were plausible, they will not prove helpful in a large class of cases, such as Mates-style synonymous predicate cases, and what I call the 'structural error' cases.

E) Inferring and synonymous predicates

The puzzle is even more pointed when it involves synonymous predicates rather than co-referring names. Consider the following example of an intuitively true inference report:

> Sparky believed that masticating was ruder than chewing, and thus he validly inferred from this that chewing was more polite than masticating.

As before, Sparky's reported beliefs appear to be necessarily false. Thus, according to Stalnaker, Chalmers et al., the content of Sparky's beliefs are really †'masticating is ruder than chewing' and †'chewing is more polite than masticating.' The trouble is that it is extremely difficult to find a possible context in which the sentence 'masticating is ruder than chewing' could have been true. Unlike the Hesperus/Phosphorus cases, there are *no* purely qualitatively identical worlds in which *our* sentence 'masticating is ruder than chewing' is true. For such a world would be a world where masticating and chewing were different properties, and this is impossible. Thus, we must consider worlds which differ in their metalinguistic facts, for example, a world in which 'masticating' refers to the act of slobbering, and 'chewing' refers to the act of chewing. Let us call this c_b. But as before, there will be many worlds with different metalinguistic facts that are relevant to the utterance of 'masticating is ruder than chewing,' including worlds where 'ruder than' and 'politer than' are not antonyms (perhaps they refer to the relations of being *ruder than* and *at least as polite as*, respectively). Let c_c be a world like c_a (where 'masticate' and 'chew' have their ordinary English referents). In c_c, however, 'ruder than' and 'politer than' do not have their ordinary English referents. Then we can begin to define the relevant propositional concepts as follows:

'm ruder than c'	w_a	c_b	c_c
c_a	F	F	F
c_b	T	T	T

c_c	T	T	T

The sentence, as uttered at c_a, expresses a necessary falsehood, since it expresses the proposition that masticating is ruder than itself. As uttered at c_b it expresses a necessary truth since it expresses the claim that slobbering is ruder than masticating[40]. Thus, as it is uttered at c_b, the sentence expresses a necessary truth. And similarly at c_c, where 'masticating is ruder than chewing' expresses the proposition *that chewing is at least as rude as chewing*. This is also a necessary truth (it is, after all, exactly as rude as chewing!)

Now consider the propositional concept associated with 'c politer than m':

'c politer then m'	w_a	c_b	c_c
w_a	F	F	F
c_b	T	T	T
c_c	F	F	F

The reasoning in the above is the same as before, with one difference. At cc, the sentence expresses the claim *that chewing is more polite than chewing*. This is of course, a necessary falsehood. Thus, there is a world (namely, c_c) in which †'masticating is ruder than chewing' is true, and †'chewing is more polite than masticating' is false. Thus, †'masticating is ruder than chewing' does not entail †'chewing is more polite than masticating,' and Sparky's reported inference, upon

[40]For simplicity's sake, I assume that rudeness is "fixed" and not relative to the societal mores of the world. This simplifying assumption does not affect the argument.

analysis, is not valid (contra our hypothesis). Moreover, c_c and c_b are, on any reasonable analysis, equally "close" to c_a, so we cannot rule out c_c on grounds of closeness[41]. They each differ from c_a with respect to only two metalinguistic facts.

F) Structural errors, diagonal propositions and inferring

Recall from the previous chapter our example of structural errors. That is, beliefs whose necessary falsity can only be explained as a result of an error in processing the *structure* of the proposition (or relevant sentence expressing that proposition). These cases differ from cases we have been discussing, in which agents appear to have necessarily false beliefs as a result of substituting individual terms. Consider, for example, the belief attributions

Sparky believes that -(A -> A) (Sparky believes that **P**)

and

Sparky believes that A ^ -A (Sparky believes that **C**)[42]

Note that **P** and **C** are logically false. Thus, the two attributions require rephrasal. To determine **†P** and **†C** (the appropriate rephrasals of the relevant that-clauses), we construct a propositional concept for each. Thus we must determine the contexts

[41]Moreover, if we rule out contexts like cc, then the propositions corresponding to †'masticating is ruder than chewing' and †'chewing is politer than masticating' will be identical, and we will not have solved the puzzle of differentiating the contents of Sparky's beliefs! I will return to this point later in the chapter.

[42]One may find P and C too transparently contradictory to be plausible beliefs. In that case, feel free to substitute something more complex, e.g., -(P ^ Q) -> -(P -> -Q), and (-P v -Q) -> -(P -> -Q).

Sparky might believe himself to be in - the ways that "the world would have to be like if what [Sparky] seems intuitively to be saying were true..."[43] (In this case, we want to know the ways the world would have to be in order for the sentence that Sparky would use *to express what he believes* to be true).

What sort of world might make S_c and S_p true? Not a world in which there are any different propositions referred to by the clause 'A.' In fact, nothing about S_c or S_p refers to any particular facts about objects, properties or relations. Thus, the only ways S_c and S_p could express a true proposition would be if structural or syntactic facts about **L** (or whatever language was used to express P and C) were different.

Are there such worlds? Perhaps so. But note that in those worlds there must be a wholesale rearrangement of the semantics of **L** (or whatever language the sentences expressing **P** and **C** were inscribed in). If S_p expresses a truth in those worlds, the change will be global. For example, consider the world c_1, in which the symbol '->' expresses disjunction and the symbol 'v' expresses disjunction.[44]

But now we can see how the inferring argument of the previous chapter may be implemented. Stalnaker's rephrasal strategy tells us that Sparky really believes †S_p and †S_c. To determine these diagonal propositions, we construct a propositional concept for Sp and Sc; to construct such a concept we consider worlds where Sp and Sc express truths. Alas, in c_2, S_p expresses a contingent sentence, whereas S_c

---

[43]Stalnaker [87A] p183

[44]Notice that the more complex S_p and S_c become, the more changes in the syntax of L must be made to make them true. For example, even if 'v' stands for conjunction, then -(P ^ Q) -> -(P -> -Q) is still necessarily false. So '-' and/or '->' must have non-standard meanings as well.

expresses a necessary truth. Thus, there is some world c_3, like c_2 in all metalinguistic respects, where the atomic sentence A is true (suppose A is proposition that Al Gore won the presidency in 2000). In this world, the sentence '-A v A' expresses the proposition that Al Gore did not win or Al Gore won, which is true in that world; but the sentence '-(A -> A)' expresses the proposition that -(Al Gore won v Al Gore won). Since by hypothesis Al Gore won, S_c is false in this world. The propositional concepts below illustrate this:

'-A ^ A'	w_a	w_2	w_3
c_a	**F**	F	F
c_2	T	**T**	T
c_3	T	T	**T**

'-(A -> A)'	w_a	w_2	w_3
c_a	**F**	F	F
c_2	?	**T**	F
c_3	T	T	**F**

It is easy to see here that the propositions †S_p and †S_c do not entail each other. For in w3, †S_p is true while †S_c is false. Thus, if †S_p and †S_c are the real contents of Sparky's beliefs, it is impossible for him to make a valid inference from the one to the other. But the inference from P to C is intuitively valid. Thus †S_p and †S_c cannot be the contents of Sparky's beliefs (or the contents of the reports of Sparky's beliefs).

One might object that in the initial construction of the propositional concepts associated with P and C, we are to consider contexts which make (both) Sparky's beliefs true. This seems implied by the original claim, that we should be "asking what the world would have to be like if what [Sparky] seems intuitively to be saying were

true..."[45]. But this has its own problems. For suppose the contexts $c_1....c_n$ are restricted to only those worlds in which what Sparky 'seems to be saying' is true. In that case, all of Sparky's apparently necessarily false beliefs are *still* identical - they are all some diagonal proposition which is actually false, but true everywhere else. We are back at square one, where we cannot explain how Sparky appears to believe one necessarily false belief but appears not to believe another. Even worse, all of Sparky's apparently necessarily false beliefs will be implied by his merely contingently false ones! For consider some contingently false proposition A believed by Sparky and some other apparently necessarily false belief B. Sparky's actual belief relevant to B is then $\dagger S_B$, a belief actually false but true in every other relevant context. But then there is no world where A is true and $\dagger S_B$ is false. Thus the arbitrary false belief implies the arbitrary necessarily false belief, which is surely a reductio.

The point may be neatly generalized. If a proposition is simply a set of worlds (or the characteristic function of that set), then attributions of necessarily false beliefs seem to require rephrasal, because we can correctly attribute one (apparently) necessarily false proposition to an agent without attributing all propositions. But the contents of these rephrased beliefs had better have different truth values in at least one world, or the rephrasals are all identical with each other as well, and rephrasal changed nothing. Yet if the contents differ in truth value in even one world, propositions that intuitively entail each other no longer entail each other, and hence

[45]Stalnaker [87A] p183

inferences that are intuitively valid are in fact invalid. Thus the diagonal propositions cannot serve as rephrasals of Sparky's beliefs (or any other agents).

Conclusion

The propositional concept and the associated diagonal proposition are ingenious tools for resolving many of the substitution puzzles brought about by externalist theories of reference. However, where substitution problems occur for reasons other than externalism - such as apparent structural differences or perhaps co-expressing predicate terms - rephrasing with the diagonal proposition is of no avail.

The central problem is the contention that propositions have no structure but are instead merely equivalent to their conditions for truth. Thus, to believe a proposition is to believe all propositions with the same truth conditions. To account for the counterintuitive belief attributions, the contents of assertions and attributions are rephrased. The rephrasal strategy involves attributing beliefs about the relations of linguistic entities to propositions. But these contingent propositions cannot underwrite the sorts of logical phenomenon - like inference - structured propositions account for. I conclude that whatever the nature of propositions, many propositions are structured and have their structure essentially.

In the chapters to come, we will investigate richer theories of belief and belief attribution, which attempt to overcome these problems by allowing for essentially structured objects of belief and/or enriching the nature of the belief relation itself. Let us move on to these.

Chapter III

Structured Propositions

A) Structured propositions

In the last chapter, we concluded that propositions must be *structured* entities. That is, they must be differentiated at least in part by their logical structures, and not merely by their truth conditions. For example, it appears possible to believe the proposition

That the cat is on the mat

while disbelieving the proposition

That the mat is under the cat,

even though these two propositions have the same truth conditions[46]. So the propositions believed appear to be individuated by more than just their truth conditions. If we hold that propositions are structured, we are free to point out that these two propositions involve different relations - the former is a structure containing the *on* relation, and the latter is a structure containing the (intuitively

[46]Of course, the *apparent* intuition that an agent could believe the former while disbelieving the latter might be explained by referring to some third 'mode of presentation' of the proposition, or some pragmatic difference between particular belief *reports* involving the two propositions. At this point, however, we will retain the simplifying assumption that belief is a two - place relation between a cognitive agent and a proposition. Moreover, we shall retain the assumption that belief attributions simply express this relation. In the following chapters, I shall address the more complex accounts.

distinct) *under* relation. Moreover, different entities fill the subject role in each structure. Thus these two beliefs are distinct structures, and it is possible to stand in the belief relation to one without automatically believing the other, as is intuitive.

Such a theory of structured propositions goes a long way toward solving the sorts of substitutivity problems we have been discussing. Nonetheless, serious problems still remain. In this chapter, I will examine some representative theories of structured propositions, and discuss several problems which still face these theories. I begin with the examination of a purely hypothetical account which extends the metalinguistic treatment from the last chapter to a theory of *structured* propositions. While purely hypothetical, such an account provides a useful illustration of many of the sorts of problems that face theories of structured propositions, and a useful springboard to more complex theories. For example, the account has difficulties accounting for cross - linguistic belief attributions; it individuates beliefs so finely that they cannot be shared by many agents that intuitively do share them; and it falls prey to a similar version of the inferring problem discussed in the last chapter. Then I shall examine Larsen and Ludlow's theory of "Interpreted Logical Forms," which treats propositions as natural-language syntactic trees with each node of the tree annotated with its semantic value (in the case of the terminal nodes, these values are individuals, properties, relations, and so on). This theory has exactly the opposite sort of problem that the coarse-grained theories have - that is, it is *too* fine-grained. Among other things, the theory cannot account for the possibility of making true

attributions of beliefs to speakers of different languages, and it individuates beliefs so finely that it cannot account for many cases of believers who intuitively share beliefs.

i) The Structures

Of course, there are many ways to model propositional structure. Take the proposition

That Venus is brighter than Mars.

One might, for example, model this as a sequence (ordered set) of entities. These entities might themselves be any number of things. For example, they might be individuals and *sui generis* properties and relations, resulting in structures such as

<the brighter-than relation, Venus, Mars>[47].

Alternately, the proposition in question might be a sequence of entities such as possible worlds constructs. For example, we might treat the *brighter-than* relation itself as the set of all possible ordered pairs in which the first individual is brighter than the second, i.e., {<Venus, Mars>,<Sun, Earth>,<Mars, Pluto>, ...}. [48] In this case, the proposition in question would be analyzed as

<{<Venus, Mars>,<Sun, Earth>,<Mars, Pluto>, ...}, Venus, Mars>

Such so-called "sequence theories" have their own specific problems, which we will not go into here (but see, for example, Bealer [93],[98]).

---

[47]As usual, we use the braces {} to denote unordered sets, and the angled brackets <> to denote ordered sets.

[48]For representatives of such theories, see (for example) Soames [87] and Cresswell [85].

Alternately, one might think that propositions derived their structure from the grammar of the natural language expressions used to express them. Viewed this way, propositions might naturally be modeled as syntactic trees terminating in values such as individuals and properties or relations. As with the sequence theories, one might also choose to analyze these constituent properties and relations as anything from *sui generis* entities to possible-worlds constructs.[49] Finally, one might treat propositions themselves as *sui generis*, irreducible entities, which have no "constituents" but instead are associated essentially with some structured entity like a tree.

Although there are many ways in which one might model the structure of a proposition, what I want to make clear is that each account will have certain problems in common. In what follows, I intend to discuss very general problems that hold for any theory of structured propositions.

ii) The Problems

Structured propositions neatly solve some of the substitutivity problems discussed earlier (especially, ch 1). For example, we have already seen how structured propositions differentiate propositions such as

that the cat is on the mat, and
that the mat is under the cat

It is also easy to see how propositional structure may be used to differentiate logically equivalent "structural" propositions, such as the DeMorgan equivalences

[49]For examples of such theories, see (for example) King, ([94], [95]), and Larsen and Ludlow (1993).

~(A v B) and
(~A & ~B).

Since these propositions are true in exactly the same circumstances, they cannot be differentiated by a theory such as Stalnaker's which reduces propositions to their truth conditions. But a theory of structured propositions can account for the difference very simply. For example, a sequence theory might treat these two propositions as:

<NEG <DISJ<A, B>>> and
<CONJ <NEG A, NEG B>>[50]

Since these sequences (structures) are distinct, it is easy to see that standing in the belief relation to one structure need not commit an agent to standing in the belief relation to the other.

Nonetheless, problems remain. For there are propositions that have not only identical structures, but also seemingly identical constituents (properties and individuals), which nonetheless appear distinct. Consider our old chestnuts:

Hesperus is bright
Phosphorus is bright

Intuitively, it is possible for an agent to believe the former, while denying the latter (and if not, we require at least some *account* of the appearance of such a possibility)[51]. However, our supposition is that Hesperus just *is* Phosphorus. Since the

[50]I have deliberately left the structure of the constituent clauses A and B unaccounted for here. This is so since however one chooses to structure these clauses, the two propositions are already distinct in virtue of the "macro" structures described here.

[51]Of course, Fregeans hold that the relevant propositions involved in beliefs do not involve the individuals Hesperus and Phosphorus. Rather, they involve distinct

structures of each proposition are identical, and the constituents are identical, the propositions are identical. To stand in the belief relation to one is just to stand in the belief relation to the other. As an example, consider that on the sequence theorist's account, both of the above propositions are analyzed as:

<The property of brightness, Venus>

There seem to be two general means for resolving this problem. First, we can attempt to give a theory of belief reports on which the that - clauses in reports such as

Sparky believes that Hesperus is bright, and
Sparky believes that Phosphorus is bright

somehow express distinct propositions. This will require a complex analysis of these propositions. Second, we can give a more complex account of belief and belief reports, according to which an agent may (for example) believe a single proposition in many ways, and it is these 'ways' of believing which account for the apparent differences in belief and belief reports. It is the first approach that we will discuss in this chapter.

B) Structured Metalinguistic Propositions

In the last chapter, we discussed a basic metalinguistic rephrasal strategy. According to this strategy, reports of the form

---

senses of Hesperus and Phosphorus; and Russellians hold that the propositions actually involve disguised descriptions of these individuals. Following Kripke [72] however, such moves have been widely rejected. What I think Kripke demonstrated was not that *senses* should be rejected, but rather, that *descriptive* senses should be rejected, for reasons I shall discuss in this and the next two chapters.

A believes that Fa

are analyzed as

A stands in the belief relation to the proposition that the sentence ‘Fa’ expresses a truth.

However, the metalinguistic proposition was itself identified with its truth conditions - that is, with the set of possible worlds in which the sentence ‘Fa’ expresses a truth. And we saw that this created seemingly insurmountable problems. A natural extension is simply to treat the above metalinguistic propositions as structured in some way, rather than equating them with their truth conditions. That is, let us assume some theory of structured propositions, and treat the problematic constituents as metalinguistic descriptions rather than as individuals or properties. For example, consider our earlier example,

Sparky believes that Hesperus is Hesperus.

According to our hypothetical account, this would have the analysis:

<Believes, <Sparky, <Identity, <the x:<names,<‘Hesperus,’ x>>, <the x:<expresses,<‘Hesperus’, x>>>>>[52]

which is, as our intuitions track, distinct from the structured proposition

<Believes, <Sparky, <Identity, <the x:<names,<‘Hesperus’, x>>, <the x:<expresses,<‘Phosphorus’, x>>>>>

The two structured propositions differ in virtue of containing distinct names for

---

[52]Again, we suppose the term

the x:<expresses,<‘Hesperus’, x>>

is to be understood as *de dicto* (i.e., as having the semantic value of the description itself). For if it is understood as *de re* (i.e., as having the value of the thing described, namely Venus), then the substitutivity problems arise even here.

Venus in the final clause (the names ‘Hesperus’ and ‘Phosphorus.’) In less technical terms, these reports state that Sparky believes that the thing referred to by the word ‘Hesperus’ is the thing referred to by the word ‘Hesperus.’ At the same time, he does not believe that the thing referred to by the word ‘Hesperus’ is the thing referred to by the word ‘Phosphorus.’ This gives the results desired - that the propositions believed and denied by Sparky are distinct. The treatment easily extends to our stripped-down version of Mates’ puzzle, explaining how it is possible that Sparky believes that John masticates and yet he does not believe that John chews. The two reports are analyzed as:

<Believes, <Sparky, < <the property x:<expresses,<‘Masticates’, x>>>, John>>>

and

<NEG < Believes, <Sparky, < <the property x:<expresses,<‘Chews’, x>>>, John>>>>

Again, in less technical terms, this simply says that Sparky believes that John performs the action expressed by the word ‘masticates,’ but he doesn’t believe that John performs the action expressed by the word ‘chews.’

Although this treatment neatly addresses several of the substitutivity puzzles, it is saddled with several flaws. First o f all, it cannot account for cross-linguistic belief attribution or sameness of belief between speakers of different languages. Second, it is really a flavor of descriptivism, and suffers all the problems of that theory (specifically, it is prey to Kripke’s modal argument, as we shall see). Finally, analyzing belief reports as involving structured metalinguistic propositions changes

the modal value of the reported beliefs, and is thus still susceptible to a modified version of our earlier inferring argument. Let us now address these points in detail.

i) Cross-linguistic belief sharing and attribution

Treating beliefs as containing linguistic elements has an obvious prima facie cost. That is, those linguistic elements are not normally shared by speakers of different languages. Thus, intuitively correct belief attributions will turn out to be false on analysis. For example, it may be intuitively true that

Pythagoras discovered (believed) that Hesperus is Phosphorus.

However, the proposed metalinguistic analysis, viz.,

<Believes, <Pythagoras, <Identity, <the x:<names,<'Hesperus', x>>, <the x:<names,<'Hesperus', x>>>>>

is surely false. Pythagoras, an ancient Greek, did not know or use the words 'Hesperus' and 'Phosphorus,'[53] so it is extremely doubtful that he had beliefs involving them. Similarly, it seems intuitive to say that modern (English-speaking) astronomers come to discover the truth of the same thing Pythagoras famously discovered. But on the above analysis, this is impossible. English speakers and Greek speakers believe very different metalinguistic propositions.

One might attempt to address this by using something analogous to Carnap's

[53]At least, not as these words are ordinarily understood. It is debatable whether or not 'Hesperus' and 'Phosphorus' are ordinarily considered to be Greek words as well as English; to make the point more clearly, substitute some attribution involving words with no obvious ties to Greek, such as "Claret is Burgundy'.

original stratagem involving translation:

> Pythagoras believed that the celestial body named by some Greek word which translates as the English ‘Hesperus’ is the same as the celestial body named by some Greek word which translates as the English ‘Phosphorus.’

Still, this too is beset with difficulties. In the first place, it is hideously ungainly. Is *that* monstrosity really correct? Did Pythagoras really believe that? Unlikely. It is even more unlikely when one considers the beliefs of, say, small children, who do not seem to possess the cognitive resources to consider metalinguistic contents[54].

Secondly, as Church famously pointed out,[55] the analysis logically commits us to the existence of the English language. Surely Pythagoras might have made his discovery even had the English language never come about. But had the English language not existed, the above analysis would be either meaningless or false, since Pythagoras would not have had any beliefs involving words that translated as the (nonexistent) English word ‘Hesperus.’ A similar point made be made more forcefully by translating the analysis into another language:

> Pythagoras glaubte daß der Himmelskörper, der durch irgendein griechisches Wort ausgedrückt wurde, das als das Englische ‘Hesperus' übersetzt; ist dasselbe wie der Himmelskörper, der durch irgendein griechisches Wort ausgedrückt wird, das als das Englische ‘Phosphorus' übersetzt.

Since translation preserves meaning, the English and the German attributions should have the same meaning. But few German speakers, if any, believe any propositions involving the *English* words that correctly translate the Greek words used by

---

[54]For example, young children do not appear able to make the sorts of use/mention distinctions required.

[55]Church [1950]

Pythagoras.

Finally, the analysis in terms of translations is not fully determinate. For, presumably, ancient Greek has *several* words which translate as 'Hesperus' or 'Phosphorus' [56]. Now suppose W1 and W2 are ancient Greek words which each name Venus. Then suppose that Pythagoras believed the triviality that Hesperus was Hesperus. On the current account, this means that:

> Pythagoras believed that the celestial body expressed by W1 is the same as the celestial body expressed by W1 (note two instances of W1)

But if this is so, and W1 and W2 each name Venus, then W1 translates as 'Hesperus' and *also* as 'Phosphorus.' Thus, it is also true that

> Pythagoras believed that the celestial body named by some Greek word which translates as the English 'Hesperus' is the same as the celestial body named by some Greek word which translates as the English 'Phosphorus.'

And if this is the case, then the analyzed report says that Pythagoras believes the informative identity that Hesperus is Phosphorus. And surely this is not a consequence of his believing the triviality!

Of course, the problems presented here for the metalinguistic strategy are merely prima facie problems. It is, of course, possible that a richer theory of belief and belief attribution could resolve these issues. Nonetheless, these *are* problems that must be addressed, and hence we have a motivation to enrich the metalinguistic strategies. Before we discuss such possible refinements, however, I would like to look at some

---

[56]Here I assume that proper names are correctly intertranslatable when they share their referents. It is, of course, possible that a more robust theory of translation can account for this by tightening the requirements for translation of proper names. One way to do so would be to give richer theory of identity for words, which is precisely what I intend to do in the second main part of this work.

further problems faced by metalinguistic strategies.

ii) Structured metalinguistic propositions and inferring

The account at hand avoids the original inferring argument. Although it changes the modal values of beliefs, it changes them in such a way that the truth values of the premise beliefs vary exactly with the truth values of the conclusion beliefs. This is so because we do not need to consider what a metalinguistic proposition would express in a world with different natural language grammars; the only relevant metalinguistic facts which change from world to world are those facts involving the referents of the individual words. To illustrate this, consider our original inferring argument:

> Sparky believed that Hesperus was brighter than Phosphorus, and from that he inferred that Phosphorus was dimmer than Hesperus.

The two propositions involved in this inference will be, upon analysis, something like:

> <Brighter-than, <the x:<expresses,<'Hesperus', x>>, <the x:<expresses, <'Phosphorus', x>>>>

and

> <Dimmer-than, <the x:<expresses,<'Phosphorus', x>>, <the x:<expresses, <'Hesperus', x>>>>.

Unlike Stalnaker's rephrasal strategy, the *sentences* used to express the beliefs are not invoked in the rephrasal, only some of the constituent words involved in the sentences. It is easy to see that there is no possible world in which the above premise belief is true, and yet the conclusion belief is false. So, as is intuitive, Sparky's inference is valid.

Still, a change as drastic as altering the modal value of a proposition believed should be expected to have repercussions elsewhere. Consider:

> Sparky believed that Hesperus was Phosphorus, and from that, he inferred that *necessarily*, Hesperus was Phosphorus.

This is an intuitively valid inference. Again, our rephrasals involve the propositions:

> <Identity, <the x:<expresses,<‘Hesperus’, x>>, <the x:<expresses,<‘Phosphorus’, x>>>

and

> <Necessity, <Identity, <the x:<expresses,<‘Phosphorus’, x>>, <the x:<expresses,<‘Hesperus’, x>>>>.

This, however, is *not* a valid inference. For consider a world in which the thing named by ‘Hesperus’ is the same as the thing named by ‘Phosphorus’ (again, we assume *de dicto* readings of the relevant metalinguistic descriptions). It is nonetheless false that necessarily, the thing named by ‘Hesperus’ is the thing named by ‘Phosphorus.’ For their may be a further world where the word ‘Phosphorus’ is used to name the sun, in which case, even though ‘Hesperus’ and ‘Phosphorus’ co-refer, they do not do so necessarily.

Perhaps, it may be objected, this argument rests on false premises about proper names like Hesperus and Phosphorus. For it asks us to consider worlds where the word ‘Hesperus’ names the sun, but according to Kripke’s (widely accepted) account of proper names, proper names refer to the same thing in every world[57]. Nonetheless,

---

[57]Recall the Kripkean doctrine that proper names are rigid designators, that is, they refer to the same thing in all worlds. It seems nonetheless true, at least on the received view of names, that it is a contingent fact of English that ‘Hesperus’ is a proper name. But suppose that the word ‘Hesperus’ is essentially a proper name. In

this objection may be surmounted by replacing the individuals/proper names in the example with properties/predicate terms:

> Sparky believed that chewing was masticating, and from that, he inferred that necessarily, chewing was masticating.

Again, it seems surely possible that ‘chewing’ and ‘masticating’ express the same property, without *necessarily* expressing the same property. And so Sparky’s intuitively valid inference is not valid upon rephrasal.

A similar example was given in chapter 1 of this work. Suppose that Sparky believes that necessarily, Hesperus is Venus. Then, upon analysis, he believes that

> <Necessity, <Identity, <the x:<names,<‘Phosphorus’, x>>, <the x:<names,<‘Hesperus’, x>>>>.

But, at least on the ordinary understanding of words, Sparky does not believe the latter. For Sparky, like most people, believes that the word ‘Phosphorus’ could have referred to something totally different from (Venus). It could have referred to Lobsters, as far as Sparky believes. Ergo, such a metalinguistic analysis of Sparky’s attributes metalinguistic beliefs to Sparky that he simply does not appear to have (and would probably vigorously deny).

Once more, this is a *prima facie* problem, which a more complex theory might resolve. Still, the problem does need to be addressed by even theories of structured

---

that case, we have already committed to the thesis of thick words - for the word ‘Hesperus’ has its referent necessarily. And this raises two points: first, what sort of a thing *is* a word if it has its referent necessarily? It surely is not an orthotype or a phonotype. And second, if some words have their referents necessarily, what reason is there to think that the claim does not hold of *all* words, and is not merely arbitrarily restricted to proper names?

propositions in beliefs and belief ascriptions. Before we discuss such refinements, let us consider one final general problem faced by metalinguistic strategies.

iii) Descriptivism

The proposed metalinguistc analyses are really just forms of descriptivism about proper names (and even of predicate terms!). According to descriptivism, beliefs about individuals and properties really are beliefs involving certain *descriptions* (or descriptive entities like Fregean "senses") of those individuals and properties, and not involving the individuals or properties themselves. Instead of describing the individuals using intrinsic characteristics of the individuals and properties (such as the property of *being the first star to appear in the morning*) , the metalinguistic strategies describe individuals using relational properties - namely, the property of being referred to by the linguistic symbols we use to represent the individuals and properties. But just as these descriptions fail to have the relevant modal characteristics to capture modal phenomena like inferring, they also fail to have the right modal characteristics generally. Recall that one of Kripke's original arguments against descriptivism was, in fact, a modal argument. Kripke pointed out that necessarily, Hesperus is Hesperus. But for *nearly any* description of Hesperus (call such a description D), it is not true that necessarily, Hesperus is D. It is not true, for example, that necessarily, Hesperus is the first star to appear in the evening. And similarly, it is not true (again, at least on the ordinary conception of words) that

necessarily, Hesperus is the thing referred to by the word 'Hesperus.' [58] And since the metalinguistic rephrasal is just a flavor of descriptivism, we should expect the rest of the common objections to descriptivism to apply here as well.

One of these further objections to descriptivism involves the sharing of belief. Intuitively, two agents may share Hesperus-beliefs even though they describe Hesperus differently. For example, agent A may describe Hesperus as *the first star that appears in the morning*, and Agent B may describe it as *the star with seasonal cycle such-and-such*. Yet intuitively, they may both be described as having the de dicto belief that Hesperus is bright. And so the descriptions do not appear to be what are involved in their shared beliefs. On the other hand, metalinguistic descriptions are by their very nature sharable and public. But as our cross - linguistic attribution examples show, even these metalinguistic descriptions are not public enough to account for shared beliefs between speakers of different languages (at least, not as the words involved in these metalinguistic descriptions are commonly understood).

A further general problem with descriptivism is that agents seem to share beliefs involving some individual, without sharing descriptions of that individual. Similarly, agents often fail to possess the descriptions required to individuate their beliefs along the correct lines. For example, most Americans know nothing of either Cicero, apart from his being a famous Roman orator (if they know anything at all of him). So it is

---

[58]The dual of the example makes this more stark. For it is true that contingently, Hesperus is referred to by the word 'Hesperus.' But it not true that contingently, the thing referred to by 'Hesperus' is the thing referred to by 'Hesperus.'

not clear at all how descriptions could solve any substitutivity puzzles involving beliefs about Cicero and Tully. The issue extends to even metalinguistic descriptions, for we can imagine two distinct individuals who are correctly described as *being known by the name 'John Smith,'* for example.

The arguments given in this section (section B) lead me to conclude that structured propositions cannot simply be augmented with metalinguistic content in this way without running afoul of problems involving modality and associated phenomena like inferring. Nonetheless, there is a lesson to be learned. Even structured propositions on their own do not *appear* to be fine-grained enough to serve as the objects of belief. Linguistic objects, on the other hand, seem to provide such fine differentiation. Perhaps it is possible to augment propositions with such linguistic content without resorting to descriptivism.

C) Interpreted Logical Forms

i) ILFs

Perhaps it is possible, as I have suggested, to augment propositions with very fine-grained objects such as words, without explicitly holding the proposition to involve the metalinguistic relations. One way of doing this is adopted by Larsen and Ludlow in their paper 'Interpreted Logical Forms'[1995]. The central idea of the paper is that the analysis of belief attributions involves a two-place relation between an agent and a hybrid linguistic entity called an *Interpreted Logical Form.*

According to Larsen and Ludlow, an Interpreted Logical Form (henceforth ILF) is

formed by taking a syntactic tree structure of some natural language clause, and annotating the terminal nodes (the individual words) with their semantic values (individuals, properties, and so on).[59] Thus, a belief attribution reports that an agent stands in the belief relation to the ILF formed from the natural language that-clause used in the attribution. So, for example,

Sparky believes that Hesperus is bright

reports that Sparky stands in the belief relation to the ILF

<<NP<'Hesperus', Venus>, <VP<'is bright', brightness>>>[60].

It is easy to see that this is distinct from the belief reported in a similar attribution such as:

Sparky believes that Hesperus is bright.

This reports that Sparky stands in the belief relation to the ILF

<<NP<'Phosphorus', Venus>, <VP<'is bright', brightness>>>.

Although these ILFs agree in what Larsen and Ludlow call their "objectual" values and maybe even their structures, they differ on their "formal" characteristics (viz., the words at the terminal nodes). So Sparky's belief that Hesperus is bright does not commit him to a belief that Phosphorus is bright. It is also easy to see how the theory differentiates coextensive beliefs with necessarily equivalent truth values, such as

---

[59]Of course, the theory relies upon an already established theory of syntax, and an already established theory of semantic values, but the choice of such theories need not concern us here.

[60]Here, for simplicity, I have omitted the "higher" nodes of the tree, such as the sentence node and its value (truth), and I have omitted the annotated values of the non-terminal nodes. Nothing turns on these nodes. Also, I have chosen to represent tress via ordered sets, again a merely notational difference.

Hesperus is brighter than Phosphorus, and
Phosphorus is dimmer than Hesperus

for the ILFs associated with each are:

<<NP<'Hesperus', Venus>, <VP <VP<'is brighter than', the brighter-than relation>>, <NP<'Phosphorus', Venus>>>>

and

<<NP<'Phosphorus', Venus>, <VP <VP<'is dimmer than', the dimmer-than relation>>, <NP<'Hesperus', Venus>>>>

which are clearly distinct.

Finally, ILFs do not *replace* proper names with descriptions, metalanguistic or otherwise. Thus, the recursive truth definitions for ILFs determines the truth values for the ILF out of the values of the objectual components alone,[61] which are still present in the ILFs. So the truth value at a world of the ILF corresponding to some that-clause will be exactly as the truth value of the clause before being annotated with lexical items. In short, the modal value remains the same after analysis, and the inferring problem does not arise.

Although ILFs seem to allow us a means of finely individuating beliefs, without incurring the problems of explicitly metalinguistic analyses, it has its own problems. Firstly, it requires a much richer theory of lexical objects - *words* - than the conventional view of words. Secondly, it individuates beliefs *too* finely; in addition to general problems of cross - linguistic attributions, it also has difficulties accounting for beliefs involving demonstratives like 'this' and 'that,' and *shared de re* beliefs. In

[61]For details, see Larsen and Ludlow [93] p 312

what follows, I will detail these claims.

ii) Words and ILFs

The theory of ILFs relies heavily on the words used in belief reports to differentiate the content of the reported beliefs. But as it stands, ILFs and any ILF-like theory will require a very robust theory of words in order to account for obvious puzzle cases. As such, even if the theory were (roughly) correct, it would not obviate the need for the theory of words offered in the next part of this work.

One can see the problem clearly when homonymous names are constituents of ILFs. Consider the following example: Rock Hudson attends a New-age camp, incognito, and one night there is a naming ceremony. Camp members are given "back to nature" names, like 'Leaf,' 'Tree,' 'Stream,' etc. When the ritual comes to Mr. Hudson, he is given the name 'Rock.'[62]. Our old friend Sparky is also attending this camp, and we report his beliefs as follows:

> Sparky believes that Rock is a movie star and
> Sparky does not believe that Rock is a movie star

Intuitively, both may be true. But upon analysis, the attributions seem to report that Sparky both believes and does not believe the ILF

> <<NP<'Rock', Roy Scherer>>>, <VP<'Is a movie star', the property of being a

[62]This example is of course inspired by Kripke's famous "Paderewski" example from Kripke [88]. The difference here is that in the Paderewski example, the believer knows Paderewski by what he mistakenly believes to be two distinct names with different referents ('Paderewski' and 'Paderewski'). In fact, both names are intuitively the same. In the "Rock' example, Roy Scherer is known by what really are intuitively two different names.

movie star>>>.

Since the names ‘Rock’ and ‘Rock’ *appear* to be the same, it is difficult to see how they be used to differentiate the ILFs Sparky believes/does not believe[63].

Larsen and Ludlow address this point directly in their discussion of Kripke’s now-famous ‘Paderewski’ example. In the example, Ralph knows a single individual (the aforementioned Paderewski) under two different contexts. He knows Paderewski the symphony conductor, Paderewski his shy upstairs neighbor. He also knows each person by the name ‘Paderewski,’ although he does not appear to know that they are the same person. So he appears to believe both that Paderewski is shy, and that Paderewski is not shy. So according to the theory at hand, it appears that Ralph both believes and does not believe the ILF

<<NP<‘Paderewski’, Paderewski>>, <VP<‘is shy’, the property of shyness>>>

Nonetheless, it does not appear that Ralph has contradictory beliefs. In order to resolve this, Larsen and Ludlow hold that the names ‘Paderewski’ occurring in each

[63]Again, there appears to be no reason why these ‘co-expressive homonym’ puzzles should be confined to proper names. One might easily imagine, for example, a pair of mathematicians working independently, each investigating the formal properties of a certain type of function. Mathematician 1 describes this kind of function in terms of the sort of machine that can compute the function, and calls the property ‘Computability.’ And Mathematician 2 describes the kind of function in terms of the sort of algorithm that can be used to build the function. He briefly considers calling the property ‘recursiveness,’ but has been thinking about computing machinery and decides to call the property ‘Computability.’ More pedestrian examples abound. The point is, just as we coin words to refer to individuals, and sometimes coin homonymous words that refer (unknown to us) to the same individual, so too we coin words to express properties, and nothing prevents us from creating co-expressive homonyms for these properties either.

ILF are actually *syntactically distinct objects*[64]. In short, they are the names 'Paderewski$_1$' and 'Paderewski$_2$.' Although they are spelled and pronounced identically, they are nonetheless completely different words - homonyms, if you will. And of course, this treatment should generalize to the 'Rock' example. So the ILF

> <<NP<'Rock'$_1$, Roy Scherer>>>, <VP<'Is a movie star', the property of being a movie star>>>

is not the same as the ILF

> <<NP<'Rock'$_2$, Roy Scherer>>>, <VP<'Is a movie star', the property of being a movie star>>>.

And one may believe the former while consistently denying the latter.

I agree that 'Rock' and 'Rock,' in the above example, are intuitively two different words. But wherein resides their difference? Not in their physical characteristics, such as shape, size, sound, etc. According to Larsen and Ludlow, the grammar '...presumably also has the resources to discriminate syntactically between homophones, say, through 'diacritics' like those distinguishing 'bank'$_1$ and 'bank'$_2$.'[65] But this simply puts the question off. For the diacritical marks simply proclaim that the words are distinct, without telling us what distinguishes them.[66]

While I do not think this is a crippling objection, it does make it clear that even the theory of ILFs cannot stand *in place of* a theory of words, but will instead, at best,

---

[64]Larsen and Ludlow[88], p 319

[65]Larsen and Ludlow [88], p 319

[66]Notice also that the response does not appear to handle the 'Paderewski' cases very well. For I think it is intuitively plausible that the two names 'Rock' and 'Rock' are actually distinct yet homonymous names. But in the Paderewski case, there is intuitively but *one* name 'Paderewski,' used twice; and Sparky fails to realize it is the very same name.

rely on a theory of words. And while I do intend to give such a theory, I think that there are further objections to ILF which are, in fact, crippling. Let us examine these objections.

iii) Too Fine-Grained

One of the selling points of ILF is that it takes seriously the idea that we seem to be able to differentiate beliefs in at least as many ways as we have (linguistic) resources to report those beliefs. And this seems intuitively correct, as all our examples have shown - at least up to a point. But ILFs are *too* finely individuated. According to the theory of ILFs, people who do not share the same words (for example, people who speak a different language) can *never* share beliefs. Moreover, we cannot accurately report their beliefs in our language. And this is simply implausible.

To see why this is so, consider Larsen and Ludlow's individuation criteria for ILFs:

> The assumption that ILFs are composed of linguistic forms and extralinguistic objects yields straightforwards individuation criteria for ILFs: two ILFs will be distinct whenever they contain distinct forms or distinct objects. [67]

This has the happy result that the of ILFs corresponding to belief reports such as our earlier Hesperus/Phosphorus examples are distinct. So far, so good. On the other end of the scale, however, consider Sparky, who believes that Frenchmen are more dangerous than Englishmen. According to the theory at hand, Sparky stands in the belief relation to the ILF

[67]Larsen & Ludlow [88], 315

<<NP<‘Frenchmen’, Frenchmen>>,<VP <VP<‘are more dangerous than’, the more-dangerous-than relation>><NP<‘Englishmen’, Englishmen>>>>

And no Frenchman, for example, could be related by belief to *this* ILF, unless of course he spoke English. So no monolingual Frenchman, such as our intrepid Jean-Jacques (who inspired Sparky’s belief), could share Sparky’s belief, or even accurately report it. As close as Jean-Jacques can come is

<<NP<‘Les Francais’, Frenchmen>>,<VP <VP<‘est plus dangereux que’, the more-dangerous-than relation>><NP<‘Les Anglais’, Englishmen>>>>

Yet these are wholly distinct ILFs - for their “formal” components are radically distinct. The problem becomes even more apparent when one considers that the logical form R(x,y) may be implemented with wholly distinct syntactic structures in different languages - one language may use an ‘infix’ structure such as S -> NP VP NP, where another countenances only a prefix structure such as S -> VP NP NP.

But this is incredible. Even in languages where the relevant proper names, predicates and even general sentence forms are shared, it may be impossible to share beliefs! In our above example, it is plausible that the words ‘Frenchmen’ and ‘Francais’ are just variant spellings of the same word, as are ‘dangerous’/‘dangereux,’ etc. Nonetheless, the ILFs would still be distinguished in virtue of their other lexical components - i.e., the determiners ‘Les’ vs. ‘the,’ and the comparative constructions ‘plus....que’ vs. ‘more...than’ [68].

[68]There really appears to be no reason for us, however, to restrict ourselves to thinking of simply natural-language syntactic forms. We might consider hybrids of logical forms, or perhaps (universal) mentalese forms. I shall not consider such a theory of ILF at present, although when we consider the ‘sequence view’ of words,

Needless to say, Larsen and Ludlow anticipate such an obvious consequence of their theory, and they consider two alternative approaches to addressing the issue. First, there is the possibility that belief attributions are a bit more complex than it appears, and that they report believers to be related to certain ILFs *or ILFs relevantly like them*. Second, there is the possibility that such cross - linguistic attributions really are, strictly and literally, false; but that they may be more or less *useful*.

Larsen and Ludlow (in my view, rightly) reject the first response after brief consideration. The complaint is basically that such an broad construal of the belief relation is hopelessly vague. It inevitably runs the danger of undoing all the work the original theory accomplished. In what way is the ILF

> <<NP<'Frenchmen', Frenchmen>>,<VP <VP<'are more dangerous than', the more-dangerous-than relation>><NP<'Englishmen', Englishmen>>>>

"relevantly like" the ILF

> <<NP<'Les Francais', Frenchmen>>,<VP <VP<'est plus dangereux que', the more-dangerous-than relation>><NP<'Les Anglais', Englishmen>>>>?

It is surely not in 'sameness of objectual content.' For then it would also be 'relevantly like':

> <<NP<'Frenchmen', Frenchmen>>,<VP <VP<'are more dangerous than', the more-dangerous-than relation>><NP<'Britons', Englishmen>>>>?

But the original goal of the theory was to account for the possibility of believing that (for example) Frenchmen are more dangerous than Englishmen, without also believing that Frenchmen are more dangerous than Britons. No real account of

---

such considerations will naturally arise.

'relevant similarity' is offered, and none seems plausible. Moreover, this response treats the verb 'believes' radically differently from other English transitive verbs, such as 'kicks.' For it would be an unacceptable analysis of 'kicks' to say that 'x kicks y' is true when x kicks y or some individual relevantly similar to y. (For an extended discussion of this issue, see Larsen & Ludlow, 1993 pp 335-339, and Segal 1989).

These points seem right-headed to me, and I see no reason to pursue this approach further here[69]. Larsen and Ludlow themselves instead opt for a 'pragmatic' solution to the problem, where such 'similarity or same-saying in such cases are fundamentally a matter of usage and not of content, and that the correct account of such phenomena falls outside the domain of semantics proper and into pragmatics.' In short, such cross-linguistic attributions are strictly speaking false, but they may be useful for various purposes, and we may say that the beliefs of speakers of other languages are 'similar' to our beliefs based on that usefulness.

The main problem is that no real theory of 'usefulness' is offered. Speakers, we are told, must use general principles of common-sense psychology in order to build a fine grained-model of both their audience's psychology, as well as the psychology which the audience believes the attributee to have. This model will be used for the speaker to select an ILF which it will be 'useful,' in the context, to attribute to some agent. But the details of this process are missing. Larsen and Ludlow simply point to

[69]We shall, however, look at ways to capture such 'relevant similarity' or 'type identity,' in the next chapter.

some research in psycholinguistics and express the belief that 'this work must be extended to the study of the way states of mind come to be described, and why subtle differences in an expression will have great consequences for the truth of attitude ascriptions.'(p 342). That is a pretty big blank check. In fact, even if such a check can be eventually cashed, it is unclear what need there will be for ILFs. Moreover, it seems highly likely that such a theory will require an (appropriately) fine-grained model of propositions[70] which the speaker supposes the hearer to have (and which the speaker supposes the hearer to attribute to others). In short, there is no guarantee that the pragmatic theory needed by Larsen and Ludlow will not itself require a theory of propositions sharable by speakers of different languages.

At any rate, I would like to make a methodological point here. That is, we should prefer a theory which, *ceteris paribus*, takes our intuitions about attributions at face value. If we are given a choice between treating cross-linguistic attributions (such as the above) as possibly true, and treating them as merely useful, we should prefer the theory on which we may make strictly true attributions of speakers of other languages. Perhaps this is not ultimately possible, but nonetheless, we should canvas the remaining options.

The cross-linguistic attribution problems are not the only example, alas, of the "too-fine-grained" issue. The issue crops up again in the treatment of demonstratives and in *de re* attributions, as I shall now discuss.

---

[70]I.e., a model other than the one provided by the theory of ILFs

iv) Demonstratives

There is a more specialized version of the above objections that ILFs are too fine-grained to serve as the objects of belief. This objection focuses on the treatment of demonstratives like 'this' and 'that.' As it turns out, ILF precludes an adequate theory of such demonstratives. This holds for whatever one's theory of demonstratives and demonstrative content is. This is so since the way that ILFs incorporate the indexical content causes them to be far too fine-grained, even without considering cross-linguistic attributions.

Consider three individuals, O'Leary, Sparky and Willard. All three are native English speakers. They sit at dinner with a bottle of Pinot Noir (which is closest to O'Leary) and a bottle of Sancerre (closest to Sparky). O'Leary points to the bottle of Pinot Noir and utters

'I believe that this is better than any white.'

Sparky turns to Willard and with a roll of his eyes utters

'O'Leary believes that that is better than any white.'[pointing at the Pinot]

According to Larsen and Ludlow, the two attributions are:

> O'Leary believes [the ILF that] <<NP <'this', I1>> <VP <V<'is better than', the better-than relation>>, <NP<'any white', all white wines>>>>

and

> O'Leary believes [the ILF that] <<NP <'that', I2>> <VP <V<'is better than', the better-than relation>>, <NP<'any white', all white wines>>>>[71]

[71]The full structure of the ILFs here has been abbreviated to remove structure that is not really relevant to the examples; in particular, I have lumped together several terminal nodes as a single node, as in the clause 'all whites' or 'any white.'

Let I1 and I2 be the content of the relevant demonstratives. (Larsen and Ludlow endorse Burge's account of demonstratives, according to which the content of a demonstrative is a pair <x, e> consisting of an object x and an act of pointing to said object e . However, it is important to note that my objection is completely independent of one's theory of the nature of I1 and I2.)

The problem is this. Intuitively, Sparky's assertion to Willard is true - Sparky is correctly reporting O'Leary's tastes and beliefs, barbaric though they may be. But even if I1 and I2 are identical, the 2 ILFs are not identical - for one ILF contains the word 'this' and the other contains the word 'that.' If the two ILFs are not identical, then Sparky is misreporting O'Leary's beliefs. And no matter what one's theory of demonstrative content, this point remains. And one cannot 'fix' the point without abandoning, or at least circumscribing, the central claim that the semantic content contributed by a word to an ILF is that word and its objectual content[72]. It is also extremely difficult to imagine that Sparky's report is 'false but useful.'

v) *De Re* belief

It has been pointed out repeatedly that belief reports may legitimately be used to express coarse-grained beliefs. For example, suppose Officer O'Leary points at a man

---

This "lost structure" does not affect the example.

[72]One might very well choose to circumscribe this claim, and hold that only *some* words contribute themselves to their semantic content. But then, where is the theory of ILFs? What we have is something like a sequence view where *some* elements of the sequence are <word, object> pairs. As previously mentioned, I shall discuss such a view in the second part of this work.

running down the street away from a murder and exclaims, 'There goes Son of Sam, the murderer!.' Unbeknownst to him, it is not the murderer, but rather, the rap impresario, Snoop Dogg. Later, Officer Sparky reports O'Leary's blunder, uttering 'O'Leary thought Snoop Dogg was Son of Sam!'[73]

Now intuitively, even though O'Leary would surely reject the claim that Snoop Dogg was Son of Sam, Officer Sparky has reported a truth. Moreover, such attributions seem to be coarse grained, in that substitution of identical names preserves truth. The above attribution seems to have the same content as, say,

'O'Leary thought that Calvin Broadus was Son of Sam.'[74]

Ordinarily, O'Leary's report would be characterized as a *de re* attribution - something like,

Snoop Dogg was such that O'Leary believed him to be Son of Sam

where at least one singular term referring to Snoop Dogg occurs outside the scope of the belief. And such a *de re* attribution intuitively has the same value as

Calvin Broadus was such that O'Leary believed him to be Son of Sam.

Examples of such *de re* attributions are common. Soames [00 (p 157-158)], for example, offers some helpful examples:

there is something that Mary believes is heavy .
Mary believes that Peking is one of the world's largest cities, but she doesn't

---

[73]I here treat 'Son of Sam' as a non-descriptive proper name. If the reader believes the name to carry descriptive content, reconstruct the example, substituting your favorite pair of non-descriptive proper names, such as the now-tired Hesperus and Phosphorus.

[74]'Calvin Broadus' is the birth name of the aforementioned rap impresario, Snoop Dogg.

know its name.
Mary believes that Beijing is one of the world's largest cities, but she doesn't know its name.
Mary believes that you are interesting to talk to (said to me).
Mary believes it will rain in Recife tomoorow (said yesterday).

Larsen and Ludlow propose a mechanism for capturing such coarse-grained *de re* reports. According to them,

> two distinct attitude reports a and b will be logically equivalent when. . .i) the values assigned to the subparts of the complement clauses are identical. . . and ii) a and b are evaluated under structures in which their formally distinct (but coreferring) subparts are given scope out of the complement clauses, beyond the highest attitude verb. (p 322)

In other words, the theory models *de re* reports by creating a tree for the report with the *de re* component given wide scope and a "trace" (the linguist's equivalent of a bound variable) occurring inside the belief-clause. For example, the logical form of the *de re* reading of a report like 'Sparky believes Snoop Dogg is Son of Sam' involve the ILF:

> <<NP <Snoop Dogg>$_1$ <VP <NP <O'Leary>> VP <Believes, the [ILF]<NP<t1, Snoop Dogg>>, <VP<<'is', predication>, NP<'The Son of Sam', David Berkowitz>>>>

where t1 is the 'trace' of node 1 (in this case, the un-annotated individual, Snoop Dogg). For all effective purposes, 'traces' work as bound variables, as we might expect in the ordinary, non-ILF *de re* analysis

> ∃x (x=Snoop Dogg & believes (O'Leary, the proposition that x = Son of Sam))

This 'trace' analysis preserves the 'coarseness' expected of a *de re* report, for it is the same report as the report one would use to capture that

> O'Leary believed Snoop Dogg is Son of Sam.

or even

O'Leary believed Calvin Broadus is Son of Sam.

for in each analysis, the un-annotated individual Snoop Dogg appear outside the scope of the belief context, and the annotation t_1 appears inside its scope (and all other subparts of the believed ILF are identical). So it appears that this "scoping-out" mechanism allows the theory of ILFs the means to capture *de re* belief reports.

But all is not well. Larsen and Ludlow's "scoping-out" mechanism is, by their own admission, of very limited scope. Although it appears to model reports involving de re occurrences of names, it cannot accommodate similar reports involving *de re* occurrences of predicate terms. Consider Larsen and Ludlow's own example,

O'Leary believes that Ludlow photocopied War and Peace
O'Leary believes that Ludlow xeroxed War and Peace.

Just as there are intuitively *de re* readings of the Snoop Dogg / Calvin Broadus examples, there are intuitively *de re* readings of the xerox/photocopy examples. But as I mentioned, the theory of ILFs cannot capture these readings. For, as Larsen and Ludlow claim,

> The scopal mechanism for producing truth conditionally equivalent attitude reports will of course be constrained by whatever syntactic limitations exist on what can be 'moved' at [logical form] and the distances such moved phrases may traverse. (p323)

And although there are well-understood rules by which noun phrases may be 'moved' across the logical form of a sentence, by Larsen and Ludlow's own admission, there are no rules in the grammar of English for moving predicate

terms across logical forms! And so, they hold that there are no analyses of the above xerox/photocopy examples on which the reports have equivalent content, and thus, no possibility of a similar *de re* analysis of the reports.

This strikes me as deeply flawed. First, whatever motivation one can give for a de re analysis of noun phrases in belief reports, seems to apply equally to predicate phrases. Oddly, Larsen and Ludlow simply deny that there can be de re readings of the xerox/photocopy examples, without giving any independent motivation for the claim. And this flies flatly in the face of intuitions. There is no discussion of why, for example, we should take our intuitions at face value in the Hesperus/Phosphorus cases, but not the xerox/photocopy cases[75].

Second, eliminating the possibility of sameness of content for the xerox/photocopy cases removes an important tool for resolving the too-fine-grained issue. Recall that cases of cross-linguistic attribution may, on &L's account, be useful in certain contexts, even if the ILFs attributed to a believer do not contain the linguistic items that the believer is actually related to. One imagines that the ILFs it is useful to attribute to a believer must at least share the same objectual content as the ILF the believer really does believe. And while we might use the scoping-out mechanism to establish "sameness of objectual content" in the Hesperus/Phosphorus cases, we cannot use it to establish

[75]It is possible that Larsen and Ludlow believe xeroxing and photocopying to be distinct properties, and perhaps they are right about this. But for this to generalize, they would have to take the radical view that *all* distinct predicate-term pairs express distinct properties, and this just seems completely unacceptable. At any rate, this is simply speculation.

"sameness of objectual content" in the xerox/photocopy cases.[76]

The objection really turns on the nature of ILFs (and the logical form of belief reports) being given by natural-language structures. If an ILF is in fact structured by a natural-language tree, then the admissible ILFs will be restricted by the grammatical rules of that natural language. It is of course possible that the grammatical rules of that language are not well enough understood, and that further research will uncover a scoping-out mechanism that applies to verb phrases (the predicate/property pairs of ILFs) as well as to noun phrases. But Larsen and Ludlow themselves do not seem to hold out much hope for this.

Finally, one might abandon the claim that ILFs are natural-language structures, and instead treat them as logical structures or perhaps deep psychological structures (e.g., some sort of universal "mentalese"). As I already alluded to, such a theory appears rather different than ILF. It looks much like a straightforward sequence theory, where some or all of the elements of the sequence are word/property or word/individual pairs. I will return to this in the second main part of this theory when I consider the sequence theory of words, in the second main part of this work.

Conclusion

[76]Indeed, Fiengo & May [96] seem to assume that this is the point of the scoping-out mechanism, that is, to restrict the kinds of ILFs that may be "usefully" attributed to monolingual speakers of other languages. But I cannot find positive evidence of this claim in L&L, so the point remains speculation.

Most of the obvious attempts to include linguistic items in propositional content fail for a variety of reasons. First, the attempts often invoke some form of descriptivism, changing the modal value of the propositions under consideration. Second, the attempts are generally *too* fine grained, in that the linguistic elements are not shared or shareable in the way we want propositions to be shareable. As I mentioned at the outset of this chapter, these points prompt at least two natural responses. One might imagine belief to be a more complex relation, one which holds between an agent, a non-linguistic proposition shared by speakers of different languages, and some very finely individuated (linguistic) "mode of presentation" of that proposition. This is the approach that we will examine in the next chapter. Or, one might hold that the linguistic elements involved in beliefs - words - were considerably more abstract and shareable than words are normally conceived of. The latter approach is the one I believe the most promising, and what I shall focus on in the second main part of this work, "Words."

Chapter 4

Triadic and Hidden-Indexical Theories

The most recent chapter concluded that even a theory of fine-grained, essentially structured propositions will suffer substitutivity problems. That is, the fine-grained, structured propositions

Hesperus is brighter than Mars,

and

Phosphorus is brighter than Mars

appear to be one and the same proposition, and yet, it seems possible to believe the former without believing the latter. Moreover, attempts to "augment" these propositions with extremely fine-grained constituents such as words, are nearly as troublesome. Among other issues, the resulting propositions are so fine-grained that they are not plausibly shared by agents that intuitively share beliefs.

One extremely natural response this problem is to treat belief itself as a more complex relation. The above substitutivity puzzles seem to suggest that though believers may have beliefs with the same contents, they may also believe these propositions in different ways. This suggests that we might treat belief as a three-place relation, one that holds between a believer, a (fine- grained) proposition, and something like the "way in which the proposition is believed." This last element is commonly referred to as a "mode of presentation," a sort of mental representation

through which the agent is related to the proposition.

Since belief reports rarely (if ever) overtly refer to these 'modes of presentation,' such a three-place theory of belief will require a more complex theory of belief *reporting* to go with it. For we still need to show how it is possible for reports such as:

Sparky believes that Hesperus is brighter than Mars,

and

Sparky believes that Phosphorus is brighter than Mars

to differ in truth-value. Proponents of a three-place belief relation customarily treat these reports as containing a covert clause which picks out the "mode of presentation." Following Schiffer (92), we will refer to these theories as "hidden-indexical" theories[77].

There are at least two main components of such a theory. Firstly, we require an account of entities that serve as the so-called "modes of presentation" (as well as an account of the way in which reports refer to this "mode of presentation"). Secondly, we require an account of the relation that holds between the agent, the proposition, and the mode of presentation. As such, there are a host of triadic belief and hidden-

[77]Of course, a theory might have a hidden indexical element without also having a three-place structure like the one mentioned here. For example, we might have a theory much like the aforementioned ILF theory, where belief related an agent to a very fine-grained ILF-like structure, while at the same time, the lexical elements of that ILF-like structure were specified via a hidden indexical (in fact, such a picture is an accurate account of Larsen and Ludlow's theory from the last chapter; Larsen and Ludlow's hidden indexicals are simply not as context sensitive as the ones considered in this chapter). For this reason, I will refer to the theories of belief as "triadic" theories, and the theories of reporting as "hidden-indexical" theories.

indexical theories. Nonetheless, I think that all such theories will share several general flaws. First, the theory of belief reports is just a bad analysis of the form and content of those reports. Agents making belief reports do not plausibly refer to three-place belief relations or to modes of presentation. Cognitive agents are rarely (if ever) aware of such modes of presentation, they often do not appear to have the cognitive resources required to refer to them, and they do not plausibly have the intentions required to refer to them (overtly or covertly). Second, and more importantly, there just are no suitable entities to serve as 'modes of representation'; all such candidate entities simply collapse into problems previously mentioned. Such modes of presentations either result in beliefs and attributions still being too coarsely individuated, or they make beliefs too *finely* individuated to plausibly be shared (or even correctly reported) by rational agents.

Of course, we cannot here address every possible variation on triadic and hidden-indexical theories. Thus, in what follows, I shall first sketch the general form such theories take, and then illustrate by giving an outline of one of the leading contemporary theories - Crimmins' account in Talk about Belief. Finally, I shall direct my critique of triadic and hidden-indexical theories at Crimmins' account. Ultimately, I believe that the form of the criticisms should suggest how they may be generalized.

A) Triadic and hidden-indexical theories - the general form

Triadic theories of belief have two main components. First, they describe the

structure of the belief relation itself, and second, they describe the nature of the entities involved in this relation. As we stated earlier, such triadic theories also require a revised theory of attribution - the so-called hidden-indexical theory - to explain how agents may correctly report these complex triadic beliefs.

First, let us consider the structure of the belief relation itself. According to triadic theorists, the term 'belief' is ambiguous. Sometimes it is used to refer to a content of belief (i.e. a proposition) and sometimes it is used to refer to the *state* a believer is in when they bear this belief relation to a proposition. It is this latter relation or state that the hidden-indexicalist is interested in. This is a three-place relation between an agent, observer, and some other entity (such as a mode of presentation). We can understand this relation as

A believes the proposition *P* under the mode *M*,

or

A believes the idea *I* that has the content of the proposition *P*,

or whatever the specifics of the theory tell us. As one can see from the above examples, the precise nature of the relation will depend on the nature of the third relatum - the mode of presentation, "Idea," or whatever entity the particular triadic theory posits.

Second, let us consider the entities involved in the belief relation, namely, *propositions* and *modes of presentation*. Triadic theorists generally adopt a theory of structured, Russellian propositions consisting of properties, relations and particulars, much like the propositions discussed in the last chapter. For example, on Crimmins'

view, a proposition is an ordered set of such properties, relations and particulars. Thus, the proposition that Hesperus is Phosphorus is the sequence

<Identity, Hesperus, Phosphorus>

As was discussed at length in the last chapter, such propositions have fine-grained identity conditions[78]. Nonetheless, such propositions are not fine-grained enough to solve the belief puzzles on their own (since the sequence <Identity, Hesperus, Hesperus> is the same as the sequence <Identity, Hesperus, Phosphorus>). Instead, the third element of belief - the modes of presentation - have the extremely fine conditions required to differentiate beliefs. The central idea here is that properties, relations and particulars *present* themselves in different ways to cognitive agents. Thus, for any particular (like Venus) or property (like masticating), there will be several distinct modes of presentation of that particular or property [79]. And one can rationally attribute inconsistent properties to an object such as Venus, by have beliefs involving Venus under one such mode (say, the mode of *being the morning star*), and different beliefs involving Venus under another mode (say, the mode of *being the evening star*). For example, the triadic theory can account for the difference between believing trivialities and believing informative identities in roughly the following

---

[78]Such that, for example, the propositions that

Hesperus is brighter than Phosphorus

and

Phosphorus is dimmer than Hesperus

are distinct, in virtue of being distinct sequences with distinct constituents.

[79]Or, in the idiom of Forbes [00], “guises” under which objects and properties appear; or, in the idiom of Crimmins, “Notions” of those objects and properties; and so on.

way:

> <Believes <Sparky, <Identity, Venus, Venus> <M*Identity*, M*Eveningstar*, M*Eveningstar*>>>>

and

> <Not <Believes <Sparky, <Identity, Venus, Venus> <M*Identity*, M*Eveningstar*, M*Morningstar*>>>>[80].

Since, by hypothesis, M*Eveningstar* and *MMorningstar* are distinct entities, the two beliefs do not contradict each other, any more than, say,

> Sparky is between a and b,

and

> Sparky is not between c and b

Of course, for such a triadic-belief account to serve as a response to the substitutivity puzzles at all, M*Eveningstar* and M*Morningstar* (and all the other relevant modes of presentation involved in these beliefs) must really be distinct modes of presentation. Now, in order for us to determine whether M*Eveningstar* and M*Morningstar* are really distinct entities, we must first determine what they might be. Schiffer [92] provides an extensive (though by no means exhaustive) list of possibilities:

> -Individual concepts, uniqueness properties of the form the property of being the unique instantiator of such-and-such property
> -General properties
> -"stereotypes," perhaps what cognitive psychologists call prototypes,
> -"Characters," functions from contexts of utterance to the expression's content
> -Public language expressions,
> -mentalese expressions,
> -Functional roles

---

[80]Where M*Identity*, M*Eveningstar*, and M*Morningstar* are the "modes of presentation" of Identity and the planet Venus, respectively.

> -causal chains [leading] from mentalese expressions to the objects and properties for which they stand.

Whatever modes of presentation are, they must meet a certain functional constraint if they are to do the work of differentiating beliefs. That is, in any proposition believed, it must be possible for the particulars or properties in that proposition to have distinct modes of presentation M1 and M2 such that it is possible for an agent may have beliefs involving M1 without having beliefs involving M2. Following Schiffer, we will call this the functional-role constraint. As noted above, if modes of presentation do not meet this constraint, then they can hardly do the work of differentiating beliefs in informative identities and beliefs in trivialities (or other intuitively distinct beliefs). Each of the possibilities above meets the functional-role constraint. For example, individuals typically have several distinct general properties or even unique properties (such as *being the first star to appear* in the evening and *being the last star to disappear in the morning*). There are typically many distinct public language expressions (i.e., words) that present Venus. And there are many different roles Venus may play in an individual's mental life (for example, the roles corresponding to the just-mentioned descriptions of Venus). (Nonetheless, I will later argue that at least some of these entities fail to solve the substitutivity problems on other grounds - either they are too fine-grained, or they are not plausibly available to attributors, and so on.)

Finally, we turn to the analysis of belief reports. By hypothesis, we are taking seriously the intuition that report pairs such as

Sparky believes that Hesperus is brighter than Mars,

and

Sparky believes that Phosphorus is brighter than Mars

may differ in truth value, even though they express Sparky's relations to one and the same proposition. But the reports do not *overtly* specify what mode of presentation he believes that proposition under. Rather, the relevant mode of presentation is somehow determined or specified by the context of the utterance. The theory is thus *semantic* in that the 'modes of presentation' appear in the semantic content of the belief report, even if there is a *pragmatic* component which is used to specify part of that content (as opposed to theories in which the "modes of presentation" are *not* part of the content of a report, but are only conversationally implied by the report).

It would be extremely time-consuming and tedious to criticize every possible variation on the hidden-indexical theory, with every possible kind of proposition and every possible type of mode of presentation. In what follows, I propose rather to focus on one particular, state-of-the-art hidden-indexical theory. This is Mark Crimmins' theory from Talk About Belief [92]. Although my criticisms will necessarily be directed at Crimmins, it is hoped that the reader will see how the criticisms can be generalized and extended to further three-place hidden-indexical theories.

B) Talk About Belief

Crimmins' book Talk About Belief presents a detailed, sophisticated version of the hidden-indexical theory. Here, I shall discuss those aspects of the theory relevant to the substitutivity puzzles.

According to Crimmins, belief is a three-place relationship between an agent, a (fine-grained, structured) proposition, and an "ideational belief." An ideational belief, in turn, is a concrete cognitive particular, an unshareable, agent-bound mental representation. It is composed of *notions* (concrete, particular, agent-bound representations of objects) and *ideas* (similar representations of properties and relations). Ideational beliefs have two salient properties: first, they are structured particulars, not shared or even shareable by multiple agents. This non-shareability follows from the fact that they are complexes of agent-bound particulars. Second, ideational beliefs have contents; namely, the content of an ideational belief is a proposition.

Finally, Crimmins gives an account of how we report these beliefs. Typically, belief reports only explicitly mention the proposition believed, and the ideational belief is an *unarticulated constituent* of the report. The belief is specified by the context in which the report is made[81]. On Crimmins' account, the content of a belief report of the form

Sparky believes that Rab (that a bears relation R to b) is:

E*t* (Sparky, (A, *t*, <R; a, b>) &
Responsible (n_a, *t*, r2) &

[81]Much as, say, *here,* or *now, or in New York* might be an unarticulated constituent of the report 'It's raining.'

Responsible (n_b, *t*, r3) &
Responsible (i_R, *t*, r1))[82]

In plain English, this means: there is some "thought map"[83] or ideational belief *t* that Sparky has, and *t* has as its content the proposition <R; a, b>. This ideational belief in turn has a structure which mirrors the structure of the proposition it represents: the notion of a is responsible for the second "constituent" or "role 2" of the ideational belief *t*, the notion of b is responsible for the responsible for the third "constituent" or "role 3" of *t*, and the idea of R is responsible for the first constituent of *t*. In brief, Sparky has a belief t with the content P, where the structure of t mirrors the structure of its content. Following Crimmins, we will represent the structure of a thought map with the familiar set-theoretic notation instead of the cumbersome 'the notion of a is responsible for...' clauses. This is for simplicity's sake and does not affect the discussion materially.

Of course, if this is not augmented with some way to specify *which* notions are involved in the ideational belief, then the familiar substitutivity problems will arise once more. For suppose Sparky has some ideational belief with the content that Hesperus is Hesperus (say, the ideational belief <*I*dentity; *N*Hesperus, *N*Hesperus>). Then we may correctly report that

Sparky believes that Hesperus is Hesperus,

and

---

[82]I drop the reference to the time of belief for simplicity's sake.

[83]A thought map is a specific sort of structured ideational belief. The details of the difference are not relevant to the issue at hand.

Sparky believes that Hesperus is Phosphorus.

Both reports will be true, since we have not yet required that the reports specify *which* ideational beliefs (i.e., which notions of Hesperus) are involved in Sparky's belief, only that *some* ideational belief is involved. We have not yet said anything about, for example, when a report specifies a "trivial" or an "informative" ideational belief.

To this end, Crimmins gives an account of how reports may add a specification of the beliefs, ideas and notions involved in the agent's believing. The idea is that the context of the belief report provides the notions involved in a reported belief. In the case at hand, the notions are supplied by the very names 'Hesperus' and 'Phosphorus' used in the belief report. But the way a context can supply a specification of the notions and ideas involved in a reported belief can be more complex and subtle than this. We will return to this point later.

Finally, let us briefly consider the modality-switching issues raised in earlier chapters (specifically, the issues exploited by the inferring argument). Although Crimmins never discusses the inferring problem or the associated modality problem, triadic/hidden-indexical theories seem well-equipped to address them. This is so because an act of inference can be naturally thought of as involving (or at least requiring) a certain relation between the *contents* of one's ideational beliefs. Thus, the hidden-indexicalist is free to offer something like this natural analysis of inferring:

> An agent A validly infers some ideational belief C from an ideational belief P, only if the *content* of P really does entail the content of C (i.e. the content of C is true in every world where the content of P is true) .

Moreover, recall that these contents are simply Russellian propositions - they have not undergone rephrasal or been analyzed in terms of anything like hidden descriptive content, and hence they have not changed their modal values. So the inference reports which are intuitively true, remain true upon analysis.

Although the triadic/hidden-indexical theories address several outstanding problems, important issues remain. For by now the reader should be accustomed with the need to steer between the Scylla and Charybdis of fine-grained individuation and shareability. Although we want beliefs to be finely individuated enough to solve our substitutability puzzles, we want them to be shareable (in some way) by multiple cognitive agents. And we want to be able to make true belief reports of agents who may have very different cognitive resources than we do (for example, agents that speak other languages, or agents that possess very different mental associations from us). And so the natural question to ask of Crimmins' theory is this: Given that notions, ideas, and all such representations are particular to an agent, how may agents share beliefs? And how may we accurately report the beliefs of other agents, if we do not have the same notions they have? In other words, when we make belief reports, how is it possible for us to specify the notions involved in the reported beliefs? I will argue that, at least on Crimmins' account, this is simply not possible in a large number of cases. There is no plausible account of how agents can share beliefs or even beliefs of the same relevant type. And correspondingly, the account of how agents in different circumstances may accurately report each others' beliefs also fails.

Before we turn to these issues, however, I would like to look at some more general issues for hidden-indexical theories. For in addition to the issues involving the constituents of such hidden-indexical reports, there are issues involving the form of these reports. For, as we shall see, whether or not reporters *can* actually specify the notions involved in reported beliefs, it appears implausible that they ever have the intentions to do so.

C) Two problems about form

i) The linguistic form problem

According to Crimmins, then, belief reports express a triadic relation between an agent, a proposition, and some hyper-fine-grained ideational belief. Moreover, this ideational belief is an unarticulated constituent of the report. But such an analysis seems unnecessarily unfaithful to ordinary language. A report that Sparky believes that a = b, for example, *seems* to tell us that Sparky believes some proposition involving the identity relation and one or two individuals. But on the account at hand, it reports something about the *has the content of* relation, some assorted mental entities (ideas and notions), several instances of the *is responsible for* relation, and the aforementioned identity proposition. When I reported Sparky's beliefs, I said all *that*? This seems *ad hoc* at best.

Moreover we are given no independent reasons to think that reports actually have

some third argument place as a part of their content. Crimmins offers analogies to demonstrate that unarticulated constituents are fairly commonplace in natural language expressions (see footnote 71, for example). But in each of these examples, linguistic evidence about the ways sentences may be manipulated gives us good evidence that there is a hidden argument slot in the sentence used to make the report. There is also semantic evidence: there is no way to assign a truth-value to 'It's raining,' for example, without adding an argument slot for *where* it is raining. But not so for belief reports. We simply are given no independent, linguistic reasons for thinking that the sentences expressing belief reports contain a hidden argument. And if there is no reason to think that the sentences have this hidden component, it is somewhat difficult to see how we can use these sentences to express (not merely convey) these hidden components. I think it would be a real drawback for a theory of belief and belief report for that theory to commit itself to the claim that these sentences must express such hidden constituents via wholly non-semantic means

I do not think this is a conclusive objection. But I do think it highlights a desideratum of a theory of belief and belief attribution: a theory of such attributions ought to remain as faithful as possible to the grammar of the sentences used to make the attributions. Otherwise, there is the danger that we simply will be unable to explain how the sentences have the resources to express the propositions they do, and moreover, the theories will look more and more *ad hoc*.

ii) The meaning-intention problem

Many philosophers simply find triadic/hidden-indexical theories *prima facie* implausible, regardless of their technical details or particular implementation. I believe this stems from two sources. First, there is often a sense that propositions such as

that Hesperus is Hesperus

and

that Hesperus is Phosphorus

really do differ from each other, and this is more than simply a mere illusory appearance of difference to be explained away via some pragmatic sleight of hand. Although I share this feeling, I appreciate that one might consider this to be *merely* a feeling, one which might perhaps turn out to be unreliable. We are after all talking about highly theoretical entities here, and perhaps many of our seeming intuitions about these entities are simply intuitions about the way we happen to use these entities in various assertions and reports.

I believe the other main reason philosophers find hidden-indexical theories so implausible is more compelling, however. That is as follows: it is extremely hard to believe that ordinary, garden-variety belief reports involve such complex structures, covert but tacitly mentioned relations, and highly technical entities (such as “modes of presentation”). Even highly competent speakers find themselves hard-pressed to discover any reference to covert “modes of presentation” - much less *particular* modes of presentation - in their beliefs or belief reports.

As Schiffer [92] points out, the ways belief reports must refer to unarticulated

constituents - their tacitly mentioned thought maps or modes of presentation - are very unlike paradigm cases of unarticulated constituents. For consider a paradigm case, such as

It is raining.

In this case, there are clearly unarticulated constituents, such as the time and place of the raining. If one asks someone who utters 'It is raining,' they will surely be prepared to say something like 'Yes, I meant that it's raining in Boulder, Colorado at 2 PM, August first' (perhaps with a roll of their eyes at the obvious pedanticism).

Similarly, consider cases of restricted discourse domains. Suppose Sparky utters:

All the beer is in the fridge.

Obviously, Sparky did not mean that all the beer *everywhere* is in the fridge; the most natural account is to say that he has tacitly restricted his domain of discourse. And although it might not be clear from context alone just what this restriction is, we can find out quite easily by simply asking Sparky what he meant: 'Do you mean all the beer in the house? Do you mean all the beer you just bought? Do you mean all the beer for the upcoming party?' And Sparky, that obliging chap, will be forthcoming in telling us exactly how he intended to restrict his domain of discourse. (He was referring to all the beer he just bought, but not all the beer in the house; there is still beer he wants us to put in the fridge for his upcoming party).

On the other hand, imagine the blank stares from Sparky if we were to ask him 'Under what mode of presentation do you believe that Hesperus is Phosphorus?' or 'What notions of Hesperus and Phosphorus do you have, such that you have a belief

with the content that Hesperus is Phosphorus, involving those notions?' Or even worse - God forbid - 'What are the properties of the notions you have of Hesperus and Phosphorus, such that you have a belief with the content that Hesperus is Phosphorus, involving those notions?'[84] The problem would be even worse if we were to consider Sparky's belief attributions, such as when Sparky reports something like

O'Leary believes that Hesperus is Phosphorus.

How will Sparky respond to the question 'What notions, ideas and thought maps of O'Leary's are you covertly referring to?' We can surely expect blank stares from Sparky here, even though Sparky is not only an extremely competent English speaker, but also a philosophically sophisticated individual who has read quite a bit about belief and belief attribution.

The moral of the example is that if belief reports have notions, ideas, and thought maps as unarticulated constituents, those entities are at best merely tacitly referred to in the reports. And this is not a usual sort of tacit reference, but one almost entirely unavailable to the reporter. Most reporters are not only totally unaware that they are tacitly referring to such modes, but would strongly deny that they were.

This insight is at the root of Schiffer's "meaning-intention problem" (Schiffer 92).

---

[84]The properties of the notions involved are extremely important for Crimmins' account. For recall that Crimmins treats notions as agent bound particulars. Thus *my* report of Sparky's beliefs cannot actually contain Sparky's notions; it must specify these notions somehow, via the properties the notions have or the type they fall under.

Schiffer points out that, while we may have reasons to accept that speakers sometimes tacitly, unconsciously believe things that they are wholly unaware of at a conscious level, tacit meaning-inetentions seem to be a different matter.

To mean something (by some utterance), it is required that one intend that the audience comes to believe that something, and that one intend that the audience comes to believe *that thing* is what you wanted them to come to believe. So on the hidden-indexical account, for Sparky to *mean* that O'Leary believes Hesperus is Phosphorus to some audience via his utterance, Sparky must intend that his audience come to believe that O'Leary has some thought map with the content that Hesperus is Phosphorus. Which thought map does Sparky tacitly intend his audience to have (tacit) beliefs about? There are many ways Sparky might describe O'Leary's thought map, were he consciously thinking about them[85]; and, as Schiffer points out, it is simply implausible that Sparky is even tacitly intending that his audience come to have beliefs about O'Leary's thought map via one of those particular specifications rather than another.

D) Belief Sharing and Reporting

Any theory which individuates beliefs via reference to fine-grained "modes of presentation" or their ilk will have two parallel problems. That is, there will be a

[85]Even Crimmins' own discussion of belief reports acknowledges this. For Crimmins points out that we might specify notions via the property of being attached to a specific name, the vacuous property of being self-identical, the property of being a 'normal notion,' and so on.

problem with the possibility of sharing beliefs, and a problem with the possibility of accurately reporting them. Let me elaborate.

Intuitively, agents may share beliefs. Consider the following argument:

> Nebuchadnezzar denied that Hesperus was Phosphorus
> Pythagoras proved that Hesperus was Phosphorus
> Thus, Pythagoras proved what Nebuchadnezzar had denied.

This is intuitively valid. But on an account like Crimmins,' where beliefs have many components, the claim that *Pythagoras believed what Nebuchadnezzar denied* must be made more precise. It is not simply the claim that the two agents had ideational beliefs with identical contents - for we would not say that a person who had the belief <I*dentity*. N*Hesperus,* N*Phosphorus*> shared a belief with a person who had the representation <I*dentity*. N*Hesperus,* N*Hesperus*>. If we think that there are *ways* of believing propositions or beliefs that have propositions as their content, it is intuitively possible that people believe propositions in the same way or have identical beliefs with the same contents. Of course, on Crimmins' account, these "ways" are agent-bound, unshareable particulars, so it is not possible for agents to share ideational beliefs. So our shareability intuition must be that it is possible for agents to believe propositions in the same *type* of way (and not simply the type *having the same propositional content*) Our problem is then: how can we specify the type of notion, representation, etc., involved in shared beliefs? How do we describe the representations involved in shared beliefs?

There is a parallel problem in belief attribution. According to Crimmins' theory,

notions, ideas and ideational beliefs are not shareable. So an agent cannot report the belief of another agent by using the very notions of the believer, since he cannot have them. The reporter can at best have notions, ideas etc of the same *type* as the reportee[86]. And so the reporter must specify or supply the notions, etc., of the reportee by describing them or giving their type (which amount to the same thing). Following Crimmins, I will refer to this as the problem of notional specificity[87]. In this section, I will argue that Crimmins' method of specifying such notions is inadequate. Crimmins cannot specify the intuitively correct notions involved in making many common and intuitively correct belief reports, and he cannot specify the types of belief required for having beliefs of the intuitively same type.

On Crimmins' account, belief attributions provide the notions involved in the reported belief by specifying them in various ways. First, they might specify that the

---

[86]Interestingly, Crimmins [92] claims that belief reports do not merely constrain or describe the notions of the reportee, but *provide* them. I am at a loss as to what it is to "provide" such notions, however, for the reasons stated here. The notions involved in the report cannot be possessed by both the reporter and the reportee (since they are particulars). And when Crimmins discusses how the notions are provided (also see the discussion below, on de dicto and normal notions), the notions are clearly provided by descriptions like 'being the notion associated with . . .'

[87]Although Crimmins' use of agent-bound mental representations makes this a particularly thorny problem for him, I think it will crop up in one form or another even for theories that take "modes of presentation" to be shareable. For whatever "modes of presentation" are, it seems possible for agents to count as sharing belief even in some cases where they possess slightly different modes. For example, agents might have slightly different conceptual roles associated with Hesperus and Phosphorus (perhaps on the basis of the agents having better or worse inferential skills than each other), yet still intuitively share beliefs involving those roles. Agents might use different (thin) words for Hesperus, yet still intuitively share beliefs; agents might have identical descriptions of someone, say the description *being a great roman orator*, and fail to have identical beliefs; and so on.

notions involved in a reported belief are the *normal notions* of some individual. Second, they might specify that the notions involved in a reported belief are the *de dicto notions* of some individual. Although Crimmins does not directly discuss shared belief, we might extend the above accounts to say that two agents share beliefs if they each have beliefs with the same propositional content and each representation of that content involves either a normal notion or the same type of *de dicto notion*. But neither alternative seems to account for our intuitions about sharing or reporting beliefs. I will now examine each in turn.

i) Normal Notions

According to Crimmins, agents share certain kinds of representations simply in virtue of being informed members of a community. These representations are called *normal notions*. A normal notion is a notion to which a speaker attaches certain properties which most members of the community also attach to their notions of some individual (and similarly for ideas and properties). Crimmins elaborates:

> Like nearly all adult U.S. citizens, Sarah has a normal notion of [Jesse] Jackson, to which she attaches the properties of *being a politician, being a Democrat, being called 'Jesse Jackson,'* and so on. (Crimmins 92, p 158).

Thus, for a notion to be a normal notion of an individual, it must be a notion associated with various descriptions of the individual - descriptions commonly used in the speaker's community[88]. Moreover, according to Crimmins, belief reports

[88]Of course, the relevant community may be quite small; it may even be just a few individuals. For example, there is a community of people who know me as a

specify that an agent has an ideational belief involving normal notions. For example, a report that

Pythagoras believes that Hesperus is Phosphorus

reports that Pythagoras has some ideational belief involving the normal notions of Hesperus and Phosphorus (and this ideational belief, in turn, has as its content <Identity, Venus, Venus>). This at least suggests that one might use these normal notions to establish type-identity of belief. For two agents to share beliefs of, for example, the form Fa is for them to have ideational beliefs of a certain type. Namely, these ideational beliefs must have the same propositional content, and they must contain normal notions of F and a (i.e., notions associated with the descriptions most people in their community associate with F and a). But this is inadequate as an account of either reporting or sharing beliefs. Consider again the report that

Pythagoras believed that Hesperus is Phosphorus

Pythagoras of course discovered that Hesperus was Phosphorus, so he was the first to believe that Hesperus was the last star to fade in the evening. Thus, although Pythagoras associated his notion of Hesperus with the property of *being the last star to fade in the morning*, his community did not[89].

---

philosopher, and a community of people who know me as a weightlifter. What is a normal notion of me will be relative to each community.

[89]I leave it to scholars of ancient Greek to tell us what such property Pythagoras associated it with. As for fading in the morning, recall that Pythagoras was (reputedly) the first to discover that Hesperus was Phosphorus. Thus, his community's notion of Hesperus was a notion that was not associated with the property of *being the last star to fade in the morning*. But by Sagan's time, it is well known that Hesperus is Phosphorus, so Sagan's community's normal notion of Hesperus is associated with the property of *being the last star to fade in the morning*.

What is to be said here? Is it that Pythagoras lacks a (his community's) normal notion of Hesperus? This is surely not correct. Pythagoras seems to share a notion with his community; in fact, that is why it is all so striking that he associates this notion with different properties than his community associates it with. In essence, what Crimmins has done is specify "normalcy" for a notion relative to a community by referring to the properties the community uses to describe the notion. But then we should expect all the classical problems of descriptivism to arise, when members of the community fail to associate these properties of the notion. So whether an agent has a normal notion cannot be determined by whether or not an agent associates his notion with what the community associates it with. For the agent may share the same type of notion with his community and yet associate different properties with it.

There are also situations where the descriptive information available to a reporter - or even a community - is just not sufficient to do the job of individuating beliefs. Most Americans, for example, know nothing of Cicero (or Tully) other than that he was a great orator. So any notion associated with the property of *being a great Roman orator* will count as a normal notion. But now we cannot use notions to differentiate the report that Cicero is Cicero from the report that Cicero is Tully[90]. It is once more

[90]It might be suggested that such reports differ in that one conveys (whether semantically or pragmatically) the information that the agent has two distinct notions, whereas the other conveys that he only uses one notion. (See Salmon [91] for an example of this.) But even this is problematic. It will not, for instance, help us when there are *three* names (say 'a,' 'b' and 'c') for Cicero. We will still not be able to differentiate reported beliefs that a=b from reported beliefs that b=c or that a = c.

too easy to share beliefs.[91] Two agents may each have beliefs with representations involving the same type of notions (normal notions) and the same propositional content, yet intuitively be in disagreement.

Finally, consider cross - community reporting. Recall that individuals belong to different communities, whereas what is a normal notion of an individual is relative to a community. So, for example, there is a notion of Devon that is "normal" relative to my weightlifting friends, and a notion that is normal relative to my philosopher friends. Now suppose Arnold is one of my weightlifting buddies, and Sparky is one of my philosophy buddies, and neither is a member of the other's relevant community (they have met briefly and in passing). Sparky reports,

Arnold believes that Devon is from Denver

Which normal notion is Sparky attributing to Arnold? Surely not the philosophy-notion; Arnold does not possess such a notion (and moreover, Sparky knows that Arnold possesses no such notion). On the other hand, it is quite possible that Sparky not only has no relevant weightlifting-notion to attribute to Arnold, but that Sparky does not even know what community Arnold is in (and hence what sort of normal notion he has). It seems that Sparky can really only specify Arnold's normal notions in unacceptably over-general terms, such as:

---

[91]The point is even worse when one considers individuals, such as the John, the clerk at the local 7-11, for whom there are surely no normal notions. It is dubious that even within his work community, people associate him with the property of being a 7-11 clerk, any more than they associate him with any other property (such as being the guy who drinks coffee without creamer). And even if there were notions of John that were normal in the 7-11 community, you and I surely do not have those notions of John, whom we may have just met on the street.

> Arnold has a belief involving *some* notion which is normal relative to *his* community (whatever it is), with the content that Devon is from Denver.

These considerations show that normal notions can neither do the work of individuating beliefs, nor of accounting for shared belief. And the matter cannot be improved by considering associations with the property of *being associated with the name 'Such-and-such,'* as I will now discuss.

ii) De Dicto Notions

Consider an intuitively true report, such as

> Nebuchadnezzar believed that Hesperus was not Phosphorus

Crimmins suggests that such *de dicto* belief reports specify that agents have notions associated with certain names:

> What happens in these de dicto reports is that the speaker raises a notion to salience by raising a name to salience. The speaker provides a notion as an unarticulated constituent by using, in the report, the very name that [the believer] associates with that notion (Crimmins [92], p 166)

In other words, a report that Hesperus is Phosphorus reports that an agent has a representation involving a notion associated with the word 'Hesperus' and another notion associated with the word 'Phosphorus.' Crimmins calls these "de dicto notions." So just as the above belief report is true because it reports that Nebuchadnezzar has a mental representation involving two distinct notions (one associated with the property of *being named 'Hesperus'* and the other associated with

the property of *being named 'Phosphorus'*)[92], we might suppose that two agents share beliefs of the same type if they each have beliefs involving the same type of de dicto notion. But this yields unacceptable results.

To see why this should be so, recall our earlier Rock Hudson example, where Rock is rechristened with a "nature name" at a new age naming ceremony. His nature name is, coincidentally enough, 'Rock.' Suppose that Rock's agent is aware of his new age rechristening, but Rock's camp counselor is unaware of his previous acting career. The agent believes that Rock is Rock, and this is informative. Presumably the counselor does not believe this. The counselor believes the triviality that Rock is Rock. And the difference in the agent's and the counselor's beliefs cannot be explained in terms of the notions involved in their representations. For each of them are using only the *de dicto* notion of Rock - i.e., the notion associated with the property of *being called 'Rock.'*[93]

In addition, the usual problems arise with cross-linguistic attribution. Consider again our earlier example:

Pythagoras believes that Hesperus is Phosphorus

If a modern English speaker reports this, he is surely not attributing *his* de dicto

---

[92]As opposed to the false report that

Nebuchadnezzar believed that Hesperus was not Hesperus

[93]Of course, this objection rests on the "received view" that words are simply physical entities like orthotypes or phonotypes. The central claim of this work is to reject such a view. And according to the theory of "thick words," "Rock" and "Rock" in this example are different words, so the objection does not hold. Nonetheless, once we have a theory of thick words, I believe there is a much simpler and more elegant account of belief and belief attribution, which I will sketch at the end of Part I.

notion to Pythagoras. For Pythagoras has no notion he associates with the word 'Hesperus.' And the modern speaker probably has no idea what words Pythagoras associates with Hesperus, and hence is not in a position to specify Pythagoras's de dicto notions.

De dicto notions thus cannot be used to report or differentiate beliefs in a wide variety of ordinary cases. Although they escape several of the flaws of ordinary descriptivism, they nonetheless inherit many other flaws.

E) Conclusion

The triadic and hidden-indexical theories buy their power at a steep price. Such theories individuate beliefs (and belief reports) by referring to very fine-grained representations or modes of presentation of those beliefs. And because beliefs (and attributions) involve both the hyper-fine-grained modes of presentation as well as more coarse propositional content, triadic theories can avoid many of the problems that would arise if beliefs involved *only* the propositions containing modes of presentation (or even descriptions). For example, triadic theorists may avoid Kripke's arguments against descriptivism[94] such as the modal arguments, since the *contents* of belief contain no descriptions. Nonetheless, many of the problems of descriptivism and Fregeanism still arise. Triadic theories have difficulty explaining how it is

[94]And this includes theories such as Frege's, where the contents of belief involve not individuals but descriptive senses of those individuals, as well as straightforwardly descriptivist theories like the theories discussed in chapter I of this work)

possible for agents to share beliefs, or even to share beliefs of a relevant type. Similarly, hidden-indexical theories have trouble explaining how agents may correctly attribute beliefs, since it does not always appear that agents have the necessary resources to specify what representations are involved in other's beliefs. The problem is especially poignant for Crimmins, because his theory holds that the mental representations relevant to believing are *never* shared but can only be specified or described. Nonetheless, any three - place account will face a similar dilemma, with the details largely depending on what the theory takes modes of presentation to be. These modes of presentation may be extremely fine grained, in which case we will have difficulty correctly reporting the representations involved in other agent's beliefs. On the other hand, the modes of presentation may be coarse grained, in which case the theory will treat agents as having beliefs which they do not seem to have.

The final straw is that the theory just appears far too psychologically complex. According to the theory, we have to have very complex intentions involving the properties of other agent's representations simply to make correct belief attributions. And whether these intentions are tacit or otherwise, it simply does not appear plausible that ordinary reporters have them, let alone unsophisticated reporters. And embedded reports are even worse. If I report that Sparky believes P, then according to the hidden-indexicalist, I am making claims (tacitly or otherwise) not only about what representations I have, but also about what representations Sparky has. But we seem to have no such intentions.

Appendix: Beliefs, Attributions and Pragmatics

The triadic and hidden-indexical theorists claim that beliefs and belief reports are individuated not solely by reference to the propositional content of a belief, but also by some further entity, such as a *way* the proposition is believed, or a mode of presentation of that proposition, etc. And the *content* of a belief report includes not only the proposition believed, but this further information as well.

There is a serious alternative to such triadic accounts which takes these points seriously, while maintaining that belief itself is a simple two-place relation. According to this alternative view, belief relates a believer to a structured proposition involving individuals and properties. Thus, an agent who believes that Hesperus is bright really does also believe that Phosphorus is bright (since the propositions that Hesperus is bright, and that Phosphorus is bright, are the same proposition). However, when we report this agent's beliefs, we *impart* or *convey* information about some further fine-grained entities (whether these are sentences the agent would accept, guises, modes of presentation, some further descriptive proposition, etc.) This information is not literally part of the content of the belief report, but is pragmatically conveyed by the belief report. According to the pragmatic theorist, it is these pragmatic implications of belief reports which individuate the report. So, for example,

Thales believes that Hesperus is Hesperus,

and

Thales believes that Hesperus is Phosphorus

in fact have the same content (and truth value). What individuates these two reports is not their *content* but rather their pragmatic implications. According to the pragmatic theorist, the former report might conversationally implicate some further information such as:

> Thales believed that the last star to set in the morning was the last star to set in the morning

while the latter report might implicate:

> Thales believed that the last star to set in the morning was the first star to arise in the evening.

Of course, the details of what information is conversationally implicated by a belief report (and how it is implicated) may vary from theory to theory. Such information might be some sort of descriptive proposition, as in my example; but it might instead be information about what sort of sentences the agent would accept, or what sort of mental representations the agent would use, and so on.

I believe this so-called pragmatic solution is misguided. For it really seems to me that the reports of Thales' beliefs above *must* differ in truth. And it really seems to me that we *do* in fact believe different things when we believe informative identities and when we believe trivialities.

In this appendix, then, I would like to consider a representative of such a pragmatic approach to belief attribution. This is the account given by Scott Soames' in Beyond Rigidity, which I take to be the most sophisticated version of this kind of theory at present. Following this, I shall raise several objections to the account.

Firstly, it is unclear what sort of pragmatic mechanisms can be used to convey this further individuating information in a belief report. Secondly, even if belief reports pragmatically convey some further individuating information, this further information cannot do the job that it is supposed to do. It cannot individuate reports like those about Thales' beliefs, such as those we have been discussing above. For this further individuating information is the same sort of information that failed to work for the triadic theorists. Finally, since belief reports pragmatically convey certain propositions over and above the propositions literally *meant* in the report, it is fair to ask which propositions - the *meant* or the *implicated* - are involved in events like inferring. And thus ultimately, I think the pragmatic theory will fall prey to the inferring argument given in chapter I. Before I discuss these objections in depth, I would like to outline Scott Soames' account of the pragmatics of belief reports.

Soames' Account

Soames' account of belief attributions falls out of an elegant general account of meaning and content. Soames points out that an assertion (whether that assertion is a belief report or some other sort of utterance) may *impart* or *convey* many different kinds of information, depending on the context. For example, suppose two agents both believe that Hesperus is the last star to fade in the morning, and moreover, each knows that the other knows it, and each knows that the other knows that he knows it, and so on. In such a context, an utterance of 'Hesperus is bright' will of course convey the information that Hesperus is bright. But it is also highly likely that it will

also convey the information that the last star to fade in the morning is bright. Indeed, it probably conveys a whole host of related propositions, depending on the context and the auxiliary information shared by the agents involved.

If assertions and reports convey or impart many different propositions, what then is the "semantic content" of a report? According to Soames, the semantic content is that (Russellian) proposition which is imparted in every context[95]. This has the effect that an assertion such as 'Hesperus is bright' has the semantic content <Brightness, Venus>; but it also imparts information such as the first star to fade in the evening is bright, or the celestial object named 'Hesperus' is bright, or any number of related propositions[96] depending on what propositions the utterer reasonably expects his audience to come to believe on the basis of his utterance in that context.

According to Soames, agents often make assertions intending to convey information apart from or in addition to the semantic content of the assertion. When

---

[95]More precisely, C is the semantic content of an utterance U iff a) C is included in the information a competent speaker would intend to convey by an utterance of U, b) U is used sincerely (i.e. non-metaphorically, ironically, etc) c) C is not defeated or canceled by some further pragmatic implicature, and (d) there is no other proposition Q such that the fact that Q satisfies the previous requirements explains why P does as well. (Soames [02] pp 60-62). Requirement (d) is present to ensure that, for example, the trivial logical implications of P (such as P ^ P, or ∃xPx) are not counted as the semantic content of U as well. (See also Soames [02] p 210)

[96]Such propositions will also be sequence-like entities. I have refrained here from committing to the form taken by a proposition such as that the first star to fade in the evening is bright. Notably, Soames also seems to refrain from committing to exactly what sorts of propositions might be conveyed. He does not, for example, explicitly rule out propositions containing things like Crimmins' "notions," or Fregean senses. However, I think that the objections I plan to offer here will apply equally to each sort of conveyed proposition (with the exception of the propositions I plan to discuss in the next chapter).

an agent uses an assertion to convey more than one piece of information, we can speak of the information that the agent primarily intended to convey, or what Soames calls the agent's "primary intention." This information will be what Soames calls the "primary assertion."On Soames' account, it is likely that an agent might primarily intend to convey something other than the semantic content. Obvious independent examples of this are pragmatic phenomena such as metaphor or sarcasm. An assertion of 'Bush makes Nixon look like a choirboy' literally means (has as its semantic content) the proposition that Bush makes Nixon appear to be a young, devout Catholic acolyte who sings in the choir; but it is intended to convey a quite different proposition (that while Nixon was evil, Bush is so evil that Nixon's evil appears to be virtue by comparison). In realizing his intention to convey this further information, an agent uttering this sentence metaphorically takes advantage of contextual features such as general conversational rules as well as knowledge of the interlocutor's mental state. And this is precisely what happens in belief attributions. An agent utters a report such as

Sparky believes that Hesperus is bright

and while this has as its content the information that Sparky believes <brightness, Hesperus>, it may be used to convey the further propositions that, for example,

Sparky believes <∃x: Bright(x) ^ the last star to fade in the morning (x)>,

or

Sparky believes <∃x: Bright(x) ^ the star named 'Hesperus' (x)>,

or any number of propositions attributing a belief in similar descriptive propositions to Sparky. What further propositions are conveyed (and even which proposition was the primarily intended) by the report may be indeterminate, and there may be more than one. Moreover, the further propositions conveyed by an assertion or report will be a function of what knowledge the asserter believes Sparky to have, and what knowledge the asserter believes their audience to have. Finally, the further information conveyed by an assertion or report is a descriptive, Russellian proposition (i.e., a sequence of individuals and properties, and no further "modes of presentation") such as the two propositions just considered.[97]

Pragmatic Phenomena

If the extra content conveyed by a belief report is in fact conveyed pragmatically,

---

[97]Soames never explicitly commits to this claim, and never tells us what the general form of such conveyed propositions must be. Nonetheless, every example he uses is some sort of Russellian descriptive proposition of the sort we are already familiar with - a proposition of the form *the individual that has such-and-such descriptive property (metalinguistic or otherwise) is such-and-such*. Moreover, Soames doesn't explicitly extend the treatment to Mates-style examples. However, I assume that the treatment of a report like

Sparky believes masticating is rude

follows suit, conveying that

Sparky believes that the property known as 'masticating' is rude.

I assume that these will all be metalinguistic descriptions, since ordinary descriptive properties will simply raise the original substitutivity problems all over again.

In what follows, I will assume that Soames' account of the conveyed content of a report always involves such descriptive Russellian content. I don't think the account *needs to* incorporate this claim. Of course, if one held a Soames-style account of belief reports without this claim, one would need to provide an account of what *kind* of propositions are conveyed by reports. Naturally, I think that the most plausible such account is one like the account I will sketch in Ch 5.

it is difficult to see how this occurs. The extra content seems to be conveyed neither by conversational implicature, nor in virtue of being a presupposition of the reports. For the extra content fails the various tests standardly used to test for implicature or presupposition.

Conversational implicatures are supposed to be cancellable or defeasable. For example, consider the assertion

Shane got on his horse and rode off into the sunset.

This assertion cannot have as its *content* that Shane did these things in that order. For one might utter - without contradicting oneself -

Shane got on his horse and rode off into the sunset, but not in that order.

Thus, that claim Shane did these things in that order is supposed to be a conversational implicature of the assertion, since it can be canceled.

However, consider the cancellability test applied to a Soames-style analysis of assertions:

Hesperus is bright (and Phosphorus is not) , but 'Hesperus' doesn't name Hesperus

Such an assertion would make no sense at all. Since the added lexical content (that 'Hesperus' names Hesperus) cannot be canceled in this way, it is not at all clear that it is a conversational implicature of the assertion, for the cancellability test is usually taken as crucial for such implicatures.

On the other hand, the extra content does not appear to be a presupposition, either. This content fails another standard test for presuppositions, the "plug test."

The presuppositions of sentences are standardly "plugged" by certain verbs, like 'said' or 'asks.' That is, when a sentence is embedded in another sentence containing a verb like 'said,' the presuppositions of the embedded sentence are not inherited by the larger sentence. For example,

> John stopped beating his wife (presupposes that John beat his wife)
> Sparky said that John stopped beating his wife (does not presuppose that John beat his wife)

We can also apply the plug test to assertions like

> Hesperus is bright (said to presuppose that Hesperus is named 'Hesperus')
> Sparky said that Hesperus is bright (seems to convey that Hesperus is named 'Hesperus,' not plugged by 'said')

However, the information that Hesperus is named 'Hesperus' does in fact appear to be conveyed by the embedding sentence. That is, 'said' does not plug the information that Hesperus is named 'Hesperus.' Thus, this information does not seem to act as standard presuppositions act.

This is not, of course, devastating. Perhaps presuppositions act differently in intensional contexts. Perhaps the "plug test" should be considered more of a "guideline," than a necessary condition for presupposition. And perhaps there is some well-behaved category of pragmatic conveyance other than conversational implicature or presupposition[98]. Nonetheless, I think that at the very least, the burden

[98]For example, perhaps this information is conveyed via what Grice called conventional implicature. Conventional implicature occurs when words become associated with some non-semantic content via long-standing convention (perhaps 'the ides of march' being associated with doom is one such implicature). But I do not consider here the possibility of such conventional implicatures. Firstly, conventional implicatures are highly controversial categories of pragmatic implicature. Secondly, the information supposedly conveyed by assertions and ascriptions does not seem to

of proof has been shifted to the pragmatic theorist. If assertions and ascriptions are supposed to convey lexical information about their constituents *pragmatically*, then the pragmatic theorist really owes us a story about the pragmatic mechanisms involved.

A Posteriori Necessities

According to Soames' account, an utterance such as:

> Hesperus is Phosphorus

expresses (has as its content) the proposition

> <Identity, Venus, Venus>

This is of course necessarily true, and knowable *a priori*. But the sentence may also be used to convey the further information that, say,

> The thing that is the last star to fade in the morning is identical to the thing that is the first star to appear in the evening

And this proposition is contingently true, and knowable only a posteriori. There is not, Soames claims, any identity proposition[99] which is both known a posteriori and is necessary. The most that can be said is that some sentences may be used to convey both necessary a priori propositions and contingent a posteriori propositions.

---

be conveyed as a matter of convention. Finally, such conventional implicatures do not seem to have the context-sensitivity that most pragmatic theorists require of belief reports (see especially Saul [99] on this point).

[99]I leave out here propositions expressed by sentences with partially descriptive names. While some of these identity propositions may be necessary and a posteriori, Soames' doctrine of partially descriptive names creates new problems in itself, such as the counterintuitive result that the sentence 'Princess Diana might not have been a princess' is meaningless or false.

I do not find this result especially problematic. I do not think that Kripke's doctrine of the necessary a posteriori should be treated the way philosophers treated Aristotle for a thousand years, as dogma and established truth (however, I am aware that many philosophers might find such a denial of the necessary a posteriori an unacceptable consequence). More importantly, the issue raises a similar point which I think is compelling. That is, on Soames' account, "mixed modals," or reports involving combinations of epistemic and metaphysical modalities, appear to (implausibly and counterintuitively) commit speakers to asserting false propositions.

False reports

Consider a situation where Sparky asserts that

Thales knew that Hesperus was Phosphorus,

or

My son knows that epidermis is skin.

On Soames' account, such reports state (have the semantic content that) some agent believes the propositions

<Identity, Venus, Venus>

and

<Identity, skin, skin>.

Thus, strictly and literally, these reports are true, although in most contexts they will be misleading because they will convey the false propositions that some agent believes the propositions:

> The thing that is the last star to fade in the morning is identical to the thing that is the first star to appear in the evening,

and

> The thing that is referred to by the word ‘epidermis’ is identical to the thing that is referred to by the word ‘skin.’

As I have indicated, I find it hard to swallow that these reports are merely misleading. But the problem is worse. Consider similar reports involving terms expressing epistemic relations like knows a priori or (just now) discovered.. On Soames’ account, intuitively true reports of this sort must express something false! For consider reports like:

> Pythagoras discovered that Hesperus was Phosphorus,

or

> My son just discovered that epidermis is skin.

According to Soames, these reports each *literally mean* (have as their semantic content) that some agent discovered or learned the proposition

<Identity, Hesperus, Hesperus>,

or

<Identity, skin, skin>.

But Pythagoras surely did not discover that Hesperus was Hesperus; he knew it all along; and the same goes for the child’s discovery about skin and skin. To say that an agent discovered something is to say he did not know it at some point[100]. Thus,

---

[100]Soames concedes this point, although it could be made even more starkly in reports such as the intuitively true:

> Pythagoras did not believe P, but then discovered it (i.e., P) was true.

according to Soames, these assertions express something that is literally false - although they are used to convey some other things that are true. But when we assert in this way that Pythagoras or some child made a discovery, it does not at all seem that we assert *anything* false - and especially not the semantic content of our report, regardless of whether it is our primary intention to convey that semantic content!

But this is precisely what Soames is committed to. Among the many propositions conveyed by an agent's utterances, agents sometimes (or often!) unknowingly assert falsehoods. According to Soames, this is so for two reasons. Firstly, agents involved in a discourse often simply are not concerned with the semantic content of an utterance (in the way that agents typically ignore the semantic content of an utterance of, say, 'Nixon was a choirboy'). Thus, they may inadvertently wind up asserting some false proposition that is ignored by all the discourse participants. Secondly, agents may simply be unaware of the content and the consequences of what they are saying, due to ignorance of the linguistic rules involved.

This seems implausible. Of course, Soames is correct that agents sometimes are unaware of what they are really saying, and that they can be led to see the content of their utterances through careful, step-by-step argumentation where the linguistic rules are all laid out. But it is difficult at best to see how an ordinary assertion (in ordinary contexts) of

Pythagoras discovered that Hesperus was Phosphorus

could be used to assert the false (and absurd) claim that Pythagoras discovered something he knew all along (viz., that Venus is Venus). Firstly, the case seems quite

different from similar pragmatic phenomena. If I assert, 'Nixon was a choirboy,' and someone responds 'I had no idea Nixon was even Catholic,' I will roll my eyes at the pedanticism, and explain that, 'alright, Nixon isn't *really* a choirboy.' I will readily agree that what my assertion literally meant was a falsehood. But no such response is likely to occur when an agent asserts that Pythagoras discovered that Hesperus was Phosphorus. Secondly, it is unclear what rules the asserter is supposed to be unaware of. We can assume that the asserter understands that the names 'Hesperus' and 'Phosphorus' corefer, and that the asserter knows what 'discovered' means and entails. Perhaps the asserter is unaware of the linguistic rule that the sole semantic content of a name (or predicate term) is its referent. But this is precisely the rule that is disputed here

This issue is really just a symptom of any theory that treats the referent of a name (or predicate) as its sole semantic content. Pragmatic theories such as Soames' attempt to explain away the counterintuitive consequences of this general claim by asking us to concentrate on propositions other than the semantic content. But even if it is correct that assertions convey propositions other than their semantic content, these theories are still committed to the claim that ordinary, intuitively unobjectionable claims must assert something absurd, in the face of all our intuitions.

Cross-linguistic attribution, and other old objections

Since the propositions that Soames claims to be "primarily asserted" by an utterance are just versions of our old friend the descriptive proposition, we can expect

many of the issues of descriptivism to be lurking in the vicinity. On Soames' account, the sorts of descriptive propositions conveyed appear to be of two varieties. First, there are general descriptive propositions, such as

The thing that is the last star to fade in the morning is bright,

and second, there are metalinguistic descriptive properties, such as

The thing that is referred to by the term 'Hesperus' is bright.

In fact, the second sort of proposition is simply a specific sort of the first sort of proposition. Both propositions are of the form *The F is G*; metalinguistic descriptive propositions simply incorporate a very particular sort of description in *the F. . .*, a description involving very fine-grained entities indeed. But there are problems supposing that a speaker conveys either sort of proposition. Let me start with the problems surrounding general descriptive propositions.

I think the main problem in claiming that agents primarily assert these general descriptive propositions is that there are plenty of cases where agents simply do not have enough descriptive knowledge to do the work. Again, most Americans (myself included, despite the benefits of a classical education) do not know anything more about Cicero (and thus Tully) other than that he was a famous Roman orator. Some of us are lucky enough to know that Cicero and Tully were the same person, but I would wager that most Americans don't know this. Thus, it is hard to see how any of these Americans could use and utterance of 'Cicero is a Roman' to convey any general descriptive proposition that is not also conveyed by an utterance of 'Tully is a Roman.' Utterances of 'Cicero is Roman' and 'Tully is Roman' will not, in this case,

differ in *either* semantic content *or* conveyed information. Thus Soames' account simply cannot account for the intuition that these two utterances convey different information in some way.

To put a finer point on it, it is unclear what true descriptive proposition Soames thinks could be conveyed by a report such as 'Sparky thinks that Cicero is a Roman, but Tully is not.' For by hypothesis, this (as asserted by most Americans) conveys the information

> Sparky believes <CONJ <being a Roman, the x: x is a famous roman orator>, NEG <being a Roman, the x: x is a famous roman orator>>

If one thought that the utterance attributed some non-contradictory proposition to Sparky, surely it is not *this* proposition. And again, by hypothesis, there is no other general descriptive propositions that either Sparky, the reporter, or the audience believe about Cicero.

The final drawback to such descriptivist accounts is that they seem to be ill-suited for handling substitutivity puzzles involving predicate terms, such as our earlier Mates-style puzzles. Soames must hold that a pair of assertions such as

> Masticating is ruder than chewing
> Masticating is ruder than masticating [101]

do not differ in semantic content, but instead differ in the general descriptive propositions they pragmatically convey. But what general descriptions of the properties of *chewing* and *masticating* are available? After all, those properties themselves serve as descriptions of objects; if descriptions of objects are too coarse-

[101] And, of course, similarly for belief reports containing these sentences.

grained, then why should descriptions of descriptions be any different? Perhaps such assertions convey information about the extensions of the properties involved, such as

> Masticating is ruder than the property had by things that are not rude, and
> Masticating is ruder than the property had by things that *are* rude.

But this strikes me not only as completely *ad hoc*, but also as being extremely unlikely to be conveyed by any assertion. For there seems to be little way for an audience to determine what general (and non-vacuous) description of the property's extension is being conveyed.

I believe it is considerations like these which drive description theorists to claim that assertions and belief reports sometimes convey metalinguistic descriptive propositions. For an agent making an assertion involving (for example) Hesperus, or a report involving another agent's beliefs involving Hesperus, will typically at least possess some name of Hesperus (i.e. the name used in the assertion or report). And even in cases where we know nothing else that we could use to distinguish (for example) Cicero and Tully, we can at least describe them as having different names. But even this approach is saddled with familiar problems.

First, there is the familiar issue of cross-linguistic attributions. Agents simply are not typically in a position to know what names are used by speakers of other languages, as I have pointed out before. The claim that a proposition is "conveyed" by an utterance rather than being the content of the utterance does absolutely nothing to change this. Moreover, absent a more robust theory of words and names, Soames' account does nothing to address the 'Paderewski' and 'Rock' (see Ch 3,C, ii of this

work) puzzles. For just as it is possible for an agent to falsely believe that one and the same general description applies to two distinct entities, these examples show that it is possible for an agent to falsely believe that one and the same name (or thin word) applies to two distinct entities.

In sum, we have already seen that neither general descriptive propositions nor metalinguistic descriptive propositions really serve to do the work of disambiguating assertions or belief reports. And this holds whether those propositions are the literal semantic content of a report, or are merely conveyed by the report.

Inferring, revisited

Consider once more a report such as:

> Sparky validly infers that Phosphorus is dimmer than Hesperus from his belief that Hesperus is brighter than Phosphorus .

This has as a necessary condition that:

> Sparky believes that Venus is brighter than Venus, he believes that Venus is dimmer than Venus, and that the first proposition he believes entails the second proposition he believes.

On Soames' account, however, the report conveys other propositions as well. For it is highly unlikely that anyone uttering this report primarily intends to convey the information that Sparky believes that Venus is brighter than Venus (for example). The primary assertion is, according to Soames, surely involves something like:

> Sparky believes that the thing named 'Hesperus' is brighter than the thing named 'Phosphorus,' Sparky believes that the thing named 'Phosphorus' is dimmer than the thing named 'Hesperus,' and . . .

And filling in the ellipsis is problematic. Does the reporter mean to convey that Sparky made an inference from the proposition <Brighter-than, Venus, Venus> to the proposition <Dimmer-than, Venus, Venus>? Surely not! Surely, the reporter intends to convey that the inference was between those of Sparky's beliefs that he intended to report - viz., the descriptive propositions he intended to attribute to Sparky.

In this case, the analysis renders the inference report true. For in every world that the proposition that *the thing named 'Hesperus' is brighter than the thing named 'Phosphorus'* is true, it must also be true that thing named 'Phosphorus' is dimmer than the thing named 'Hesperus.' Nonetheless, Soames' account will still fail to give an adequate explanation of inference reports such as

> Sparky believes that Hesperus is Phosphorus, and from this he validly infers that necessarily, Hesperus is Phosphorus.

For the same problem arises that we discussed in chapters 2 and 4. For intuitively, the report claims that Sparky is making an inference between two propositions - not the uninformative propositions that are the semantic contents of the reports, but the propositions that the reporter primarily intended to report. In other words, the report primarily intends to convey that:

> Sparky believes that the thing named 'Hesperus' is identical to the thing named 'Phosphorus,' Sparky believes that necessarily, the thing named 'Hesperus' is identical to the thing named 'Phosphorus,' Sparky inferred the latter from the former, and the former does in fact entail the latter.

The trouble here is, of course, that just as we discussed previously, it is highly unlikely that Sparky believes that

> Necessarily, the thing named 'Hesperus' is identical to the thing named

'Phosphorus.'

Not only that, but the inference report will be, upon analysis, false. For the "premise" proposition does not in fact entail the "conclusion." (For a fuller discussion of this, see ch 4).

In short, all of the inferring problems of metalinguistic descriptivism seem to apply to Soames' pragmatic theory. For on Soames' account, inference reports primarily assert that agents have metalinguistic descriptive beliefs, even if these beliefs are not the literal semantic content of the report. Thus, while the problems with inferring will not apply to the semantic content of the report, they will nonetheless apply to what is "primarily conveyed" by the report.

In the end, this really seems to be the central flaw of pragmatic accounts. Theories which incorporate descriptions into the semantic content of assertions and reports face certain problems we have discussed. If we simply shift these descriptions into what is pragmatically conveyed by a report, we simply shift these problems as well. Perhaps the pragmatic accounts need not involve such straightforwards descriptivism. In this case, it would be reasonable to ask just what sort of non-descriptivist propositions are conveyed by assertions and reports, and correspondingly, what sorts of propositions are the objects of belief. And I intend to do this in the next chapter, although I will not commit to whether these propositions are "literally meant" or "conveyed."

Chapter 5

Word, beliefs and propositions

In previous chapters, we have investigated many theories of fine-grained content. The central theme of these theories has been that the contents of beliefs will often be individuated not only by the objects, properties and concepts involved in those beliefs, but also by linguistic entities - words - used to express those beliefs. But these theories were all found wanting on multiple grounds. Such linguistically augmented beliefs often impermissably change the modal value of the beliefs, they involve overly complex and apparently ad hoc structures, and they are not appropriately shareable across linguistic communities. It is now time for me to offer an initial sketch of my own positive view.

A) Thick words

My central claim is that the words involved in such beliefs are not words as commonly understood. Rather, they are what I shall call *Thick Words* - abstract objects similar to properties, which have their referents necessarily. (It will become convenient hereafter to use double quotes “ ” to indicate thick words, in addition to their normal function as “scare” quotes. I will be explicit about which I am using if the context does not make it clear.)

This central claim enables us to avoid the objections which beset other theories

of belief involving linguistic entities. First, since words have their referents necessarily, beliefs involving words retain whatever modal value they had before analysis. Thus, we will not be prey to problems that arise surrounding phenomena that are sensitive to modality, such as inferring. Second, the structure of propositions involving such abstract words, as I will explain, mirrors the structure of the clauses used to express such propositions. We need invoke no further metalinguistic relations in analyzing propositions. An attribution like

Sparky believes that a is F

will state that O'Leary is related to a proposition which actually has the form

a is F,

and not some more complex proposition involving a hidden relation of expression, such as

the thing expressed by 'a' is F.

Finally, the treatment of words as property-like abstract objects allows us to describe the structure of these objects. Just as properties have independent instances, so too words have instances (or what I will call presentations). These instances are the particular spellings and pronunciations of the words, which may vary considerably. Given such variance, the possibility remains open that different languages may contain distinct presentations of the same word[102]. As such, although presentations may vary from language to language, the words presented may be

[102]I say 'the possibility remains open,' because much depends on the form one's theory of abstract words takes. I do intend to endorse a theory according to which words are shared by languages.

shared by languages.

B) The Logical Structure of Beliefs and Attributions

Belief is a (two-place) relation between a cognitive agent and a proposition. There is no further "mode of presentation" of the belief which a believer 'has in mind' or is otherwise related to by the belief relation. The propositions believed by an agent are sometimes ordinary objectual propositions involving only properties, relations and individuals. However, in certain cases, agents instead believe propositions involving words. These propositions, however, do not involve any further metalinguistic relations, syntactic trees, annotated nodes, or word/object pairs. Rather, they have the form of ordinary singular propositions.

Obviously, there is a small problem. Ordinary singular propositions predicate properties of individuals, and not of words. The claim that Hesperus is bright should not be confused with the (false) claim that the word 'Hesperus,' thick or thin, is bright. So we cannot simply place abstract words in the subject position (or the predicate position, if it is the predicate which raises the substitutivity problems, as in our masticate/chew examples).

What is required is a richer notion of predication. Bealer (1993) and Moffet (2001) discuss propositional constructions in which predication appears to work differently from ordinary, 'singular' predication, or $pred_s$. We can think of such singular predication as something like a function which takes a pair consisting of a property and an individual, and maps that pair onto a (proposition that is) true just in

case the t individual has that property. However, not all predicative sentences exhibit this sort of structure. As Moffet [03] points out, generic sentences (i.e., sentences involving genera) such as

The lion has a mane

do not appear to involve such singular predication. ‘The lion’ presumably refers to the genus of lion, and the genus itself does not have a mane, but rather the *members* of the genus have manes . Moreover, certain constructions involving predicating properties of properties do not appear to involve singular predication. For example,

Blue is pretty

does not appear to say that the property of blue (or set of blue things, or whatever analysis of properties one prefers) is itself pretty, but rather that blue things are pretty. Finally, propositional constructions involving definite descriptions also do not seem to involve singular predication. If the semantic content of ‘The King of France ‘ is some individual concept (as many theorists seem to hold), then

The king of France is bald

surely does not predicate baldness of the concept of being the king of France, but rather of its extension.

Let us call this sort of predication *descriptive predication,* or $pred_d$, after Bealer. Let us say that descriptive predication is a function taking a pair of individuals and/or properties (or property-like objects, such as words) and mapping them onto a true proposition just in case the extension of the first entity is in the extension of the latter property. What I propose is that in certain problematic belief contexts, propositions of

the form

that Hesperus is bright

be analyzed as

$Pred_d$ (brightness, “Hesperus”)

and propositions of the form

Sparky masticates

be analyzed as

$Pred_s$ (“masticates”, Sparky)[103]

This promises to resolve several puzzles. First, it gives a natural and semantically “innocent” account of the objects of belief. They are not ad hoc constructions, functions, possible worlds and so forth; they do not contain predicates and relations (such as the ‘expression’ relation) that agents are unaware of; their logical form looks just like the logical form we expected of propositions in belief reports. Moreover, the device of descriptive predication is no ad hoc device, but one in common use in natural language.

Second, the treatment easily resolves the belief report substitutivity problems for both names and predicate terms. For the pair of propositions

[103]It is not immediately clear how to generalize such an account to n-place predication. One might, for example, opt to treat any n-place predicate as a 1-place relational predicate, e.g., the result of applying the two-place loving relation to an individual such as Sparky, to create the 1-place predicate of loving Sparky. This seems inelegant and counterintuitive. One might also attempt to give a generalized account of n-place descriptive predication. This, however, is a technical topic for another day.

$Pred_s$ (“masticates”, Sparky)
$Pred_s$ (“chews”, Sparky)

are clearly distinct from each other. The former predicates a property of Sparky via the linguistic mode of presentation that is the word ‘chews,’ the latter does it through a different linguistic mode of presentation. To believe the former does not commit one to believing the latter. This holds for our Hesperus/Phosphorus problems as well.

Moreover, recall that thick words have their referents necessarily. Thus thick words cannot be identified as shapes or sounds (or ‘types’ of shapes and sounds). For shapes are contingently, not necessarily, associated with what they express. Even though the word “masticating,” on my account, necessarily expresses the chewing property, we might have used a different shape or sound as an instance of the word “masticating.” (I will speak at length on this point in the next chapter.) But if words are not identified as their spellings and pronunciations, it seems highly likely that the word “masticate” is the same word as the word “macher” (French), and the Greek word “mastikhan,” and the Latin word “Masticare.” If this is so, then we can in fact use such “linguistic propositions” to make correct belief reports about speakers of other languages. We can correctly report that Jean-Jacques, Aristotle, and Cicero all believed the proposition that $Pred_s$ (“masticates,” Socrates),

Finally, as was noted earlier, the analysis does not change the modal value of the propositions analyzed. For example, consider the proposition

that Hesperus is Phosphorus.

This is a necessary truth. On my account, in certain belief contexts, the proposition actually believed is:

$Pred_d$ (Identity, “Hesperus,” “Phosphorus”)

Since “Hesperus” and “Phosphorus” each necessarily refer to Venus, this too is a necessary truth.

Finally, I want to stress that this is not a universal theory of belief and attribution, nor a wholly general response to the substitutivity puzzles. There are doubtless cases where the substitutivity puzzles arise in contexts where beliefs do not seem to involve language, and I do not intend to address those cases (although I think the general form of the answer will mirror mine). But there are undeniably cases where our relevant doxastic access to entities is only through words. I think in many cases, we can have beliefs about things simply in virtue of having picked up the right words. And these cases I think are admirably suited to the analysis I have offered.

This, then, is the general outline of my theory of belief. Of course, important questions remain. First, it is not immediately obvious that words do have their referents necessarily. Second, it is not clear what sorts of objects words are. Third, it is not clear what relation holds between thick words, their ‘presentations,’ and the things they express. Finally, it is not clear to what extent such words are shared among languages, if any. These issues will be the guiding concerns of the second part of this work.

Part II

The Ontology of Words

The ontology of Words

Overview

The previous part of this dissertation has advanced the claim that agents sometimes believe propositions involving what I have called *thick words*. These thick words are abstract objects which have their referents necessarily. I will call this general claim about words the *theory of thick words.* But according to most conventional wisdom, words are not thick. Thus, if the best solution to the substitutivity puzzles involves the existence of thick words, so much the worse for the theory of belief and belief attribution.

Among the intuitions reported by the conventional wisdom about words are that words have their referents as a matter of convention, that words thus have their referents contingently, and that words can change their referents. I will call this, broadly, the *theory of thin words.* The theory of thin words is itself the conjunction of two component theses - the thesis that words are merely physical objects or types of physical objects, and the thesis that words only have their referents contingently. I will refer to these theses as the *reductive thesis* and the *contingent-convention thesis*, respectively. According to conventional wisdom, these theses are both motivated by the body of linguistic practices, usages and embedded intuitions which I will refer to as "folk linguistics." Far from using the term dismissively, I think that folk

linguistics is our best source of intuitions about the nature of words. But the evidence from folk linguistics is not as the theory of thin words would have it.

Supporters of the theory of thin words are quite right to point out that we already have a host of usages, practices and embedded intuitions about words. They are also correct to point out that, if the only motivation for thick words is (for example) to solve some problems in the philosophy of mind or the theory of belief attribution, well, then, so much the worse for the philosophy of mind. Unfortunately for the thin words theorist, this is a point that in the end *supports* the theory of thick words. For once we focus on folk linguistics, it becomes evident that there are in fact very few ordinary usages that give evidence for the theory of thin words. More startlingly, I will argue that the great weight of evidence from folk linguistics implies that the sorts of words we use in communication are not merely physical entities, nor are they contingently related to referents. On the contrary, this folk-linguistic evidence appears to show that many words are in fact individuated by their referents. Consider, for example, our treatment of homonyms (i.e. distinct words within a language which share both pronunciation and spelling), and false cognates (i.e. distinct words in different languages which share pronunciation and/or spelling). Such words are usually - and intuitively - individuated solely by their referents.

What, then, are we to make of the remaining (and rather forceful) folk-linguistic intuitions which *appear* to support the theory of thin words? Here I argue that the intuitions invoked by the thin words theorist are not necessarily intuitions about

words *simpliciter*; rather, they are intuitions about shapes, sounds, etc. And it is not obvious that these are the only kind of words there are. To claim that these shapes and sounds are the only kinds of words is simply to beg the question about the identity of words. For we also have common practices, such as those alluded to in the previous paragraph, according to which we treat words as being *distinct* from sounds, and even being differentiated by their referents. In fact, deeply embedded in our folk linguistic usage is an ambiguity about the referent of the word 'word.' Rather than say that one set of usages (and the associated intuitions) is/are "correct," and the other set is "incorrect," I hold that ordinary practices commit us to two sorts of words - a thick and a thin word. 'Word,' then, is *ambiguous* as to which sort of word is meant[104]. We sometimes use "word" when we are interested in talking about the physical manifestations of thick words, i.e. shapes and sounds; but more often we use the distinct homonym "word" to refer to thick words.

In Part II of this dissertation, then, I intend to focus on folk linguistics and the ontology of words. First, I embark on a more in-depth discussion of the two theses involved in the theory of thin words - the reductionist thesis and the contingent-convention thesis. Then I give a lengthy treatment of the information from folk

---

[104]Recall that single quotes are treated as being ambiguous between being used to mention thick and thin words (or, there are two homonymous sorts of single quotes, one used to mention thin words and the other used to mention thick words). Thus one of the claims made here is that there are two words, "word" (referring to thin words) and "word" (referring to thick words). This is not a case of one ambiguous word but rather of two distinct homonyms.

linguistics, which indicates that ‘word’ is ambiguous between thick and thin words. The discussion of folk linguistics leads to a further consideration of intuitions about the nature of words. For example, consideration of folk-linguistic intuitions reveals the intuitions that words are created, that words exist contingently, and that words have presentations (e.g., their spellings and pronunciations) as well as referents.

At this point, however, we have not discussed what kinds of abstract entities thick words are, except in the most general of ways. There are many possible sorts of abstract entities that might, at this point in our inquiry, fit our description of thick words. For example, ordered pairs consisting of thin words and referents are abstract objects that have their referents necessarily. But certain kinds of trees are also abstract objects that have their referents necessarily. Which, if any, of these entities are the thick words we use in linguistic communication? In answering this question, I will develop and critique three accounts of thick words which are consistent with the intuitions to be uncovered from our discussion of folk linguistics - the ‘tree’ view, on which a word is identified with a tree that has its origin or root node necessarily; the sequence view, on which a word is an ordered set such as <some orthotype, some phonotype, and the referent of the word>, and the *sui generis* view on which a word is an unanalyzable simple or “marble.”[105] Finally, using the resulting, more fully developed account of thick words, we return to the positive theory of belief and belief

[105]The word ‘marble’ is here intended to connote a seamless, spherical object that cannot be broken apart without losing its identity.

attribution of the previous part, and show how this theory of belief and attribution can be used to address common puzzles and counterexamples.

Let me make a final comment on terminology here: I will use, as I have been, double quotes as in “cat ” to indicate that I am mentioning a thick word. I will use dividers as in |cat| to indicate that I am mentioning a thin word, and my use of single quotes as in ‘cat’ will be ambiguous between mentioning thick or thin words.

I now turn to a consideration of the received view of words - the theory of thin words.

Chapter 6

Two Theses About Words

The Contingent-Convention Thesis

According to the contingent-convention thesis, a word is merely contingently associated with its referent. This association is usually established through a convention (although it is possible that a mere individual intention could establish the association). Thus, the word ‘Lobster’ actually refers to the family of lobsters, but it (that very word ‘lobster’) could have referred to the species *Felis Domesticus*, had we established a different convention. In support of this claim, philosophers typically invoke the commonplace intuition that words can and do change their meanings. Moreover, conventions are arbitrary and could have been otherwise. Support for the contingent-convention thsis also comes indirectly, from the thin word thesis (which we shall discuss presently).

It is worth noting that the contingent-convention thesis is deeply embedded in the practice of formal logic. To see this, look at the way central logical concepts like ‘validity’ and ‘tautology’ are defined. A sentence is a tautology *iff* it is true under every interpretation, where an interpretation is a function assigning that sentence an extension (truth value). A sentence of logic, then, does not have its extension necessarily, since the very same sentence can be assigned a different extension and

remain the same sentence.

Finally, the contingent-convention thesis is not a theory (at least not directly) about the ontological status of words, although it may imply such a theory. It is simply a theory about the semantic relation that holds between words and their referents. However, the contingent-convention thesis is closely allied with a thesis about the ontological status of words, namely, the *reductive thesis*.

The reductive thesis

The reductive thesis is a thesis about the ontological status of words, that is, a theory about what words are. According to the reductive thesis, words are physical objects (or perhaps types, sets or classes of such objects). Thus, words are something like marks on paper, or sound waves, or combinations of the two. I call this the "naive reductive thesis." Of course, this initial claim is implausible for many reasons, such as the reasons brought up by CS Peirce in his famous type/token distinction[106]. Peirce pointed out that there is a sense in which one word - 'words' appears several times on this page. This single word which appears many times is what Peirce referred to as a word-type of which there are many tokens. Exactly what a *type* was supposed to be is somewhat unclear, but it is evidently an abstract object much like a set or a class. But while word-types and sets of words are themselves abstract objects, they are (at least

[106]Peirce [31-58], vol.4

according to received wisdom) types or sets of physical things - namely, the previously considered marks on paper or sound patterns. I will henceforth refer to these as *orthotypes* (types of orthographical objects, i.e. written marks) or *phonotypes* (types of phonological objects, i.e. sound waves). The variant of the reductive thesis I will consider is the thesis that words are reducible to entities such as orthotypes, phonotypes, or combinations of thereof. Additionally, my considerations will apply equally to the naive thesis.

The reductive thesis is embedded (albeit usually only implicitly) in the practice of logic, which treats words and sentences as merely strings of physical symbols which we may manipulate. More importantly, the thesis also has the weight of simplicity behind it, for it accounts for words without postulating any new sorts of entities. The world is already full of physical objects like tables, chairs, gases, inkstains, and so on; on the thin word thesis, words are just more of the same. And we already have reasons to countenance sets or types of such objects, along with a (supposedly) fairly good understanding of what sets and types are. And the theory is (prima facie) intuitive - for we utter sounds and we write marks on paper, and intuitively that is all there is to writing or uttering a word (or tokening a word).

I believe that both of these theses are false. I believe that words are abstract objects (not merely abstract sets of physical objects) and that words have their referents necessarily. First, I think that Part I of this dissertation shows that there are good theoretical reasons to treat words this way. But just as importantly, I think there

are solid reasons embedded in our ordinary linguistic practices that support these seemingly radical claims. In what follows, I will offer arguments and considerations from folk linguistics against both the reductive thesis and the contingent-convention thesis.

Of course, folk linguistics is hardly the final authority on the nature of words. Nonetheless, the opposition to the thick words theory appears to be mainly motivated by folk-linguistic intuitions. If one simply dismisses the evidence from folk linguistics - or if, say, the evidence from folk linguistics turns out to be inconclusive - then, presumably, theoretical considerations decide the issue. And theoretical considerations are precisely what have been offered in the first part of the dissertation.

Chapter 7

Folk linguistics and thick words

In this chapter I discuss the evidence from folk linguistics. First, I shall discuss the folk-linguistic evidence *against* the two thesis comprising the theory of thin words (that words are physical or physical types, and that words have their referents contingently). Then I shall discuss the way in which these two different sets of arguments support each other, in that if either the reductive thesis *or* the contingent-convention thesis fails, the other thesis is either severely undermined or rendered fully untenable. Finally, I discuss the putative folk-linguistic evidence for the reductive thesis and the contingent convention thesis, and show that the evidence is ambiguous at best. Such evidence is really evidence that shapes and sounds are contingently associated with referents. But, I argue, the claim that words are only shapes and sounds is false. I will conclude with the claim that we use the word "word" to refer to shapes and sounds (thin words) and we also use a homonymous but closely related word "word" to refer to thick words**.**

A) Folk-linguistic evidence against the reductive thesis

The following arguments focus on the way we associate physical entities like shapes and sounds with words**.** The first two arguments show that the possibilities of phenomena like mispronunciation and alternate pronunciation preclude the possibility

of the reductive thesis. The second two arguments show that our practices of calling two physical objects *the same word*, or even calling a given physical object a word at all, appear to preclude the possibility of the reductive account.

i) Malapropisms, mispronunciations and misspellings

In ordinary usage, we think of people as sometimes mispronouncing or misspelling a word, while nonetheless still using the original word. For example, as David Kaplan points out, when Tom Brokaw says 'a mi-we-un dollars,' he is clearly using the word 'million,' regardless of his inept pronunciation. When George W. Bush says 'Nuc-u-lar power,' he is using the word 'nuclear.' Similar points hold for misspellings. When I write 'The Indians lived on the great planes,' I am using the word 'plain,' but misspelling it. I am not using the word 'plane' and falsely stating, although I do not realize it, that the Indians lived on large flat geometrical surfaces or airborne vehicles.

Furthermore, conceive of almost any phonological or orthographic errors a speaker may commit. Ordinary speakers nonetheless still will considered the mis-speaker by to be using the word he misspells or mispronounces. Importantly, we may conceive of a word being misspelled so that it shares an orthographic presentation with another word (as in the plane/plain example above), and yet is still distinct from that other word. So, again, ordinary usage motivates the claim that a word is more than a phonotype or orthotype (or combination thereof).

There is a related point. The very concept of misspelling - as opposed to malapropism, using the wrong word - implies that a word cannot be identical with its spelling. For if one misspells a word, one must be using that very word, and not some other word. Consider the ‘plane’ example again. How can we say that ‘plain’ was misspelled unless it was the word used in the sentence? If the word in the sentence was in fact ‘plane,’ then how could it be misspelled? ‘Plane’ is in fact, after all, spelled |p l a n e|. One might hold that words are not misspelled simpliciter, but rather are somehow misspelled relative to the words one intended to use. But such a treatment would still suppose that the word used was a distinct word from the word intended, and hence was a malapropism rather than a misspelling. A malapropism is the use of an incorrect word in a context, such as stating that ‘we have parakeet flooring.’ The user is not misspelling ‘parquet’ but incorrectly using a word used to refer to pet birds. Ordinary usage treats misspelling and malapropism as distinct errors. Sometimes misspellings result in malapropisms, but not always. Of course it is often difficult to determine whether a given mistake is a malapropism or a misspelling, but sometimes it is intuitively clear (such as if our hypothetical mis-speaker went on to say something like ‘the floor material is so named for its yellowish wood finish.’)

Similar points may be made about mispronunciation. If a word is a sound, then how is mispronunciation possible? How can one mispronounce a sound? One can, after all, produce the incorrect sound, but that is a different matter entirely.

ii) Alternate pronunciations

Consider the words ‘Route’ and ‘rout.’ The former refers to a path, the latter means to inflict a severe defeat upon (as in ‘the Roman legions routed the Franks.’) The former word may be pronounced in at least two alternate ways; |root| or |rowt|, respectively. But it is incorrect to say that one |rooted| an enemy (although one might have *rooted for* an enemy, with a very different meaning!)[107]. The alternate pronunciation is unacceptable for ‘rout.’ But if there is an acceptable alternate pronunciation for one word which is not acceptable for the other, then the two words cannot be the same. Moreover, common practice would be to treat the spoken words |route| and |root| as being identical - in fact, ordinary speakers would say they are alternate pronunciations of the same word. So a word cannot be associated with a phonotype.

Perhaps, then a word is associated with its spelling and pronunciation, for the above words do differ in spelling. But we can easily extend the above sort of argument. For consider the words ‘bark’ and ‘bark.’ The former refers to the outer covering of woody plants, and the latter refers to a square-rigged ship. The latter has an alternate spelling, |barque|. But this is not an acceptable alternative spelling of the

---

[107]Further field research indicates that Australian English contains a transitive verb ‘root,’ with a shockingly vulgar meaning. An Aussie who roots for the home team is doing quite a bit more than an American who roots for the home team! (See section v and vi, below)

former word.

It seems moreover at least possible to that there be similar pairs of words which share spelling and pronunciation, but one of which has an alternative pronunciation the other lacks. Suppose we began to pronounce the word ‘plane’ (aircraft) with the alternative |planned| or |plan|, but did not accept these alternates for ‘plane’ (flat surface). Or consider the actual pair of words ‘peaked’ and ‘peaked.’ The latter word, which means *having peaks*, has an acceptable (if floridly poetic) pronunciation |peak ED| which the former word (which means *wan and pale*) lacks[108]. We would surely not say that one and the same word both had and did not have an acceptable alternative pronunciation. So again, we do not ordinarily treat a word as being identified with its orthotype, phonotype, or combination thereof. [109]

iii) Written and spoken words

Natural languages are used in at least two salient ways - they are spoken and written. The word ‘cat’ may be inscribed on a piece of paper, as I just have, or it may be spoken aloud. Intuitively, it is one and the same word which is both written and

---

[108]Thanks to Michael Zerella for the example of ‘peaked.’

[109]One might instead treat words as some sort of disjunctive set of orthotypes and phonotypes, where ‘peaked’ (wan and pale) was identified with {|peaked|(spoken), |peaked|(written)} and ‘peaked’ (having peaks) was associated with {|peaked|(spoken) v |peakED|(spoken), |peaked|(written)}. Still, this leaves us with a similar problem: any given utterance of |peaked| is an instance of *both* words, and is thus an instance of a word with a certain alternate pronunciation and an instance of a word without that alternate pronunciation.

spoken; a folk linguist will say that I just spoke/said/uttered the very same word that I wrote. Moreover, if one is asked what word was previously written, one may respond either by saying the word again or by writing it. It follows, by Leibniz's Law, that words cannot be associated with an orthotype or a phonotype, because phonotypes are distinct from orthotypes. Intuitively, the spoken or written words are *presentations* of the word. [110]

Of course, this intuition on its own is not enough to motivate thickness. For it simply shows that a spoken word and a written word intuitively have something in common other than an orthotype or phonotype, something of which the orthotype or phonotype is a manifestation. This need not be an entity which has its referent necessarily (such as a thick word), although that seems a plausible candidate. Perhaps, in keeping with the thin view, a word is to be identified with an {orthotype, phonotype} pair. But this pair seems wholly unnatural. There are, after all, many such pairs, and only some of them are words or associated with words. {|lobster|(written), |squirrel|(spoken)} is neither a word, nor even associated with one. The pair {|lobster|(written), |lobster|(spoken)} is at least associated with a word. So again, a word cannot be identified with just any such pair. Moreover, there is a reason some

[110]Of course, the written and spoken words of a language need not be related in this way. Perhaps Chinese ideograms are not intuitively the same word any some spoken word, or perhaps Egyptian hieroglyphs are not the same words as any spoken Egyptian words. We shall discuss these cases in chapter 9. For now, the point remains that in many languages, we have a strong intuition that spoken words are identical to certain written words.

pairs are associated with words, and other pairs are not, and this reason does not seem explicable in terms of 'thin' words.[111]

iv) Embedded words

Consider the word 'catastrophe.' This has a certain shape and sound. A constituent of this shape and sound is the shape/sound [cat]. But it is odd to say that the word 'cat' is part of the word 'catastrophe,' although cats are certainly parts of many catastrophes.

In a similar vein, consider a series of sound/shapes that have been run together, much as in written Latin - such as

justtoolovely.

The orthotype/phonotype |tool| is part of this string. But it would be unusual to say that this string |tool| was either a word in the sentence or a word in the string. And surely the reader can create examples in which the embedded shape shares a pronunciation with a genuine word. But, once again, we are not inclined to think that the orthotype/phonotype is a word, at least not the sort of word that speakers would use to communicate.

There is a final point to be made here. Whether or not a word has some referent

[111]For a discussion of treating words as pairs of thin words and referents, see the next section of this work, 'The Ontology of Words.' Such a pair is, notably, thick; it has its referent necessarily.

necessarily, a word has to have *some* meaning. The string of letters |wrkstrf| is not considered a word. And the embedded string above, namely, |tool|, has no meaning, whether necessarily or contingently. So it is not, intuitively, a word.

The above examples show that words cannot plausibly be reduced to physical objects or physical object-types (or sets of physical objects) such as orthotypes or phonotypes. Even so, it is still possible that words are some sort of abstract objects which have their referents merely contingently. In this next section we will look at folk-linguistic evidence against this contingent-convention thesis.

B) Folk-linguistic evidence against the contingent-convention thesis

The arguments in this section fall into three rough types. First, I consider ordinary practices by which we individuate words by their semantic properties - as in the cases homonyms, cognates and equivocation. Second, I consider practices by which we individuate words based on their syntactic qualities. Finally, I consider the practices of word individuation commonly employed by experts, such as linguists and lexicographers.

v) Homonyms

Natural languages are considered by its speakers (and specifically, by the

presumed “masters” of that language, i.e., its teachers) as having homonyms [112] - distinct words that have identical orthotypes and phonotypes. For example, English has the word ‘plane’ and the word ‘plane’; one word refers to a flat surface, and the other word refers to an aircraft. Grade school grammar classes instruct quite explicitly that these words, *although they sound and are spelled alike,* are different words. Moreover, what distinguishes them is, as the example points out, their referent. It is also worth noting that grammar teachers are sensitive to both distinctions in words and distinctions in sounds - for there is also the category of *homophones*, or words that share a phonotype but not an orthotype (such as ‘plain’ and ‘plane’). The fact that we make such a distinction between homonyms and homophones suggests that we treat homonyms as being distinguished by more than merely thin characteristics. Clearly, if a given utterance of ‘plane’ had been used to talk about aircraft, it would not have been an utterance of the other word ‘plane’ (the one involving surfaces). The referent of ‘plane,’ even according to folk linguists, is an essential part of its identity. If it had referred to something else, it would have been a different word - a homonym.

The example of homonyms shows clearly that words cannot be thin. But it also seems to show that what makes an occurrence of ‘plane’ an instance of one word

[112]From the Greek, *homo,* same + *onomo,* name. Whether this was originally intended to mean that there is one name for two distinct words, or that the two homonyms are actually the same name/word, is unclear to me.

rather than another is its referent. Hence, the referent is essential to the word.

vi) Cognates and false cognates.

In natural languages, it is quite common for distinct languages to share orthotypes, phonotypes, and even ortho-phonotypes (strings of symbols which are both inscribed and pronounced the same way). When these entities share referents, they are referred to as "cognates"; when they do not share referents, they are referred to as "false cognates."

For example, in German, the symbol written as |r o t| and pronounced |*rOt*| refers to the color redness. The English word with the same pronunciation |*rOt*| and the spelling |r o t e| refers to a mechanical method of learning or memorizing. The English entity and the German entity clearly share the same phonological presentation. But ordinary language users treat these as different words, and intuitively they *are* distinct words. Thus, a common way of naturally understanding words is incompatible with treating those words as phonotypes. The same point can be made at a sentential level. The German sentence 'Empidokles liebt,' which means that Empidocles, a person, loved, shares the same phonotype and pronunciation with the English sentence 'Empidocles leaped,' which expresses the proposition that Empedocles jumped. But again, intuitively, the sentences are distinct, which entails that sentences are not to be associated with their pronunciations or with a phonotype.

The example of entities in different languages which share both orthotype and

phonotype is even more stark. For example, consider ‘si’ in Spanish and ‘si’ in French. The former expresses the affirmative, the latter expresses the conditional (actually the latter has a homonym, but this homonym need not concern us here). The rather strong intuitions of most speakers are that these are, in fact, two distinct words. As a thought experiment, ask yourself how many words French and Spanish share. We do not, for example, answer that the Spanish word for the affirmative *is* a French word. These words cannot then be identified with their orthotype/phonotype pair. For suppose the words are these phonotype/orthotype pairs. Then

> ‘Si’(French) = {|Si|(sound), |Si|(shape)}.
> And of course, ‘Si’(Spanish) = ‘Si’(French) = {|Si|(sound), |Si|(shape)}.
> Thus by Leibniz’s law, ‘Si’(French) = ‘Si’(Spanish), or, in other words,
> ‘Si’(French) belongs to Spanish.

This directly contradicts the intuition reported above - the intuition that the Spanish word for the affirmative is not a French word.

What intuitively distinguishes these words is that they have different referents, and this is the intuition most natural language speakers will report. One might object that this conclusion is too hasty, and there are perhaps other features which differentiate ‘si’ (Spanish) and ‘si’(French). The obvious starting point is that they are symbols in different languages, and being in a different language is what makes them different words. But I believe this to be mistaken for several reasons.

First, it seems to miss the point. For the thin words, as shown above, belong to both languages. So a word-in-a-language must not be a thin word.

Second, it requires dismissing the intuitions of folk linguist. After all, the initial claim was that intuitions from "folk linguistics" motivate the identification of words with orthotypes, phonotypes or some combination thereof. But folk linguists appear to report not merely that 'si'(F) and 'si'(S) are distinct words, but that they are distinct because of their referents. So intuitions from folk linguists, if taken at face value, support the thickness of words; and if not taken at face value but reinterpreted, then we should feel free to reinterpret other intuitions from folk linguistics - specifically, the intuitions supposedly supporting the thinness of words.

Third, the claim intuitively gets it backwards. Words are not distinct because they belong to different languages, rather, languages are distinct (in part) because they contain different words.

Finally, the claim is incompatible with intuitions about borrowed words. The Dutch language, for example, borrowed the word 'toast' from English to mean a piece of toasted bread. But this is not a mere cognate; rather, it is the same word in both Dutch and English. Examples of words English has borrowed are even more abundant. 'Qua,' for example, is a word that English uses, which it borrowed from Latin; they are one and the same word, presumably because they have a certain sort of intentional, historical connection. The point also holds for borrowed morphemes like 'In,' as in 'inflexible,' which English borrowed from Latin to mean the same as 'not,' just as it does in Latin. It also holds for borrowed idiomatic phrases such as

'Vis-a-vis.' So being in a different languages cannot be a sufficient condition for words being different.

vii) Equivocation[113]

There is a prominent school of logic which treats logical relations as relations holding between words or sentences, and not between "propositions" or the "meanings" of those sentences. Call this the "sententialist" view of logic. But such a view, coupled with the reductive thesis, cannot account for obvious phenomena like equivocation. For example, it is clear that the argument

1) John is fishing from the bank
2) a bank holds money
3) thus John is fishing from a place with money

is invalid.1 and 2 may obviously be true without 3 being true. Intuitively, the reason why is that 'bank' in 1 and 'bank' in 2 do not refer to the same thing. But if validity is a relation between words and sentences, the meanings of those sentences are not directly relevant. And if words are not differentiated by their referents, then the same word 'bank' appears in both 1 and 2. If one wants to hold that validity is a relation between sentences, then one must hold that words are not merely orthotypes and phonotypes, or one will be forced to treat the above (obviously invalid) argument as being perfectly valid. Moreover, if one denies that the words 'bank' in 1 and 'bank'

[113]Thanks to Chad Carmichael for the beginnings of this argument.

in 2 are the same word, then it is natural to claim that this distinction exists because these words refer to different things.

Of course, if one abandons this sententialist conception of logic, then this consideration is moot. Still, I would view an argument against such a sententialist conception of logic as being very substantive in its own right.

viii) Alternate syntactic constructions

The point to be raised here is a development of the point raised earlier in the homonyms section, and a variation of the point raised about alternate pronunciations. That is, there are words that share a thin component - orthotype, phonotype, or some combination of the two - which nonetheless enter into different syntactic constructions. For example:

‘I planed the surface of the board’ and
* ’I needed to get there faster than taking a car, so I planed to Dallas.’

The latter sentence is ungrammatical (and no editor of academic books would let it pass!). ‘Plane,’ when it refers to aircraft, cannot be used as a transitive verb, despite the irritating trend in modern English to transform verbs into nouns (‘officing,’ ‘transitioning,’ etc.) But if ‘plane’ is a single thin word, then it is at best unclear why it cannot be used to form a grammatically well-formed (if semantically meaningless) sentence with each of its referents.

The point may be framed even more starkly. The word ‘plane,’ according to the

thin words theory, is both a noun and a verb. Syntactic categories are properties of words. But nouns and verbs have incompatible syntactic properties. So, one and the same word has incompatible syntactic properties. Thus, ‘plane’ (n) and ‘plane’(v) are two distinct words, and they are individuated at least partly by their referents (in this case, the ontological category of the thing referred to by the word).

Sometimes we do say that a word is both a noun and a verb, especially when the two are closely related. For example, we sometimes treat ‘saw’ as both a noun and a verb, as in the cutting tool and the act of using that tool. Nonetheless, it is extremely unclear how this usage can be reconciled with the above point. Nouns can enter into constructions that verbs cannot, and vice versa. We would not want to say that one and the same word both can and can not enter into a certain construction. When we say that ‘saw’ is both a noun and a verb, then, it is not clear that we are talking about a word. Rather, what I will suggest is that we are claiming that one and the same thin word is associated with two thick words. This will be discussed more fully in the next subsection, on the ambiguity thesis.

ix) Deep structure

Consider for a moment not words, but similar entities, sentences (complexes of words). Like words, sentences which share orthotype/phonotype may sometimes express different meanings. For example, consider the sentence

‘The man saw the girl on the hill with the telescope.’

This might mean something like either

a) The man, using a telescope, saw the girl who was on the hill, or
b) The man saw the girl who was using a telescope while she was on the hill

One common intuition is that this is a single sentence, with two distinct meanings. But it seems by no means uncommon for native speakers to treat two instances of this string of thin words as being distinct sentences (or sentence-types). After all, we think of *sentences*, and not only propositions, as having a structure, specifically a syntactic structure. And of course one sentence has one structure and the other sentence has an incompatible structure. We do not speak of one and the same sentence as having incompatible structures. Rather, we usually take the syntactic structure of a sentence as being essential to that sentence, much as I have argued that we take the referent of a word to be essential to the word.

We do of course speak of syntactic ambiguity and scope ambiguity. A case of syntactic ambiguity is a case where one and the same sentence has multiple distinct deep structures. And if a sentence has multiple structures, then it does not necessarily have one of those structures rather than the other. Nonetheless, it is not at all clear what entity we are speaking of when we say that the sentence has multiple deep structures. We will attempt to resolve similar conflicts of intuitions when we reach the next subsection, 'The Ambiguity Thesis.'

x) Linguistics textbooks

Linguists commonly use a theoretical concept similar to that of a word, that is, the concept of a *morpheme*. A morpheme is the smallest unit of meaning, and words are composed of morphemes. For example:

A single word may be composed of one or more morphemes...a morpheme may be defined as a minimal linguistic sign, a grammatical unit that is an arbitrary union of a sound and a meaning.[114]

It is worth noting that this is not an advanced survey of possible linguistic theories, or a theory put forward as a hypothesis which is to be argued for or against, but a textbook. The claim about the concept of a morpheme is presented as a claim which a substantial number of linguists agree on [115]. And morphemes - the building blocks of the words studied by linguists - are thought of as being a union of a sound (or, presumably, a shape) and a meaning. If a given morpheme had a different meaning than it did, then, it would be a different morpheme than it was. This is analogous to the claim that if a number had a different successor, it would be a different number. But this is just to say that linguists treat the objects of their study -

---

[114] See, e.g., Fronkin & Rodman, Introduction to Language, pp 127 128, Saeed, Semantics, 1997, pp 56-58, O'Grady Contemporary Linguistics, p 133, and Lyons, Linguistic Semantics, 1995, pp25-27 for similar points.

[115]Notably, a large number of linguists prefer a purely formal definition for words or sometimes morphemes, such as “the largest linguistic unit which resists interruption” or “the smallest unit which can be moved around in a sentence without thereby making it ungrammatical” (as discussed by Lyons 25-27) ; this appears motivated by the urge to do linguistics as pure syntax and hence rather theory-driven. Although it appears that most linguists would expect these 3 definitions to converge, the purely syntactic definitions seem violable (‘abso-bloody-lutely’) and are often thought of as heuristics.

words and morphemes - as being thick.

xi) Dictionary Entries

In general, philosophers are wary of arguments that begin ‘Webster’s defines ‘justice’ as....’ Such argument by lexicography is rarely fruitful for conceptual analysis. Nonetheless, in the present circumstance, dictionary entries may indeed provide indirect evidence for ordinary language practices involving words.

Look, for example, in the concise OED under ‘gram.’ The first entry is for a unit of mass; the second entry, also a noun, is for a method (of staining bacteria with an iodine solution). I chose ‘gram’ purely at random; similar examples may be found on nearly every page of any dictionary. The point is that dictionaries presumably list words, and dictionaries treat these as two words, distinguished from each other by their referents. So ‘gram’ cannot be identified with a phonotype, orthotype or union of the two, since each entity shares both phonotype and orthotype.

Notice also that dictionaries are in fact sensitive to the distinction between single entries that have several alternate, albeit related, meanings (“polysemy”) and distinct entries with identical spelling and pronunciation. One entry for ‘wave’ notes that it is 1) an intransitive verb meaning to vibrate or be stirred and also 2) a transitive verb meaning to brandish (notably, they are related meanings, the latter is to impart something like the former motion to some object). But a separate entry notes that ‘wave’ is a noun referring to a sort of ridge of water. Again, this is an extremely

common example; similar ones may be found on nearly every page of a dictionary. The dictionary, then, seems to treat ‘wave’ and ‘wave’ here as different words with different meanings, not as a single word with different meanings (related or otherwise).

In the example I gave, the two distinct entries are nonetheless semantically related; the vibrating or stirring imparts a motion that looks like the motion of the ridge of water. How do lexicographers draw the line? I do not have an answer for this, and the fact that different dictionaries draw the lines differently indicates to me that it is a pretty inexact science at best. Nonetheless, where the meanings are unrelated (e.g. ‘bank’ and ‘bank’) there is the widespread use of separate entries. And for our purposes, the result that meaning plays a role in word individuation is enough to at least refute the reductive thesis.

C) Support relations between the theses on words

The theses discussed here are distinct claims. It seems at least conceivable for the contingent-convention thesis to be true independently of the reductive thesis, and conceivable that one subscribe to the reductive thesis without holding the contingent-convention thesis. For example, although the considerations in this section show that words are more than orthotypes or phonotypes, those considerations alone do not necessarily mean that words are abstract entities, or that they have their referents necessarily. To wit, one might conceivably reject the reductive thesis while holding

that words are a sort of, say, abstract historical entity which can change its referent.

Since the theses are conceptually independent in this way, I offered separate sets of arguments against each thesis. Nonetheless, at this point it would be wise to examine the support relations between the theses; for although the theses are conceptually independent, they do support each other. And more importantly, if one thesis fails, the support for the counterpart thesis will be undermined; and in some cases, if a thesis fails, the counterpart simply fails with it. If this is so, then evidence or arguments against one thesis of a pair may actually be evidence against both theses. Thus many of the arguments presented in this section will indirectly pull 'double duty'; if the direct arguments against one (but not the other) thesis prove uncompelling, the arguments against the counterpart thesis may still indirectly attack the original thesis. Moreover, the relations between the theses are interesting in their own right.

First, note that if the contingent-convention thesis fails, then the reductive thesis looks incredibly implausible. For intuitively, no physical entity like shape or sound has its referent necessarily. While I have argued against the view that, for example, the word 'cat' could have referred to the family of lobsters, it is surely the case that the sound |cat| could have been associated with a different word (and referent) - in other words, it is surely the case that the sound |cat| does not necessarily refer to the species *Felinus Domesticus*. Moreover, the intuitions about the sound |cat| and its relation to a referent seem to hold for any concrete object we might care to use as a

symbol. I suppose it is *logically possible* that there is some concrete object (or type of concrete object) we have not canvassed yet that has its referent necessarily, but this appears extremely counterintuitive at best[116] .

On the other hand, suppose that the reductive thesis fails. In this case, while it may be again logically possible that words are abstract objects which are merely contingently related to their referents, much (if not all) of the support for the contingent-convention thesis is gone. This is so since the prima facie intuition that words have their referents contingently seems to stem from the intuition that we could have used the sound |cat| to refer to the family of lobsters, and that the sound |cat| might come to be related to different referents over time. But if words are not sounds (or shapes, or other thin entities), these intuitions are not intuitions about words, and the contingent-convention theorist has little or no support for his claims. Further intuitions supporting the contingent-conventions must then be found.

Conversely, suppose that words have their referents necessarily. As we just discussed, this alone seems to indicate that words are some sort of abstract object, for concrete objects do not seem to be the sorts of things that can be necessarily related to referents. On the other hand, suppose that words are abstract objects. This alone does not ensure that they have their referents necessarily. It would be consistent to hold

---

[116]Moreover, any proposed concrete object (or type or set of concrete object) would still have to be a plausible analysis of a word or other symbol. Not just any object is suited to playing the role words play, and not just any object plausibly has the properties that words have.

that words are (for example) abstract causal-historical entities whose identity is determined by causal connections to an initial tokening of an orthotype/phonotype. Such an entity need not have its referent necessarily. Nonetheless, I think some of the arguments presented in this section rule out such a possibility.

Recap

Many of our ordinary-language considerations about the use and nature of words speak against the reductive thesis and against the contingent-convention thesis, and support the theory of thick words. Our intuitions about individuating cognates and homonyms indicate that we individuate them by their referents, and that a referent is thus an individuating characteristic of a word. Our intuitions about mistakes involving spelling and pronunciation show that words cannot be identified with characteristics such as their shape or sound, for then all misspellings and mispronunciations would simply be malapropisms. And the fact that we consider some written and spoken words be the same word speaks against identifying a word with thin (orthotypic or phonotypic) entities.

Considerations about relational properties of words also speak against their being thin. Two words may share an orthotype and phonotype, yet one of them have an alternate pronunciation the other lacks. Since a single object cannot both have and not have a pronunciation, these words must be distinct; and since they are distinct, the words cannot be identified merely with a thin word. Two homonyms may also enter

into incompatible syntactic constructions; thus by the same reasoning, they must be distinct words and not identified with any thin words. And two sentences may have distinct deep structures, yet be the same string of letters, with similar results - that a sentence, like a word, cannot be identified with a thin entity.

Finally, we showed that experts in the field of words have several widely shared practices which treat words (or similar objects) as being thick. Dictionaries, which presumably list words, list homonyms separately - that is, they explicitly identify words by referent (or at least similarity of referent). And linguists often - although not universally - treat morphemes as being individuated by their referents.

Thus, we have widespread and deeply embedded intuitions that words are, surprisingly enough, thick. In the next section, I shall address the evidence from folk linguistics that putatively supports the theory of thin words.

D) The Putative Evidence for Thin Words, and the Ambiguity Thesis

The Contingency of Coining

In ordinary discourse, we do say things like

> The word ‘cat’ could have meant something different, such as the family *Mesocritecus* (hamsters).

And this is a fairly strongly held claim. But as the contingent-convention theorists point out, this does not appear compatible with the claim that words have their

referents necessarily. For the thick word theorist holds that ‘cat’ necessarily refers to the family of felines. A word that did not refer to this family would not be the same word ‘cat,’ it would be a new word that was merely a homonym (shared the same sound and spelling). So the very word ‘cat’ we are using now, if thick, could not in fact, have referred to the family of hamsters.

Nonetheless, the intuition deserves closer scrutiny. For there is at least one other intuition that it seems to conflict with. This is the intuition that one cannot simply make an established word, by force of intention, refer to something other than what it does. Consider Lewis Carroll’s character of Humpty Dumpty in Through the Looking Glass. Humpty Dumpty tells Alice that the word ‘glory’ now mean that there’s a nice knock-down argument for you, because a word means exactly what he intends it to mean, neither more nor less[117]. But of course ‘glory’ does not mean that; it refers to a state of exaltation. Once the referent of the word is established, however that might be, an individual cannot simply change the referent of the word or use it to refer to something else. Of course, this does not rule out a collection of people getting together and creating a convention that the word ‘glory’ means that there’s a nice knock-down argument. But intuitions amongst language speakers do not appear so clear here. Have the speakers bestowed a different referent on the original word? Or have the speakers coined a new word? The intuitions of ordinary speakers do not

[117]Lewis Carroll, Through the Looking-Glass, ch 6.

seem clear cut on this issue. The claim that words might have referents other than the ones they do have, is not so clear. (But see the next subsection).

Nonetheless, there is still an intuition that seems fairly strongly held: that when the word ‘cat’ was introduced, although it referred to the family of felines, *could* have been used to refer to the family of hamsters. Let us call this the thesis of the *contingency of coining*. And I think this is in fact the most accurate statement of the claim we started with. The result is that if the word ‘cat’ has its referent necessarily, then the contingency of word introduction would entail that the word ‘cat’ might not have been the word ‘cat,’ which is absurd. And so the thesis of the Contingency of Word Introduction seems to speak against the theory of thick words.

But even this intuition about word introduction is not clear cut. Do we mean that the very same word ‘cat’ that we are currently using, could have referred to the family of hamsters? Or do we mean that when we introduced ‘cat,’ we could have introduced a word with the same sound and pronunciation, but referring to the family of hamsters, instead? I submit that there is no clear intuition deciding this issue. Moreover, even the apparent absurdity that the word ‘cat’ might not have been the word ‘cat’ does not appear so clear-cut once we consider the point just made about pronunciations. For the *apparent* absurdity that the word ‘cat’ might not have been the word ‘cat’ may really just be:

> There might have been a different thick word “cat” with the same pronunciation/spelling as the word we actually use.

And this is not an absurdity at all.

In short, the intuitions surrounding the contingency of coining seem unclear. As I will argue at the end of this section, this is because the thesis itself is ambiguous about what entity - a thick or a thin word - is introduced to contingently have its referent.

The Changing of Meaning

In addition to the above "folk-linguistics" intuition, we have a related intuition about words. Namely, words change their meanings over time, as language users engage in metaphor, or the extension of both thought and speech, or as their interests change, or even as they come to misunderstand what others have said. 'Hustle' originally meant to shake back and forth, and now it refers to illegal selling, or to "scamming" (as in "hustling someone at billiards.") And this appears incompatible with the thesis that words have their referents necessarily. Moreover, it is a stronger claim than the thesis of the contingency of coining. For the intuition is that a word, regardless of its already established meaning, may change its meaning, and remain the same word.

However, once again, the nature of the intuition is not fully clear. There are two points I would like to raise about referent change.

First, it appears rare that a word ever completely changes its referent and remains intuitively the same word. The above example of 'hustle' in fact has an (albeit

hidden) similarity of meaning which can be observed in an etymological dictionary. Its original referent was to shake, from the Dutch 'hutselen'; it was extended to meaning pushing or shoving in the 1700s, and to hurry (since pushing hurries something along) in the 1800s. In the late 1800s it meant to sell in a hurry. Indeed, the very practice of etymology embeds this insight: For us to consider that a word has changed its meaning, and thus trace that change through instances of the same word, we generally have to show that the meanings are related. And if the meanings are in fact related, then a word does not *change* its meaning so much as extend or modify its meaning[118]. In most cases, if two words have radically different meanings, etymologists list them as two distinct and unrelated words, and not as a single word which has changed its meaning (much like the point made in the earlier section about dictionary entries). So even the folk-linguistic claim that a word may change its meaning is rather heavily qualified.[119] Once again, then, we have somewhat conflicting intuitions over what appeared to be a clear cut intuition about words. And

---

[118]Thanks to Keith Gresham for this insight.

[119]Nonetheless, there may be occasions where words do appear to change their meanings radically. Consider a case in which a listener has misunderstood the meaning of a word. Suppose then that this listener transmits this word to another community, using it entirely differently from the way it had been used, and the new community comes to use this word having a new and wholly different meaning. For example, the idiom 'to beg the question' is synonymous with something like 'to assume the truth of the point raised by the question.' But journalists and undergraduates have adopted it to mean 'to ask a question' which is a radically different meaning, based on a crude misunderstanding. It seems to me that these are in fact the same idioms which have radically different meanings, and an etymological dictionary lists them as such.

again, I hope to resolve these conflicting intuitions in end of this section, when I discuss the ambiguity thesis.

The second point to be raised about meaning change is this. Even in cases where the meaning of a word clearly changes, our intuitions about what occurs are not wholly clear. Moreover, putting pressure on those intuitions can reveal surprising results.

Consider a case of putative meaning change. Suppose that over, say, five years, the word 'sick' comes to be used to mean outrageously good (as it already has in some slang)[120]. Now suppose that our old friend Sparky, while stargazing and ruminating about Hesperus and Phosphorus, hits his head and falls into a coma prior to the meaning shift, and awakens after the supposed meaning shift. Upon waking, he says to the nurse, 'Boy, I feel sick.' If 'sick' unambiguously changed its meaning, as the theory of thin words holds, then Sparky does not in fact know what he meant (if in fact he meant anything at all). But clearly he said that he was feeling ill, and he is aware that he meant this. There is surely a sense in which the nurse would say,

---

[120]At this point in the example, I wish to remain neutral as to exactly what changed its meaning. I am simply describing a case where an observer would say that the 'sick' changed its meaning, without reflecting on possible analyses of that claim.

Incidentally, the example is drawn from personal experience. As a college freshman, not aware of contemporary slang, I denied that a certain woman I found highly attractive was 'sick.' My friends knew that I was saying that she was not ill or grotesque. They were not tempted to say 'Devon is - although he does not realize - denying the outrageous beauty of that woman.'

‘Sparky knows not what he means; he doesn’t realize that he is saying that he feels outrageously good,’ and this is the sense motivated by the claim that the word Sparky used had changed its meaning. But it seems a metaphorical sense at best.

The Ambiguity Thesis

The above examples show that there are at least cases in which we use ‘word’ to indicate a thin object. And despite the presence of my ‘considerations’ about how clear our intuitions are in these cases, I do not mean to imply that these intuitions are incorrect. Nonetheless, these intuitions seem to be, upon further investigation, somewhat qualified; and they appear to conflict with other intuitions. All things considered, the best theory of words would be a theory which did not simply eliminate or discount any of our intuitions, but rather resolved these apparent conflicts while preserving most of the intuitions. And I believe that there is an approach which does exactly this: the ambiguity thesis.

Simply stated, the ambiguity thesis is this: the word ‘word’ is itself ambiguous[121]. That is, sometimes (perhaps most or much of the time) ‘word’ is used to indicate a thick entity, which has its referent necessarily, and sometimes (often in philosophical

[121]Note that here I exploit the ambiguity thesis, by remaining uncommitted to whether ‘ ‘word’ ’ expresses a thick or thin word ‘word.’ For if ‘word’ is ambiguous, as per the ambiguity thesis, then presumably the practice of using quotation marks, which are used to indicate *mention* of words, is equally ambiguous. It will remain convenient to retain the use of ‘ and ’ as ambiguous.

contexts) ‘word’ is used to mean a thin entity, an orthotype, phonotype or combination of the two. Folk linguistics thus provides us with cases prompting intuitions about words - that is, some intuitions about thin words, and some intuitions about thick words. Intuitions about the Contingency of Word Introduction appear naturally to be intuitions about thin words. Certainly, folk linguists will agree that shapes and sounds may be used to refer to whatever we decide them to refer to. We can use the marks on paper ‘glory’ to refer to a knock-down argument. And our intuitions about what happens when a community gets together and actually *does* use the word in this way are unclear, because it is not clear which word ‘glory’ we have intuitions about. If English speakers came, tomorrow, to adopt the convention that ‘glory’ means a knock down argument, it is not clear whether or not a new word has been created. We have seemingly conflicting and unclear intuitions, the ambiguity thesis says, because we have intuitions about the thin word |glory|, which has changed its referent, and we have intuitions about the thick word “glory” which is merely homonymous. In the case of the thin word, we think that the word came to refer to something else; in the case of the thick word, we think that one word actually became another word, much like an acorn might become a tree, a gamete becomes a foetus, or some clay becomes a statue -a certain process is shared, but at some point the identity of the object actually changes.

Similarly for the conflicting intuitions in the Sparky coma case discussed in the previous section, ‘The changing of meaning’. There is, intuitively, a sense in which

Sparky does not know what was meant when he says the words ‘I am sick.’ This is the sense in which Sparky is taken to be using the thin word ‘sick,’ which, unbeknownst to him, changed its referent behind his back. But there is also, compellingly, a sense in which he clearly knows what his words mean, and that is the sense in which he is using a word which did not change what it meant.

Now, the reader may imagine that these cases are arguable. The considerations eliciting thick-word intuitions above are, admittedly, somewhat complex. But the above considerations were meant to be the strongest support for the thin-word thesis. And the considerations show that even in the strongest cases for thin words, there is uncertainty about the elicited intuitions. And the preceding section gave what I take to be (some of the) evidence eliciting intuitions that words are thick. But this need not show that our intuitions about words are hopelessly at odds with each other. Rather, I take it as support for the ambiguity thesis - that apparently conflicting intuitions about words, are actually intuitions about two distinct (albeit closely related) kinds of entities. And since we use ‘word’ to ambiguously refer to both kinds of entities, there is the appearance of conflict. But the appearance is merely illusory.

E) Conclusion

This chapter was designed to achieve two important results. The first result is to give ordinary language evidence for the existence of thick words. I can imagine no other prima facie intuitions *against* the theory of thick words other than that they do

not accord with our intuitions about the ordinary ways we think of words; but if my arguments here are successful, I have shown that there are very few intuitions which really speak against words being thick. On the contrary, many if not most of our folk linguistics intuitions actually motivate the theory of thick words. At the very least, I believe this establishes that the onus is on the thin word theorist to produce arguments against the theory of thick words; at best, I believe this establishes that the concept of a word we most often use in natural language communication is that of a thick word.

The second important result is the ambiguity thesis. The thin word theorists are not wholly incorrect; some of the ways we use words imply that words do in fact have their referents contingently, and are associated with shapes or sounds. Or to put it another way, some of the objects involved in communication are shapes and sounds. But this is no reason to suppose that all words are thin, or that the only entities involved in communication are thin. Moreover, it is no reason to suppose that our intuitions about words are contradictory or divided. Rather, we are interested in at least two (and perhaps more!) distinct kinds of words which are very closely related, and which we equally refer to with the expression ‘words.’ The ambiguity thesis resolves seemingly contradictory intuitions without having to discount any seemingly clear intuitions, while at the same time maintaining that the relationship between each sort of word is a very close one (but see below).

Nonetheless, a positive account of thick words remains to be given. Moreover, as I mentioned, such an account has been complicated by introducing distinct entities -

the thick word and the thin word. Any account of words must explain the relationship between these two entities. A complete account of this is the task for a further work. At this time, however, I would like to offer a brief account of some of the salient properties of thick words which we have gleaned from the inquiry. This will be the task of the next chapter.

Chapter 8

Some salient properties of thick words

The thesis that words (of the relevant sort) have their referents necessarily has been motivated in the first part of the dissertation, and defended against folk-linguistic objections in the previous chapter. Further, the previous chapter has demonstrated that folk linguistics not only provides little evidence for thin words, but is in fact committed to thick words. More appropriately the thesis that there are categories of words that have their referents necessarily (that there are thick words) has been motivated and defended. However, as has been mentioned, there are many possible ways a word might have its referent necessarily; to put the point differently, there are many conceivable objects, with different characteristics (apart from thickness), that we might choose to identify as thick words. The task of these final chapters is to flesh out the theory of thick words by examining their ontology. In this chapter, I will discuss some of the more salient folk - linguistic intuitions about thick words, revealing several prominent characteristics of thick words.

Both the first part of the thesis (the critical section which establishes the need for thick words) and the previous section (folk-linguistic intuitions about words) have resulted in several considerations about thick words. These considerations will provide us with a starting point for the discussion of the ontology of thick words.

After all, some sorts of entities are clearly not compatible with the characteristics suggested by these considerations. In doing so, we will look at questions about the nature of thick words that have not been addressed by our previous considerations, and use these issues to further enrich the theory of thick words.

A) Words have their referents necessarily.

This has been discussed at length. To recap, we have two sources for this claim. First, the insight that many of our beliefs actually have hidden metalinguistic content, coupled with the observation that this metalinguistic content must have the same modal value as the apparent beliefs, means that words must have their referents necessarily (see part I of this dissertation.) Second, a thorough examination of our linguistic practices and intuitions about the nature of words indicates that words have their referents necessarily (see part II, section I).

B) Words are abstract objects.

Thick words are not concrete or physical entities. It seems clear after all, that any concrete entity imaginable - a sound, a shape, a gesture, etc - may be associated via convention with almost any referent. Such physical entities, if they bear referent, only contingently do so. Nor are Thick words *types* (as opposed to tokens) of such physical entities, for the same reason.

Thus thick words must be abstract entities of some sort. Abstract entities exist

independently of their instances. This implies that thick words exist regardless of whether they are being spoken or written. Whether thick words are *ante rem* or *in re* (or even strongly *in re*, where they only exist at the moment they are instanced!) remains to be seen.[122]

C) They are artifacts

Thick words are created by the intentions or conventions of individuals or cultures. This means that they are very different from other abstract objects. Platonic properties, sets, numbers, and similar objects are not plausibly artifacts. Such entities are discovered by rational agents. But words are not discovered, they are created, much like works of art (especially pieces of music).

The creation of a word, then, is a power of a mind. It is not merely connecting a concrete symbol with what that symbol is to express. It is rather the creation of a brand new abstract object with special properties all its own. Nonetheless, on reflection, it is seen to be a very commonplace and familiar sort of power. Empiricists may be inclined to accuse this view of being somehow mystical, and to insist that minds cannot simply bring abstract objects into being via will and fiat. But this is a

[122]Even most of those who think words are thin, concede that some words are abstract objects. For, words are usually taken to be associated with *types*. The two tokens 'cat' and 'cat' are distinct tokens of one word - namely, 'cat'(or rather the word-type 'cat'). Since there is but one word-type 'cat' of which all the tokens are tokens *of*, this type cannot be identified with its multiple tokens. It does not exist, for example, on any page. So a type is taken to be an abstract entity.

misguided accusation. Uncontroversial examples of humans creating or bringing abstract objects into existence are legion. As I claimed, there are a great many examples of the human creation of abstract objects. For humans create, not discover, works of music and other art.[123] Humans create inferences when they discover proofs. And these are very much like the sorts of abstract entities that words are. Moreover, rational agents can be plausibly said to create - or at least to bring into being - certain other more pedestrian abstract objects, for example, some sets. For sets are abstract objects which only exist if their members exist. And sometimes humans bring such members into existence. For example, the set {the first Plymouth Fury} presumably did not exist until the first Plymouth Fury was created by a human designer. So that designer simultaneously brought the set {the first Plymouth Fury} into existence as he created the first Plymouth Fury. Notice that this is plausibly very much like a familiar relation between a thin and thick word - commonly, that a thick word does not exist at least until a corresponding thin word is spoken or written (or otherwise tokened).[124] Moreover, breeders and geneticists create new varieties of plants and animals, and sometimes, new species. And species or varieties are abstract objects in the sense we are concerned with here.

Another source of the human creation of abstract objects comes from the *in re*

---

[123]For an example of a theoretical treatment of this, see Levinson [1980]

[124]Of course, there may be other ways to introduce a new word; producing a token is only one such possible way.

views of universals. On this view, several ordinary universals are brought into existence by the creative act(s) of a mind(s). For on the in re view, universals depend on their instances for existence. And many properties or relations can only have instances brought into being by minds. For example, consider the relation of expressing, as in

expression e expresses m.

Presumably, there are no instances of expressing without expressers to express. An expresser brings about an instance of expressing. So the universal expressing depends upon an act of an expresser (a presumably mental act) for its existence, and on the usual *in re* view, does not exist before this act. So the first expresser brings the universal of expressing into being (and similarly for the relational properties of expressing P, expressing Q, expressing the concept of a Plymouth Fury, etc.)

Finally, let me make a broader point. It is difficult to see why the creation of an abstract object is so much different from the creation of a concrete object. Of course, we cannot create concrete objects tout court - that is, we cannot bring new material substance into existence. Nonetheless we may create physical objects that are plausibly distinct from the matter they are created from. A building, a piece of paper, a car[125] - all are physical objects which are created. Why should it be so different to

[125]The reader may choose to consider these as examples of abstract rather than concrete objects. So much the better for my overall point - that we are committed to the claim that humans create abstract objects.

imagine humans creating abstract objects as well? [126]

D) They are not necessary existants

This point follows from the fact that words are artifacts. Words are created entities, not timeless entities; they do not exist before there are any communicators to create them. However, once they are created, they may or may not depend upon speakers (not on being spoken) for their continued existence. They do not exist merely at the times and places that they are uttered, inscribed, or otherwise instanced. In this sense, words are like thin-word-types. For a thin word-type exists in between the times it is tokened, and presumably for some time after its final tokening. So it is with thick words. For the same thick words are used by individuals across time: I may utter the same thick word as an Englishman uttered 200 years ago. That word has a history even when it is not being used.

Thick words do not exist in worlds without communicators. In a world where there are no language users, there are no words. This follows from the intuition that words are created. But if we take this intuition at face value there will be theoretical consequences (see the section on "What are the possible objects which could have these properties?").

---

[126]The accusation of "mysticism," in the end, seems to stem from empiricists' general suspicion that abstract objects and all our relations to them must be somehow mystical. While I will provide no arguments against this radical empiricism in this work, I will note that it is clearly a radical position.

There is a final point. Although a word is a sort of *in re* universal, it is nonetheless an instance of a certain *ante rem* universal - namely, the property of being that very word. So, for example, the thick word "lobster" is a culturally created entity, which does not exist until it is first coined. It is then an instance of the property *being the word "lobster,"* a property that is (unlike the word itself) a necessary existant (given, of course, that the *ante rem* view of properties is correct).

E) They have presentations

Words are pronounced or spelled (or otherwise presented, say via the gestures of sign language.) But, as discussed in the previous section ("Misspelling, Mispronunciation and Malapropism"), words cannot be thin, for a thin word just is a sound or a shape (and a sound cannot be mispronounced, just as a particular set of letters (shapes) cannot be misspelled.) So a word must be a thick entity. Nonetheless, words are presented via spellings and sounds. That is, a thick word is pronounced by uttering a sound and written by inscribing a shape; or, more generally, thick words given by producing thin words.[127]

Although this much is clear, it remains to be seen exactly what the nature of this

[127]It is not at all clear to me at this point what the correct locution for this relationship is. What is it that one does with a thick word when one utters a thin word? What is the relationship between the utterance, the thin word, and the thick word that results? At this point, I am not even certain if natural languages have settled on a "correct" practice here.

presentation relationship is (and, as noted, it remains to be seen what the correct linguistic practices we use to describe this relationship are). Is it analogous to the type-token relation? The sentence-statement relation? The traditional universal-instance relation? Or is it wholly new?

F) They may be shared by languages.

Thick words have histories. This follows from the very practice of etymology. Moreover, it follows from the fact that thick words may change pronunciations and spellings. Thus, a word may continue to exist in a single language over time, even as it is pronounced or spelt radically differently from when it was first created. But if a thick word can continue in this way in one language, there seems to be no principled reason that a thick word cannot continue to exist as that language changes into another language. For example, there is no reason why Latin and modern Italian cannot share the same thick words (perhaps pronounced differently, perhaps not). In fact, Latin and modern Italian intuitively do share some words in this way. And there is no reason why two languages at the same time cannot share thick words in this way. Such a sharing may be accomplished in at least two ways. First, one language may directly borrow a word from another, as French borrowed ‘Le hamburger.’ Surely the *Academie Francaise* would not be so upset at the bastardization of the glorious French language if this were a case of the creation of a new and distinct word, as opposed to the encroachment of an English word. Second, a “parent”

language may change over time into two or more languages, as Latin gave birth to the so-called Romance Languages. And a Latin word may continue on into both, say, modern Italian and to French. The Latin words are now intuitively part of the "daughter" languages intuitively, despite the fact that they are pronounced differently than the Latins pronounced these words. And thus, by transitivity, the Italian word is identical to the French word. This point will become important in dealing with cross-linguistic belief attributions.

Let us move on to a consideration of the entities that could possess these qualities. Let us, in short, give a systematic folk-linguistic theory of words.

Chapter 9

The Ontology of Words

In the last chapters, we have canvassed folk linguistic intuitions on the nature of words. This data has offered compelling evidence that words are abstract objects that have their referents necessarily. In addition, folk linguistics has yielded several intuitions about what qualities such words have. In this chapter, I will consider three theories of the ontology of thick words. These theories are Katz's *ante rem* universals view, on which words are simply traditional universals; the sequence view, on which words are ordered sets of orthotypes and/or phonotypes and semantic values (meanings or referents); and finally the abstract causal-historical view, on which words are abstract causal trees.

In the course of the chapter, I will argue against the *ante rem* universals view and the sequence view. The *ante rem* universals view, for example, does not account for our intuitions that words are created artifacts, or contingent extistents. Moreover, the view fails to respect, or is simply silent about, many intuitions about the historical properties of words. And the sequence view also has severe flaws. For example, it cannot account for substitutivity puzzles involving homonyms (such as our earlier Rock Hudson puzzle).

Ultimately the theory of words I develop is the causal-historical view, which

treats words as abstract entities which give rise to tree-like causal chains of their instances. This view has its roots in David Kaplan's seminal work "Words"; however, Kaplan never fully developed the view nor did he explicitly show how to apply it to the substitutivity puzzles (although that was his stated goal). In effect, then, this development of the theory of thick words can be seen as a development of the view originally put forth by Kaplan.

A) Words as *ante rem* universals (Katz)

Jerrold Katz, through several books[128], promoted the view that linguistics was an *a priori* discipline much like mathematics. A central feature of this claim was the thesis that words are abstract objects that exist outside of time and space. Apart from this claim, however, the thesis was not extensively developed; for example, we are not told if such objects are structured universals (and if so, what their structure is, or what essential features these universals have.

Katz does present us with a few remarks that we might use to develop a Katz-style theory of thick words, however. First, his claim that words, qua abstract objects, exist timelessly (and presumably necessarily) (insert quote here). Thus, on Katz's account, words are something like traditional, Platonic *ante rem* universals, which exist independently of their instances. Such words are related to other entities which

[128]Notably, Katz [80] [90][00].

are contingent existents, namely, the shapes and sounds we associate with words (in our parlance, thin words)[129]. While thick words have their referents necessarily, then, tokens of those words are only contingently associated with thick words (and through those thick words, their referents).

While Katz's remarks on the relationship between thick and thin words are suggestive, I think that the theory really does violence to our folk-linguistic intuitions about words. That is, we think of words as human artifacts, which exist merely contingently, and have beginnings. For example, it seems intuitive that had Shakespeare had not existed, the English language would have been poorer by about 1700 words. This is because Shakespeare created those words, he did not discover them. And it seems intuitive that had the island of England sunk, Atlantis-like, beneath the waves before the arrival of the Angles, Saxons and Jutes in the 5th century AD,, the English language would not exist. And intuitively, this is partly because the English language - began with the arrival of these three tribes. Most of the words of English did not exist, intuitively, before this.[130]

---

[129]See, for example, Katz 2000: 'Linguistic utterances and inscriptions are composite objects' (p.162), where a composite object is some sort of union of an abstract and a concrete object. It is these composites which are created (p 152), and which have the contingent characteristics of words (being spelled a certain way, having so many letters, etc).

[130]The discussion of these intuitions raises several interesting questions about the relation of words, grammars and languages. How do words compose a language? Are grammatical rules essential to a language? Are sentences thick in that they have their grammatical structures necessarily? While these questions are interesting in their own right, they are beyond the scope of this work.

It is, of course, possible for a Katz-style theory of words to accommodate at least the *appearance* of these intuitions. Katz holds that while Shakespeare did indeed discover (rather than create) these 1700 new words, he did create (or at least initiate[131]) the association between these words and certain thin words[132]. This general approach can be extended. According to Katz, while words exist necessarily, the association of thick with thin words is contingent. The association is created by the creation of a contingent entity, the *token* of the word. It is this which explains the apparent contingency of words. Nonetheless, I think this approach just does too much violence to our folk-linguistic intuitions. It seems extremely implausible and counterintuitive to hold that words - the elements of linguistic communication - are discovered the way we discover numbers or other abstract objects. For one thing, we have a very clear idea of how we go about discovering new numbers. At the simplest, we merely apply the successor function to some number we already know, and eventually we come across numbers we had not previously known. But what is the corresponding process for discovering new words? This seems a mystery.

Moreover, Katz is simply silent about many aspects of the nature of these words (or their tokens). Can different languages contain the same words? How does one

---

[131]I insert this qualifier, because it is unclear whether the intention of an individual alone can create such an association. It is possible that Shakespeare initiated the convention which associated the thick with the thin word, and the creation of the convention itself was necessarily a group effort.

[132]Or rather, "Semi-thin," since Katz's word tokens are hybrid abstract/concrete entities.

token a given word? How are the pronunciations (and other presentations, like inscriptions) of a word related to each other? One might simply add principles to a Katz-style theory in order to decide these issues. But what would the status of such principles be? Would they be analytic truths? If so, can they simply be added? On what grounds would we add such principles? All things considered, it would be preferable for such issues to be resolved by considering the very essence of the entities involved - the thick words.

Thus, although a Katz - style approach to words is not completely ruled out, and indeed has some advantages, I would like to pursue a theory of words that does not do violence to these intuitions. I shall pursue a theory of words on which words are artifacts and genuinely contingent existents. Still, should such a theory prove ultimately untenable, and words really must be *ante rem* universals, many of the claims made about the causal-historical theory of words can be applied to *ante rem* universals as well.

B) Words as sequences or sets

On the sequence view, a word is identified with a pair: {orthotype, referent}, or perhaps a triple {orthotype, phonotype, referent}. Such an identification is really just the explicit ontology of the claim noted earlier, made by linguistic textbooks: that a morpheme is an arbitrary union of a symbol and its referent.

The sequence view has at least two great attractions. First, it is a very simple,

parsimonious view which adds very little to the ontology. The ontology already has thin words, sets, and referents; thus, there are sets composed of thin words and meanings or referents. Second, On the sequence view, linguistically augmented propositions turn out to directly involve the properties or objects one would expect, rather than *merely* symbols. For the word contains (set-theoretically) the referent, so a proposition involving the word directly involves the referent.

Nonetheless, the sequence view is flawed. First, the view inherits many of the problems that face the theory of thin words. For example, it does not allow for many instances of mispronouncing or misspelling words. This is so since the orthotype/phonotype is a necessary constituent of the word. To produce a different orthotype/phonotype is to produce a different word. However, when Tom Brokaw says 'one *mi-we-un* dollars' he is mispronouncing the word "million." On the sequence view, he must be using an entirely different word.[133] For on this view, the misspeller has not misspelled a word, but has rather substituted a completely different

[133]Even supposing that the thin word or sound [million] is vague, it is certainly possible to mispronounce 'million' so egregiously as to fall outside the boundaries of the vague object [million]. The word 'nuclear,' for example, is commonly mangled in such a way ([Noo-kyew-lar]) that it cannot plausibly be said to be an instance of any putatively vague phonotype [nuclear]. One does not have to search far afield, unfortunately, for even more egregious examples, especially where the *spelling* or orthography of a thin word is concerned. A doctor's bad handwriting may cause a given inscription to fall outside the boundaries of any vague orthographic object which is a constituent of the set the sequence view identifies with a certain word, yet that inscription may still be a badly written instance of that word. Even worse, an inscription or sound can be so badly mangled as to actually fall inside the boundaries of a completely different phonotype!

word (sequence). And this is implausible. (See the section Misspelling, Mispronunciation and Malapropism in Ch 7 of this work for further details of this argument)

Another problem with the sequence view is that it raises issues similar to those facing the *ante rem* universals view. For sets exist iff their elements exist. And if this is so, then the theory holds that many words existed long before they intuitively did. For the corresponding <orthotype, referent> sequence will exist long before we think the words existed. For example, consider the verb 'frag,' which refers to the killing of one's own (unpopular) military officers by their soldiers in battle. This word was, to the best of my knowledge, coined in the Vietnam war, i.e, the 1960s. Of course, the practice of fragging has been around at least since the Napoleonic wars, and probably since the dawn of time. So even if the act of fragging or the property of fragging is an *in re* universal, it is a universal that has existed since well before the 1960s. And of course, the phonotype and orthotype |frag| have surely existed since well before the 1960s; it is nearly certain that those sounds and marks on paper were made (perhaps as the result of some accident!) long before this century[134]. If this is so, the sequence <|frag|, fragging> has existed long before this century. But then, the word 'Frag' existed long before the Vietnam war, and that American soldier who coined the term

[134]I am supposing here that entities like orthotypes and phonotypes are something like in re universals, in being some kind of temporally bounded abstract object. If orthotypes are *ante rem* universals, then the problem for the sequence theorist is even worse.

in the heat of battle did not create it, he simply discovered it. But this is implausible, as we discussed in the previous section on words as *ante rem* universals.

More problematically, the sequence view cannot account for true homonyms *that happen to have the same referent.* Names are a particularly obvious example of this. Since it is possible for an entity to have multiple names, it is surely possible for an entity to have multiple names which happen (as a matter of circumstance) to be spelled and spoken the same way. Certainly some cases which initially appear to be cases of distinct homonymous names turn out to really be cases of a single name[135]. Nonetheless, it is implausible to suppose that all cases are like this. Consider again our earlier Rock Hudson example. Suppose Rock Hudson participates in a new age ceremony where the participants are given new "back to nature names." In this ceremony, he is dubbed 'Rock' (his colleagues are dubbed 'Leaf,' 'Stream,' etc). It seems counterintuitive to suppose that the first half of the name under which Mr. Hudson was credited with starring in Magnificent Obsession is the same name as his "new age" name. But on the sequence theory, they must be. Moreover, the substitutivity puzzles arise for the sequence theory in the context of these homonymous coreferential names (coreferential names that are spelled/pronounced the same way, and yet are distinct). For no matter what one thinks the logical form of

[135]For example, in Kripke's "Paderewski" case, an individual is introduced to a man named 'Paderewski' on two different occasions, and although he believes these to be different and homonymous names, it is really the same name used in both cases.

the belief attributions are, if they involve these sequences, they will be unsurprising and trivial. For example, while it seems intuitive that Mr. Hudson's new age colleagues might be genuinely surprised to learn that Rock is Rock. it is trivially true that

the man denoted by <|Rock|, Rock> is the man denoted by <|Rock|,Rock>[136]

Finally, as with the *ante rem* universals view, many questions about the nature of words are simply unanswered on the sequence view. In short, the sequence view of thick words not only seems to conflict with, or remains silent about, folk-linguistic intuitions, but it also seems to do little to ameliorate the substitutivity problems like the ones we discussed in part I of this work. I conclude that it is an inadequate theory of thick words.

Before we develop the causal-historical view of words, it may be instructive to look in passing at a theory with similarities to both the causal-historical view and the sequence view. This is Mark Richard's theory of Russellian Annotated Matrices.

Discursus: Rusellian Annotated Matrices

A view closely related to the sequence view is Mark Richard's theory of

[136]One might, of course, simply accept that this proposition was trivially true, although it conversationally implied some proposition that was false; say, that the person with name associated with the movie star is identical to the person with the name associated with the new age camper . Nonetheless, we discussed such a strategy in the appendix to chapter 5, and found it lacking. No new reasons have been given to vindicate such a strategey.

"Russellian Annotated Matrices" (a name chosen, no doubt, for the acronym of 'RAMs,' which we shall henceforth use). In brief, Richard holds that the propositions attributed by belief attributions are RAMs, or sequences of entities such as individuals and properties (i.e., Russellian propositions), where each entity is 'annotated' with the appropriate word, much as in ILF theory (see ch 3, C). So, for example, an attribution such as

Sparky believes that Hesperus is Phosphorus

claims that Sparky is related by belief to the RAM

<<'is', Identity>, <'Hesperus', Hesperus> <'Phosphorus', Hesperus>>

Unlike a sequence-words theorist, however, Richard does not hold the claim that words are ordered pairs. Rather, he holds that 'Part of what makes for sameness of name in the requisite sense is being part of the same "causal chain" of transmission.'(Richard 1990, p 183) Finally, the words which annotate the RAMs reported by belief attributions are determined by a very sophisticated analysis. A belief attribution attributes not only a RAM, but also (tacitly) a *correlation restriction* to an agent. The correlation restriction determines which words of the believer's RAM will correlate with the word used in the attribution. Thus, the attribution is true just in case the agent believes some RAM containing words that correlate to the word used by the attributer.

Richard's view has several defects. First, it is ultimately a sequence theory, and as such, heir to many of the problems of sequence theories. Second, Richard's view of

the linguistic entities involved in RAMs is incomplete. There are many sorts of causal chains that might be necessary to words, but - as I shall argue - the sorts of causal chains best suited for the task (i.e., consistent with our intuitions about the individuation of words) are not suitable parts of RAMs. Moreover, Richard's account of how such words might be used to solve the "Paderewski" puzzles is itself unsatisfying. Let me address these points in greater detail

First, the fact that RAMs are simply sequences or ordered sets of entities causes several problems. First, there are the so-called 'Benacerraf" problems. That is, there appears to be no principled means of determining just which sequence a proposition is. Is it <subject, predicate> or <predicate, subject>? Neither sequence could be a preferable to the other as an analysis, but of course, they cannot both be the correct analysis. Is the correct sequence <word, annotation>, or <annotation, word>? The same point arises. Second, suppose that the sequence is (as in the ILF theory) determined by the deep structure of the sentence used in the attribution. If that is so, then the theory suffers many of the same problems as the ILF theory. For it seems plausible that different, even causally connected languages, have different deep structures. So an attribution must not only involve a correlation restriction on the possible words involved in the attributed RAMs, but also a restriction on the possible structures of the RAMs themselves. Moreover, even if the Benacerraf problems were resolved, the general point remains: do we *really* believe sequences of <word,

referent> pairs? It just seems wildly implausible to imagine that such a set-theoretic construct is the object of belief.

Second, as I claimed earlier, Richard's analysis of the words themselves is problematic for his overall theory of RAMs. Richard holds that a necessary condition of two tokens being tokens of the same name is that they are part of the same causal chain. But of course, there are many ways to delineate such a causal chain. A causal chain may be the actual history of the word. But this is a counterintuitive way to individuate words. Intuitively, a word might have had a different history (to a point). I might heard the word from a different individual. It might have been pronounced differently than it was. So the actual causal chain cannot be necessary to any given word. For reasons that I will argue in the next section, however, any such causal chain must be necessarily tied to the original referent of the word. But if Richard's causal-historical words have their referents necessarily, then RAMs simply contain redundant information. There is no point in annotating such a word with its semantic value. The semantic value is *already* necessary to the word.

Our earlier "Paderewski" puzzles also remain problematic. For these puzzles involve two tokens of a name that are part of the same causal chain (regardless of how one construes causal chains). Recall that the puzzle is that some agent may both affirm and deny that a person (say, Paderewski) has a certain property. Moreover, they may use one and the same name 'Paderewski' (or even "Paderewski") in these reports, because they mistakenly believe that there are two different names spelled and pronounced |Paderewski|. Thus, it seems as though Richard must hold that the

agent both believes and disbelieves some RAM correlating to

> <<‘Paderewski’, Paderewski>, <‘Being Musical, the property of being musical>>

(Constrained by a certain correlation restriction). In cases like these, Richard holds that the lexical constituents of the RAMS are not words, but rather word tokens; and the two RAMs actually involve distinct tokens ‘Paderewski’$_1$ and ‘Paderewski’$_2$.[137]

I think Richard is basically correct here. The Paderewski cases critically involve an agent’s mistakenly believing that two words aren’t really the same word. And I think this is best understood as the agent’s failing to realize that two distinct tokens - or some other finer-grained lexical items - are not tokens of the same word, rather than that a given word is not identical with itself. Still, I think Richard’s view is problematic. He holds that such attributions indicates that the agent believes some RAM correlating to

> <<‘Paderewski’, Paderewski>, <‘Being Musical, the property of being musical>>

where the correlation function restricts the possible RAMs to those containing tokens of Paderewski of the same *representational type* as the token involved in the attribution. And the notion of being of the same representational type is left uncharacteristically vague. But Richard does say: ‘Let’s call the representation the talented tokens determine ‘P_t,’ the representation the untalented tokens determine ‘P_u.’’(Richard 1990, p 185) . This certainly makes it look as if representational types

[137]It isn’t immediately clear if Richard thinks that *all* RAMs involve word tokens, or only these RAMs involved in the problematic “Paderewski” cases. The subtlety is not relevant for our purposes.

are individuated by some sort of descriptive knowledge (i.e. the type associated with being untalented, or the property of being unmusical, and the type associated with being talented, or musical). And in many cases, this simply cannot do the job, for reasons previously discussed. In brief, it seems possible to have a Paderewski-type puzzle, where the agent believes that the names or tokens ‘Paderewski’ and ‘Paderewski’ are distinct, but he has *no* descriptions associated with either name that individuate the tokens. I might, for example, falsely believe that both Anderson and Anderson are distinct philosophers of language (both named ‘Anderson’). If all I believe about either Anderson is that he is a philosopher of language, then every token of ‘Anderson’ will be of the same representational type.

Still, as I indicated, I think that Richard’s position is basically correct. However, the view of words as causal-historical entities needs further development, and I think that in the course of this development, I will be able to offer a similar (but preferable) solution to the “Paderewski” puzzles. And I think that the account of propositions given in chapter 5 of this work obviates the need for entities as baroque and cumbersome (not to mention as implausible) as RAMs. I turn now to developing such a causal - historical view.

C) The causal-historical view (Kaplan)

David Kaplan, in ‘Words,’ pioneered the view that words were to be identified at least in part by their histories. On his view, a word is a causal - historical entity, which begins with a particular introduction. I think that something very like this is

correct.

What sort of an entity is this causal historical entity? I propose that this entity is a member of a class of very common, yet largely undiscussed objects. I call these created in re universals. These are universal-like entities which, instead of being timeless, are created. Such entities are not as unfamiliar as they might at first appear. For there are other branches of philosophy in which such abstract, culturally created (or otherwise in re) objects are familiar, even if they have not been explicitly theorized about. For example, consider a familiar and analogous problem from aesthetics: what is a piece of music? We have many intuitions leading us to believe a piece of music is an abstract object of some sort; it is certainly repeatable. We also have intuitions leading us to believe that music is created, not discovered; and hence that a piece of music is not a necessary existant. The difficulty of providing a reductive analysis of a piece of music that answers to these intuitions seems to be grounds for holding that a piece of music is a sui generis object just like the sort I am proposing - abstract and culturally created.

Consider also a problem from the philosophy of biology: what is a species (or the similar question, what is a lineage)? Again, we have many intuitions leading us to believe a species is an abstract object of some sort. Likewise, we have intuitions that a species does not exist until its first member exists. For example, we have intuitions that *homo sapiens* did not exist until a few millennia ago, that species evolve as their

members change[138], and so on. And just as in the case of words, the difficulty of providing a reductive analysis that answers to both these intuitions might lead us to think that a species is a sui generis object like a word - it is abstract, and while not created by its first member, somehow brought into being by or with its first member.

Similarly, we can consider common abstract objects like statements or beliefs. A belief is intuitively a sort of abstract particular, which does not exist until someone has that belief (as opposed to the shareable content of a belief).

In short, we should remember that although the theory that words are a new sort of sui generis abstract object may be without established theoretical precedents, it is certainly not without precedents from our pre-theoretical thinking in widely disparate fields like philosophy of biology and aesthetics. And just as the theory of these created in re universals objects may provide answers for the philosophy of language, they may also prove useful in those other areas.

On my account, a word is associated with a tree. But I do not mean a 'tree' in the mathematician's sense of the word (i.e. a set with a particular ordering) or the sense of a chart representing a lineage. Rather, I wish to distinguish trees from lineages. For a lineage has its parts necessarily. If I had had a different aunt, for example, I would have had a different lineage. But this is implausible for words, species and other created *in re* universals. The word "Lobster" might have been pronounced with a silent 's,' for example; the species of lobsters might have evolved to lack large, tasty

[138]The analogy to thick words here is obvious - words evolve as their pronunciations change.

tails, the fifty-seven Chevy might have gone out of production for a year in the sixties. This is much like the relationship between ordinary universals and their instances. Redness might not have caused bulls to rush, and pure whiteness might have been instanced by this very piece of paper you are reading (instead of a whiteness admixed with black marks).

Let us say then that thin words are *presentations* or instances of thick words[139]. When one utters a word, one presents a thick word via its thin presentation. In what follows, I intend to flesh out an abstract causal-historical account of words by providing answers to the following questions:

1) What is the relation of thick words to trees (and branches)?
2) When are two thin words presentations of the same thick word?
3) When are two thin words *directly* related?
4) When are two thin words *indirectly* related?
5) What is a *tree* in the aforementioned sense?
6) What is a *branch* of a tree?

On my account, words themselves are not trees. Rather, the *presentations* of thick words form patterns of tree-like, branching, causal chains. This tree-like structure will be given by providing criteria for presenting thick words, criteria which themselves rely on causal and intentional connections between the thin presentations. Such criteria will also tell us what it is for two thin words to present the same thick word, either indirectly or directly; and these conditions will give rise to trees and branches of presentations of the same thick word. Thus, the answers to questions 2-6 will

[139]We can view this as roughly similar to Peirce's type/token distinction. However, Peirce's distinction seems to holds between thin word types and thin word tokens, whereas I allow that both thin word type and tokens may be presentations of thick words.

provide the full answer to the original question: what is the relation of thick words to presentations and trees of presentations? Let us, then, begin with the question of directly and indirectly related presentations of the same word.

Directly related presentations of a thick word:
w_1 and w_2 are directly related presentations of the same word iff:
1) w_2 and w_1 are either tokens of the same (or relevantly similar)[140] thin word-type, or w_2 *represents*, via some convention, the thin word w_1, or
1a) w_1 and w_2 are involved in the same *coining* of a word (see the following chapter for a more thorough discussion of coining), or
1b) the utterer of w_2 intends to use a token of the same thin word type as w_1, where this intention is causally related to the original use of w_1, and the utterer of w_2 intends that w_2 is a presentation of the same word that w_1 is a presentation of.

Indirectly related presentations of a thick word:
w_2 and w_1 are indirectly related presentations of the same word if and only if::
2) w_1 is the *linguistic ancestor* is of w_2. (That is, w_1 stands to w_2 in the relation that is the *ancestral* of the relation of being directly related presentations of a thick word. This relation, the linguistic ancestor, is defined thus: w_1 is the *linguistic ancestor* of w_2 iff it is a member of every set S that contains w_2 and all the directly related presentations of members of S).[141]
Some points about indirectly related presentations:
First, note that the utterer of w_2 need have no intentions about w_1, and as long as he has intentions about some word w3 which is directly or indirectly related to some word that is causally and intentionally related to w_1.
Also, note that being indirectly related is not a symmetric relation, because being a linguistic ancestor is not a symmetric relation. That is, w_2 may be an indirectly related presentation of the same word that w_1 presents, without w_1 being an indirectly related presentation of the same word w_2 presents.

Branches of a word:
3) w_1 and w_2 are *on the same branch* of a thick word W if and only if w_1 and w_2 are indirectly related.
A *branch* of a word is all the words that are on the same branch. We will designate branches a and b of a thick word W as "W"$_a$ and "W"$_b$.

---

[140]"Relevantly similar" remains vague here. This reflects what I take to be an ontological vagueness in just how much a word can be mispronounced/misspelled before we are ready to say that it is not even the same word.

[141]For a thorough discussion of the ancestral of a relation, see Quine (1982)

Trees associated with a thick word:
4) Two branches "W"$_a$ and "W"$_b$ belong to the same tree if and only if "W"$_a$ and "W"$_b$ overlap - that is, they have a thin word presentation in common.

Presentations of the same thick word:
5) Two thin word tokens w_1 and w_2 are *presentations of the same thick word* if and only if they are on branches belonging to the same tree.

Thin Word Types:
6) A thin word type W_1 is a presentation of a thick word W at a time T in a language L iff the weighted majority of speakers of L, at time T, present the word W using tokens of W_1. (I.e., there is a linguistic convention, at the time, to inscribe or pronounce the word in a certain way).

Let us now discuss some consequences of these criteria.

i) Essentialism

The word itself is not identified with the tree; rather, it is the causal and intentional requirements for presenting a word which give rise to the tree - like structure of its presentations. This is a desirable feature. For intuitively, although we now pronounce "cat" with a short 'a,' we might have come to pronounce it with a long 'a.' But if we identified the word with its branches[142], then it is hard to see how to accommodate this intuition. For the word "cat" has no such branch where it is pronounced with a long 'a.' Thus if the word is identified with its branches, the possible word that has the long-'a' pronunciation branch is simply a different word. But our account avoids this, for the actual branches are not essential to the tree. Any presentation can be a branch of a word, as long as it is causally and intentionally related to the other presentations of the word.

[142]Say, by treating a word as sequence of ordered <time, presentation> pairs.

ii) Humpty-Dumptyism and related issues

In our previous discussion of Humpty-Dumptyism, is was clear that words cannot simply refer to what individual speakers intend them to[143]. Such a result is ensured by the clauses of the presentation criteria, which require certain sorts of intentions to accompany presentations of a word. Moreover, the word that is itself presented has its referent independently of these intentions. The intentions of the speaker do not determine the referent of the word, only which word a token is a presentation of.

There is a related issue here, however. That is, can one utter a thin word w_2 that is wildly different from w_1, and claim it as a presentation of the same word as w_1? I think the answer depends on the circumstances. If you utter 'plane' to refer to aerial vehicles, and I hear this and utter 'refrigerator,' even (unreasonably) intending this to be the same word, my best intentions have gone awry. Intuitively, the answer here is no: 'Plane' and 'refrigerator' are not tokens of the same word. The thin words fail to be relevantly similar, and fail clause 2b[144]. Moreover, if I utter the thin word |plane|, but I have no intentions that this is the same thin word as some previously used word |plane|, then I seem to have introduced a new (albeit homonymous) word. For if I have no such intentions, then I have no intention that individuals will recognize that my thin word is a token of the same thick word. And thus I can hardly have any reasonable intention of presenting the same thick word.

---

[143]Of course, one is free to coin *new* words that mean what one intends them to mean, but when one coins such a homonym without revealing that they have coined it, this is disingenous and impedes communication.

[144]Nonetheless, we have an intuition that words may change pronunciation over time. We shall return to this below in the section on *word evolution*.

On the other hand, the pronunciation of a word may surely evolve. Words can change their presentation quite radically over time. The transitivity clause ensures this is so, while ruling out sudden radical changes in pronunciation. Suppose, as before, Sparky presents the word “Lobster” by making the sound |lobster|. suppose that Speaker B slightly mis-hears Sparky, and utters some thin word which is at least similar to |lobster|, such as |lobsturr|. Then suppose speaker C hears this, and utters some thin word which is similar to the thin word speaker B used. Anyone who has ever participated in this parlour game knows that by the time the word is transmitted even a small number of times, its pronunciation may change quite radically, even to the extent of sounding something like |refrigerator|. Intuitively, the thin word at the end of such a transmission chain is nonetheless a presentation of the same word as was presented by the thin word token at the beginning of the chain. This need not mean, however, that wildly different a thin word types may count as presentations of a word in a language. This is covered by our account of what is required for two for thin word *types* to present a thick word, as I shall discuss presently.

Finally, notice that two thin words may be presentations of the same thick word even if they are used by speakers of wildly different languages. As long as those speakers are causally and intentionally connected in the right way, the pronunciations and spellings of the word may vary wildly. A Latin word, for example, may come to be spelled and pronounced in a different way by English speakers, as in our earlier example of the Latin “masticare” and English “masticate”; but it is still the same thick word, for utterances of these words are causally connected to each other in the

appropriate way. A French speaker may be similarly related to the original Latin speakers and utter the word “macher,” which is also the same word as the Latin “masticare.” Thus, in virtue of definitions (1)-(5), above, these three thin word tokens present the same thick word.

iii) Coining

It is quite simple, on this account, to coin a word. One can coin a word that no one else understands (but one has to at least intend that others understand it). But this does not mean that one has coined a word of any particular language. Others speakers of a language must actually accept this word in order for it to be a part of the language. And so, while the creation of words need not be a matter of cooperative conventions (although it surely involves such conventions in getting an audience to understand one’s intentions) the creation of a language is a matter of such convention (see the section on thin word *types*, below).

Intuitively, a word may only be coined once. If I introduce the word “lobster” and token it with the sound |Lobster|, and another individual (causally unconnected to me) also coins a word he pronounces |Lobster|, these are intuitively different coinings and different words. It is a mere accident that we used the same sound |Lobster| (much like the Rock Hudson case from previous chapters).

Nonetheless, statements like

> He might have coined the word “Lobster” earlier, or even
> He might have coined the word “Lobster” with a slightly different spelling, such as |Lobsturr|

are presumably true. And this is so despite the fact that the actual coiner and the possible coiner have no causal relations or intentional relations. I do not intend my thin word |lobster| to be the same as (for example) some merely possible thin word |lobsturr|, nor do I cause the merely possible Devon to utter the merely possible thin word |lobsturr|. Nonetheless, the merely possible thin word |lobsturr| is surely a token of the same word that |lobster| is a token of. What is intuitively required here is that the two thin words are related to the same *coining* of a word "Lobster," where the coining of a word is an event that may be identical across possible worlds. I will discuss such event-identity in the section on cross-modal identity, in the next chapter. For now, I simply note that definition (1a) ensures that two thin words in different worlds may be tokens of the same word if they are used to coin the same thick word, regardless of whether they are causally or intentionally related. And the account of indirectly related presentations, branches and trees (definitions (2)-(5)) ensures that two thin words in different possible worlds may be tokens of the same thick word if they are each on the same branch containing the some presentation involved in the coining of that word in different possible worlds.

iv) Spoken and written words

Sometimes, agents create symbol systems where the symbols express their referents mediately - that is, the new symbols are intended to represent some other symbols. Phonetic alphabets are a good example of this. The symbols of a phonetic

alphabet (letters and spellings of words) are thin words that represent sounds (phonotypes)[145]. Insofar as they represent some thin word, they present the same word the phonotype presents. Thus, the phoneme |lobster| and the written symbols |lobster| may be presentations of the same word (if the rest of the above conditions are met). Similar examples include American Sign Language (ASL). If an agent cannot hear, then he needs a communicative medium to replace hearing spoken words; ASL thus contains symbols which represent spoken words. On this account, signs in ASL and in Spoken or written English may be presentations of the same word. The idioms of ASL use reflect this: we say, 'so-and so is signing [the word] "cat" '

Borderline and problem cases arise. Consider Chinese calligraphy, or Egyptian hieroglyphs. Are these orthographic systems intended to represent spoken words? This seems unlikely. Thus, on this account, spoken and written Chinese are (probably) different languages, containing distinct words. But this, I think, reflects our intuitions. Suppose a society arose that used two wholly different symbol systems, in different communicative media (i.e. sounds and written presentations). Suppose then that, at least at the outset, the written symbols are not introduced with any cooperative intentions (i.e., intentions to use the same word as some spoken presentation). Nor are they intended to represent any sounds (and in fact, Heiroglyphs and Chinese symbols are not phonetic in that sense). I think our intuitions in this case

[145]Of course, it is possible that the order is reversed, and spoken sounds are intended to represent written words. Given the usual order of linguistic development, i.e. spoken languages being developed before written ones, this would seem to be rare.

are clear: the spoken and written symbols simply present different words. Now suppose that over time, speakers of the language(s) come to be so familiar with each set of symbols that they translate freely between the different types of symbols. How does this, intuitively, differ from a strictly bilingual population? It does not.

v) Branches

Presenting a word gives rise to a certain tree-like structure; that is, a set of thin word tokens connected by causes and intentions. We might be interested in many of the parts of this tree-like structure, such as the parts corresponding to different languages. What I am here interested in are the parts of the tree that are connected chains of beliefs and intentions - I call these *branches*. A branch is, roughly, a presentation, together with the 'parent' of that presentation (that is, the presentation directly related to my presentation via cause and intention), the 'parent' of that presentation, and so on, back to the original presentation used to coin the word.

A word thus originally branches when a single speaker presents a word twice without intending to that each presentation to present the same word, or without intending to use a token of the same or relevantly similar word-type. Those two presentations fail clause (1B), so they are not directly related presentations; and they fail clause (2), since neither will be the ancestor of each other, although they presumably have some common ancestor.

Suppose speaker A presents (by uttering) some thick word such as "Paderewski" twice. Let us call these presentations p_1 and p_2. Suppose he intends the second

utterance, p_2, to be a presentation of the same thick word as the first. Both utterances of the word are on the same branch. But then suppose speaker B hears A's utterances, and mistakenly believes p_1 and p_2 to refer to different things, and hence mistakenly believes p_1 and p_2 to be presentations of different words .

B then utters "Paderewski" twice (p_3 and p_4). Nonetheless, in each case, he intends to present the same words that A presented - he intends p_3 to present the same word p_1 presents, and he intends p_4 to present the same word p_2 presents. Thus, p_3 and p_4, by the presentation clauses, are presentations of the same word that p_1 and p_2 present. Moreover, since p1 is the linguistic ancestor of p_3, p_1 and p_3 are on the same branch. And since p_2 is the linguistic ancestor of p_4, p_2 and p_4 are on the same branch. p_2, however, is *not* on the same branch as p_4, because it is not a linguistic ancestor of p_4. This is due to the fact that the *linguistic ancestor of* relation is not symmetric, and hence the *indirectly related presentation of* relation is neither symmetric, nor transitive.

An agent may fail to realize that two qualitatively similar thin word tokens present the same word, as in the above case. And so, he may fail to intend that two qualitatively similar thin word tokens present the same thick word, and hence, these qualitatively similar thin word tokens may be on different branches. But of course, an agent may fail to realize that two qualitatively *different* thin words present the same word, as in the case of |catsup| and |ketchup|, or pairs of presentations in different languages. And so, the same reasoning may apply to qualitatively different thin word tokens, which may be on different branches of the same thick word.

Finally, note that a branch is not a thin word itself, but rather a set of thin words. A branch is necessarily a branch *of a word*, and as such has its referent necessarily. We shall make extensive use of the concept of a branch in the following chapter.

vi) Thin word types and thin word tokens

It is fairly easy, on this account, for a thin word *token* to present a word. One can mispronounce it fairly egregiously, or at least one can pronounce it very differently from the accepted pronunciation. And regional dialects may pronounce the same word very differently. But there is surely a sense in which the utterance |new -KEW- ler| is not a presentation of the word "nuclear" that educated English speakers use. On my account, this is an intuition, about not thin word *tokens*, but thin word *types.* A thin word type is simply a class of thin word tokens. The thin word *type* |new -KEW- ler|, that is, the set of all the thin word tokens |new -KEW- ler|, is not a presentation of the word "nuclear" that educated English speakers use. And according to the criteria for thin word types, above, the thin word type |new -KEW- ler| is not a presentation of the word "nuclear" we use, simply because the (weighted) majority of English speakers do not use tokens of the thin word |new -KEW- ler| (i.e., members of the type of that token) it to present the word "nuclear."

Obviously, this is highly amorphous. First, I say "weighted majority," because the usages of some (e.g., lexicographers or other experts) presumably count for more than the usages of others (e.g. undergraduates). What the relative weightings are must be extremely vague. And what constitutes a majority is obviously vague. But this

mirrors the vagueness in lingusitic drift. At some time in the unfortunate future, the best efforts of English teachers will be for naught, and the official pronunciation will be the painful |new -KEW- ler|. This has happened countless times in the language; consider 'Worcester,' 'colonel,' 'aluminum,' or 'jaguar,' all of which are now pronounced radically differently than they originally were (presumably because their pronunciation was rather difficult). Famously, there are also examples like the gradual substitution of 'who' for 'whom[146].' For a thin word *type* to be a presentation of a word is then simply a matter of the conventions of the speakers at the time, and it changes over time.

Let us turn finally to the application of the theory of thick words to some outstanding issues.

[146]Thanks to Mark Moffet and Liz Slokar for some of these examples.

Chapter 10

Three puzzles

In this chapter, I discuss three outstanding issues: The so-called ‘Paderewski’ cases, the problem of cross-linguistic belief attribution (and belief sharing), and the problem of cross-modal attribution. Each of these issues presents a problem for any theory of belief and attribution which incorporates linguistic content, whether that content is in the form of thick words, thin words, ordinary descriptive propositions or the propositions described in chapter 5. Nonetheless, I believe that the theory of thick words may respond to these problems with considerably more success than previous theories.

A) Cross-modal attribution and “coining”

The problem of cross-modal attribution may be further divided into two problems. First, there is the problem of cross-modal attributions such as ‘if Sparky had spoken Japanese, he would have still believed P’; and as a limiting case of this, ‘if Sparky hadn’t even had a language, he would have believed P’ (or alternatively, ‘Even nonlinguistic beings believe P’). Such cases seem to me to fall under the aegis of cross-linguistic attribution, and I would like to save the discussion of them until that section.

There are, however, cases of cross-modal attributions within what is intuitively

the same language. For example, 'If Bush had lost the election in 2004, this would not have affected Sparky's belief that Hesperus was Phosphorus' (i.e., 'in a close possible world, Sparky would still have believed that Hesperus was Phosphorus.' On our analysis, such an attribution is analyzed as

in w_2, Sparky believes $pred_d$(identity, "Hesperus", "Phosphorus")

The question is, how can the Sparky in w_2 present tokens of the same thick word we are presenting in w_1? And if he cannot present tokens of the same word we present, how can his beliefs involve the same thick word?[147]

Such cases necessarily boil down to a question of trans-world identity for words. And I believe that the issue further devolves into the issue of identity for *coinings* or initial uses/ creations of a word. For if we can establish identity conditions for such coinings, then the identity criterion discussed in the previous chapter will ensure that any token which is a token of the same word as the coining-in-w1 will also be a token of the same word as the identical coining-in-w2 (and moreover, all of the tokens in w2 that are tokens of that coining in w2).

A coining requires a tokening of a word. After all, if a word is not even tokened, then how could it be created? For a word is essentially a communicative and public entity. Let us say that two coinings are coinings of the same word, and that they are thus the same coining, if at least the following conditions are met:

---

[147]Of course, it is the thick words themselves, and not their tokens, which occur in Sparky's beliefs. Nonetheless, it must be the case that the *tokens* of these thick words in each world are tokens of the same word.

1) each token is relevantly similar in its thin characteristics (e.g., orthotype, phonotype, gesture type, etc).
2) each token is used to refer to the same thing (individual, property, event, etc).
3) each coining uses a relevantly similar description to fix the referent. (again, this is deliberately vague, for the same reasons. Moreover, the criterion uses similarity, and not identity. Surely, for example, the coining of the expression "relativity theory" in two different worlds may be the same coining, even if the theory is formulated slightly differently; and the coining of "Hesperus" in two different worlds may be the same coining, even where the reference-fixing description in W_1 is *the last star to fade in the morning* and in W_2 it is fixed with *the final sun to disappear in the AM*.)

Let me discuss some consequences of this criterion. First, consider the similarity-of-thin-characteristics clause. The coining of a word is much like an event; intuitively, if an individual had used a radically different sound or shape to coin a word, he would have coined a different word. If the first English speaker to refer to cats had instead used the sound |lobsters|, he would have coined a different word. But also intuitively, if he had used the sound |cutt|, or something similar to |katt|[148], he would have coined the same word. The criterion of "relevant similarity" is deliberately vague, and is intended to reflect the vagueness in our word-identification practices.

Next, consider the sameness-of-referent clause. This simply reflects the arguments of the preceding chapters - that a word has its referent necessarily. If some individual coined a word referring to lobsters by tokening the sound |katt|, it would intuitively not have been the word "cat" but a mere homonym.

---

[148]Of course, the individual who coined the thick word "cat" probably used a sound very different from |katt|. For simplicity's sake, however, I will simply assume in this example that he used |katt|.

Now consider the similarity-of-reference-fixing-description clause. This reflects the intuition that identical words should have been coined under roughly the same (relevant) circumstances. And I think that the relevant circumstances are the circumstances that lead the coiner to pick out the referent in the first place. That is, the descriptive features of the referent that the speaker was considering (consciously or not) when he thought of the referent.

Suppose I look up at the last star to disappear in the morning and say 'I will call this 'Hesperus'.' I have thereby coined the word 'Hesperus.' There are surely many circumstances under which I could have coined that same word. But suppose in some alternate possibility, I look up at the evening sky, and see the first star to appear, and say 'I will call this 'Hesperus.'' it seems no more than an accident that I have used the same sound |Hesperus|, and the circumstances of the coining seem very different. In short, it seems that these two words are mere homonyms.

Philosophers have always had a strong sense that names are associated with *some* sort of descriptive content. On this account, the description is not part of the content or referent of the word, but it nonetheless is part of the identity of the word. Once the word is coined, of course, speakers need not have any such descriptive content "in mind" in order to use that word, and this is reflected in the criteria given in the previous chapter. Under these criteria, one speaker may token the same word as another speaker just in case they stand in a certain kind of causal/ intentional relation, and there is no reference to any reference-fixing content the speaker has 'in mind.'

One serious issue remains. This problem is what I shall call the problem of

"modal ductility." For the above criteria allow two thin word tokens to be presentations of the same word if they are *relevantly similar*, and does not require that they have identical thin characteristics. But then, even in worlds otherwise identical, two vastly different thin word tokens - say, |lobster| and |refrigerator| - might be coinings of the same word. For there is a world in which there is a thin word token, say, |lobsteref|, which meets all the other criteria mentioned, and is thus a coining of the same word we coined with |lobster|. And if we continue this reasoning, it is easy to see that eventually, |refrigerator| will coin the same word that |lobsteref| coins, and hence it is a coining of the same word we coined with |lobster|. And this is implausible. (Note that this simply mirrors the process of linguistic 'drift' described earlier, where just as the pronunciation of a word may drift over time, it may drift over modal space. And the same issue holds for clause (3), requiring relevant similarity of reference-fixing description.

I am not sure at present how to respond to this. Here, I will simply sketch some possible responses.

First, one might note that the above is intended to be necessary, and not sufficient conditions. It is surely the case that there will be pairs of worlds w1 and w2 where w2 has more than one coining meeting the conditions above. But it is surely not the case that those coinings are the same event or the same word as each other. The same word cannot be created twice, after all. But if this is so, they cannot both be the same word as the word in w1. It will simply be the task for a further work to fully determine the sufficient conditions for identity of coinings.

Second, one might hold that one can have coinings of the same word in different possible worlds only in cases where the two worlds overlap with regard to the region where the word is coined. In worlds that do not overlap in this way, one could then talk about coinings of words that are very similar to our actual words. I find this position sub-optimal, however. If some tiny historical fact about ancient Rome had been different in w_2, then w_2 would not overlap the actual world at the region where the name "Cicero" was coined (suppose the reigning empress had eaten one less lark's tongue in aspic, the morning Cicero was baptised.) Thus the token "Cicero" in w_2 would not have been a coining of the same word we use in the actual world. (Merely similar). And my beliefs in w2 involving "Cicero" would be different from my actual beliefs. And this seems implausible.

Similarly, one might "tighten" clauses 1 and 3, so as to require that coinings of the same word be phonetically/orthographically identical. This avoids the problem of modal ductility, by simply removing any ductility from the identity conditions. However, I find this response also counterintuitive. If we admit that words change their pronunciations over time, then why not also over modal space?

Finally, we should notice that this is not a unique problem for a theory of words, but rather, a very general problem for any entities which have (intuitively) ductile identity conditions. For example, it is fashionable to hold some ductile version of the 'necesssity of origins' thesis for certain physical objects; e.g., if this table had not been originally formed of at least 95% identical wood particles, it would have been a different table. But if we consider a succession of possible tables, it is easy to see that

the same issue will arise - table 1 might be identical to table 2, which is identical to table 3, and so on to table 10 - but tables 1 and 10 were not originally formed of *any* of the same wood, and hence are not identical. I will here leave the general problem of modal ductility for another day.

B) "Paderewski" cases

So-called "Paderewski" cases have proven to be problematic for several theories both of mental content, attitude ascription, and possible accounts of the ontology of words (specifically, the set-theoretic account). I believe, however, that the account of words and their structures given in the previous chapter will allow us to address the 'Paderewski' problems more elegantly.

Consider again the case of Sparky's beliefs about Paderewski. Sparky knows a single individual (the aforementioned Paderewski) under two different contexts. He knows Paderewski the symphony conductor, and Paderewski his shy upstairs neighbor. He also knows each person by the name 'Paderewski,' although he does not appear to know that they are the same person. So he appears to believe both that Paderewski is shy, and that Paderewski is not shy. [149] As Kripke put it, the problem appears to be that Sparky has contradictory beliefs that no amount of logical ability could resolve. On our account, those beliefs are

[149]And, as we discussed in ch 4, it is possible to create cases where there is no further non-metalinguistic information available to Sparky, by considering Sparky's beliefs about some individual only known to him under some historical names which he mistakenly believes to refer to distinct individuals, e.g, 'Cicero' and 'Cicero.'

$Pred_d$ [“Paderewski”, shyness] and not-$Pred_d$ [“Paderewski”, shyness].

Nonetheless, I think this is correct. Sparky really does believe some contradiction. But I think that where words are involved, such contradictory beliefs are common. Suppose, for example, Sparky’s friend lies to him about the names. He says,

> ‘Paderewski’ is not the same name as ‘Paderewski,’ or perhaps ‘Paderewski’ refers to Paderewski, but ‘Paderewski’ does not refer to Paderewski.

Sparky believes these falsehoods, and so it appears that he simply believes some contradiction. He believes that one and the same word both does and does not refer to his neighbor. And whatever account of words one has, I do not think this can be avoided. Moreover, if Sparky’s friend is really lying, then the names involved in his assertion are in fact identical. [Unlike other abstract objects like properties, words are opaque[150].]

Nonetheless, *some* different information appears to be conveyed by each use of ‘Paderewski.’ And this information is conveyed pragmatically, as in the accounts discussed in chapter 4. What differs is the sort of information conveyed.. I think in this case, it is information involving the branch that ‘Paderewski’ is on. Given that Sparky mistakenly intends each token of ‘Paderewski’ to refer to a different individual, then each token is on a different branch of the thick word “Paderewski.” And so while Sparky believes the contradiction

---

[150]Perhaps this is because properties are constitutive of or otherwise connected to possibilities, and possibilities are open to rational, a priori investigation; while words are not constitutive of such possibilities. Words are rather connected to conventions, and such conventions are only open to empirical investigation. But this speculation is an issue for further work.

$Pred_d$ [“Paderewski”, shyness], and
not $Pred_d$ [“Paderewski”, shyness].

He also believes

$Pred_d$ [“Paderewski”$_a$, shyness], and
not-$Pred_d$ [“Paderewski”$_b$, shyness].

Here, “Paderewski”$_a$ and “Paderewski”$_{b,}$ as before, represent distinct branches a and b of the thick word “Paderewski.” And although these propositions in this latter pair of beliefs cannot both be true, these are not contradictory beliefs about one and the same object. And I think this is really what is going on in the overtly linguistic contradictory beliefs as well. Of course Sparky believes that the word “Paderewski” is the same word as “Paderewski.” He simply fails to believe that a given token or utterance ‘Paderewski’ is an utterance of *that* thick word. He is still committed to contradictory beliefs, such as

‘Paderewski’ is a token of “Paderewski” and not a token of “Paderewski.”

But as I mentioned before, I think contradictory beliefs like these are simply unavoidable.

Notice, moreover, that this account neatly handles the “mixed-modal” examples from the Chapter 4 appendix. Suppose Sparky asserts that

I just learned that Paderewski is Paderewski - and since I believe that identities are necessary, I just learned that necessarily, Paderewski is Paderewski!

On our account, what Sparky learned was something like

(Necessity, [$Pred_d$ [Identity, “Paderewski”$_a$, “Paderewski”$_b$]]).

Unlike thin words, branches have their referents necessarily. And thus the above

modal claim is a truth, and accurately captures what Sparky could have learned.

C) Cross-linguistic attribution

Many cases of cross-linguistic attribution are elegantly handled by our account. As we noted in the last chapter, one and the same word may be presented in different languages. Where this is this case, as with the Greek “mastikhon” and English “masticate,” it is a simple matter for Greek and English speakers to share beliefs about, and attribute beliefs to each other, such as the belief that all who masticate, chew.

The real problem arises with what I shall call “causally isolated languages” (or causally isolated words). These are cases where groups of language-users are causally isolated from each other, such that no words from either language are used or adopted by the other. (In the terminology of the last chapter: no words of the one language are tokened or presented by the speakers of the other; that is, no speakers of the second language intend to use the same words as the first, no tokens of the second language are caused in the appropriate way by tokens of the first, and so on). And of course, even in languages which are as closely connected as, say, Italian and Latin, there may be cases where a word used in one language simply is not used in the other. After all, languages constantly add words. Moreover, two languages may independently coin a word with the same referent, even using the same reference-fixing procedure or description, and these words are intuitively distinct (just as, say, the English words “Hesperus” and “Phosphorus” are distinct). I shall refer to these as causally isolated

words.

The problem is this: how may one correctly attribute fine-grained beliefs to speakers of causally isolated languages? How may an English speaker, for example, correctly attribute the belief that Hesperus is not Phosphorus, to some Japanese speaker? (I will here presume that Japanese and English are paradigm cases of causally isolated languages. In actuality, there is some interpenetration at this point; but we can ignore this for our purposes, or, if you please, imagine pre-Shogunate Japanese). For neither of the thick words "Hesperus" and "Phosphorus" are elements of the Japanese language, and few if any Japanese speakers have beliefs involving these thick words.

The first thing to say about these cases is to remind the reader that the theory of thick words is not necessarily intended to solve *all* the substitutivity puzzles. Rather, the account is tailored to the many cases where our only real "access" to the things we hold beliefs about is via words for these things that have been passed down to us by other speakers. Cases like the Cicero/Tully cases are surely like this for most Americans, who really have no other knowledge of the Roman orator apart from the knowledge of his name. And I think the account of propositions involving thick words given here neatly and elegantly accounts for these cases. Still, I do not rule out the possibility that there are other non-standard or "non-Platonic"[151] "modes of presentation" of objects; some perhaps created with conventions or intentions, some

[151]See Bealer, [93] for both the introduction of the term 'non-platonic modes of presentation' (NPMOPs) and for possible further examples of NPMOPs.

mind-dependant, others not. I think that such NPMOPs will probably have many of the features of thick words - being *in re* abstract objects, having their referents necessarily, and so on. Moreover, I think the propositions involving such NPMOPs will have roughly the structure already discussed (cf. Chapter 5 of this work).

Nonetheless, I think I can here sketch at least the outline of an approach to the problem of cross-linguistic attribution for causally isolated languages. This approach will involve only the thick words and propositional structures I have argued for in this work.

The natural response is to suppose that our attributions somehow involve not the English words, but the corresponding translations. But this response has been historically problematic.

First, it raises the issue of the Langford-Church translation test again. If my attribution in the above case involves the Japanese word that translates as "Hesperus," then the issue simply arises once more when we ask how attributions by speakers of causally isolated languages can mean the same thing, since (for example) a Basque speaker making an attribution of a Japanese speaker will surely make no attribution involving "Hesperus."

Second, there is the issue of translation criterion. What is the correct Japanese translation of "Hesperus"? Sameness of referent is not sufficient for translation in the sense we require; for the Japanese, who had conducted some astronomy, presumably were in a similar situation to the ancient Greeks - at one point they falsely believed two heavenly bodies to be distinct, and hence gave them distinct names. But these

names both referred to Venus. Further, they had the same referents as “Hesperus” and “Phosphorus.” Yet these Japanese terms intuitively do not translate as *both* “Hesperus” and “Phosphorus.” Even if they did, we could not use translations to help us solve the issue of cross-linguistic attribution, for our attributions such as

Hiro believes that Hesperus is bright

would attribute beliefs to Hiro involving *all* of his expressions for Venus!

Finally, even supposing that we can resolve the issue of translation, I think that even translation is too coarse-grained. Consider the following example - the “Tale of two cities” example.

Shingen, in Hokkaidu, looks up at the sky in the morning, notices that a certain star seems to always be the last one to fade as the sunlight comes, and decides, ‘I will call this star “A”.’ His fellow Hokkaiduans pick this word up and use it. Meanwhile, Hiro, in Kyoto, has a nearly identical experience, and, noticing that a certain star seems to always be the last one to fade, decides, ‘I will call this star “B”.’ And his fellow Kyotans follow suit. Later, each discovers that the star they were thinking of is really the star Venus, which they know as “bi-nasu.” Shingen and his fellow Hokkaiduans come to realize that this star is the same one as the star they call “C,” the one which first appears as the evening sunlight begins to fade. Now suppose the Hokkaiduans and Kyotans intermingle, sharing their astronomical terminology. But the word is generally introduced in the following way: the Hokkaiduans say ‘Oh, A -

it's just good old bi-nasu,.'[152] However, Hiro, and his fellow Kyotans, who also use the name "C" for this star, do not seem to have realized that the stars B and C are the same; they will say (with the appropriate Japanese words), 'B is not C.'

The point of this example is that both 'A' and 'B' seem to be accurately translated as 'Hesperus' (Although Hiro and his Hokkaiduans do not realize this!), and 'C' as 'Phosphorus.' Nonetheless, although it seems plausible that we might correctly attribute to Shingen something like

> He believes that Venus is Hesperus

based on the fact that Shingen believes a proposition involving Japanese words that translate as "Venus" and "Hesperus," it seems unclear just what beliefs we should attribute to Hiro in this way. For Hiro will reject the claim that B is C, while accepting the claim that A is C. He then believes one proposition involving words that translate as "Hesperus" and "Phosphorus," and disbelieves a nearly identical proposition involving (different) words that translate as "Hesperus" and "Phosphorus."

Now consider Sparky, a native English speaker. Sparky might make an attribution of Hiro, something like

> Many Japanese, including Hiro, believe that Venus is not Phosphorus, as their astronomical sciences are not very advanced.

Suppose we then tell Sparky all about the two cities and the facts surrounding Hiro's

---

[152]Moreover, it is simple to amend the example so that both "A" and "B" have identical pronunciations and spellings, while the inhabitants of Hokkaidu and Kyoto believe them to nonetheless be mere non-coreferential homonyms, as in the Rock Hudson case.

beliefs. Does Sparky really mean to attribute to Hiro a flat-out contradiction - that he believes V = P, and also that he does not believe V = P? Presumably not. But Hiro seems to have beliefs that correspond to these - i.e., he believes that bi-naisu is C, and he does not believe that bi-naisu is B. Faced with this, I think Sparky will agree that his attribution does not literally capture Hiro's beliefs. Moreover, he will realize that he really does not know what exactly Hiro believes, and the best he can do is roughly characterize some features of Hiro's beliefs[153].

Let us deal with these issues singly.

Word Correspondence

What we require for two words in isolated languages to *correspond* to each other is that they have been coined using relevantly similar reference-fixing procedures or descriptions. It is not necessary for subsequent speakers of the language to be aware of these procedures or descriptions; for example, few English speaker know that "Hesperus" was initially coined while the speaker was thinking about the last star to disappear in the evening. Moreover, even if few Japanese speakers realize anymore that "A" was coined while Shingen looked up at the morning sky, and C was coined while the speaker looked up at the evening sky, it seems intuitive that "A" corresponds to "Hesperus" and "C" to "Phosphorus." This mirrors our intuitions

[153]To put a finer point on it, suppose the words "Venus," "Hesperus" and "Phosphorus" are then introduced to Hiro. Suppose Hiro comes to assert (in Japanese) that Venus is Phosphorus. Presumably, Hiro learned something. So his belief that bi-naisu = C could not be a belief that Venus is Phosphorus.

about cross-modal identity of words: for coinings in different worlds to be coinings of the same word, they need to have been at least coined in relevantly similar ways. Similarly, although these Japanese and English words are not identical, and cannot be coinings of the same word, the words in one language that correspond to words in another are those that were coined in relevantly similar ways.

Multiple Correspondence

As the "Tale of two cities" example shows, one language may have multiple words that correspond equally well to a single word of another language. But more importantly, I think we can redescribe the words "A" and "B" introduced in Kyoto and Hokkaidu in such a way that under just about any plausible criteria of "correspondence" - whether sameness of coining, or translatability, or sameness of connotation - "A" and "B" both correspond to "Hesperus." What can we (English speakers) say about the beliefs of the residents of Hokkaidu? We can say that they have some identity belief involving words that correspond to "Hesperus" and "Phosphorus." And moreover, they have some non-identity belief involving the words that correspond to "Hesperus" and "Phosphorus." But as attributors, we cannot do better than this. We simply are not in a position to say exactly what it is that Hiro believes. And our attributions, taken literally, will be ultimately false. Moreover, while Hiro and I share many beliefs about Hesperus and Phosphorus, the belief I express (whether pragmatically or semantically) by 'Hesperus is not Phosphorus' is not a belief we share. But this seems correct. If we think that we make different

attributions by

‘Venus is Phosphorus’ and ‘Venus is Hesperus,’

then we should make different attributions by

‘Venus is Phosphorus’ and ‘Venus is B,’

and thus different attributions by

‘Venus is Phosphorus’ and ‘A is B’

But if ‘Venus is Phosphorus’ expresses a different belief from ‘A is B,’ then the best we can say is not that Hiro and I share beliefs, but we have some beliefs that *correspond* to each other in that those beliefs involve modes of presentation - words - that correspond to each other. Thus, an attribution such as

A believes that P(a)

will, in these causally isolated cross-linguistic cases, be strictly and literally false where considerations of fine-grainedness lead us to analyze them as

A believes that pred_d (P, “a”)

Nonetheless, as we noted in our discussion of Soames in chapter 4, speakers generally convey a host of information with their attributions, and if a speaker has reason to believe that he is making an attribution of a person speaking another language, he probably also intends to convey the information that this agent believes some proposition like

there is a word w such that Hiro believes pred_d (P, w), and w corresponds to a.

To paraphrase Mark Richard, this is a theory of words, not of non-literal or

conversationally conveyed/implied meaning. As such, I do not really intend to discuss the details of how such content is conveyed. But it is worth noting that this is not the "semantic content" of the utterance. When, for example, Japanese and English speakers make attributions of (say) Basque speakers, they do convey slightly different information, in virtue of the very words they use, for part of the information used is information about particular words. And moreover, the literal meanings of their attributions are in fact different. The literal meaning or semantic content of what a Japanese speaker attributes to a Basque speaker is in fact different in many of these cases from the semantic content of what an English speaker attributes to him. So what are we to say about the Langford-Church translation test? Simply this: That two intertranslatable attributions in different languages need not have the same semantic content. It seems to me that two sentences may, in some cases, be appropriate translations if they have corresponding content (in the sense of "corresponding" described above), that is, if their contents are identical save for the substitution of corresponding thick words. This should not surprise us, for translation seems to involve both something more and something less than sameness of content. It is more than sameness of content, for a sentence such as

'The ides of March are here, so we should worry'

will not translate into a Japanese sentence involving the Japanese expression for March 15. The correct translation will involve something more than the Japanese expression with the same meaning; it will require an expression with a similar

connotation. And the correct translation will thus require *less* than sameness of content, for the it may be a fact about Japanese that the day when fates and dooms are decided is a not March 15 - perhaps it is April 4 (the fourth day of the fourth month, four being a particularly fateful and doom-laden number for the Japanese). Moreover, translations involving direct discourse often do not have the same literal meaning. Suppose, for example, a judge asks a court translator about a dispute between two Japanese-speaking individuals. 'Tell me what Shingen believes were Hiro's exact words in the contract,' asks the judge. Obviously, the translator will (correctly) offer an attribution like

Shingen believed that Hiro said [some English sentence],

and not an attribution involving the exact Japanese sentence Shingen thought Hiro used. So there is no reason why sentences of Japanese and sentences of English might not be good translations of each other, despite the fact that they have contents involving distinct (albeit corresponding) thick words. Translation does not require sameness of meaning, but rather correspondence of meaning. Thus we deny the initial premise of the Langford-Church translation test - that two intertranslatable attributions always have the same semantic content.

D) Conclusion: comparison with the other theories

I believe that at this point, it will be instructive to briefly compare the theory here with the previous theories I have criticized.

i) ILFs vs. Thick words

The theory of thick words bears some superficial similarities to Larsen & Ludlow's theory of ILFs. Both theories hold that the objects of belief involve linguistic items, without being about the relation of words to their referents. But the similarities end here.

First, ILFs hold that the objects of belief are deep structure syntactic trees (annotated with the semantic values of the words involved). But this is surely implausible for several reasons. It seems dubious that you or I ever believes a syntactic tree. And since the structure of the tree is given by the deep structure of the sentence used to express the belief, the possibility exists that the deep structures of two languages is completely distinct. In this case, it would be impossible for speakers of these languages to share beliefs, even in ordinary, otherwise unobjectionable cases of beliefs that involve no issues of fine-grained content[154].

Second, as long as the theory of ILFs gives us no theory of the linguistic items involved in the ILFs - in short, as long as it gives us no theory of words - we have absolutely no solution to the Paderewski problem, and no explanation of how beliefs may be shared even in causally connected languages (such as our earlier mastikhon-masticate case). To solve these, I think the theory of ILFs would simply have to import a theory of thick words.

---

[154]It is, of course, possible that necessarily, all languages share deep structure. But this hypothesis seems to have been met with considerable skepticism. At any rate, it seems unwise to have a theory of belief and attribution depend so crucially on such a speculation.

ii) Notions and three-place beliefs vs. thick words

The theory of thick words also bears some superficial resemblance to the account of belief presented by Crimmins in chapter 4. On both accounts, beliefs involve some kind of non-standard mode of presentation of an object (what we have called an NPMOP); on Crimmins' account, this is a concrete cognitive particular, on our account, it is an abstract, shareable linguistic item (a thick word).

Again, the similarities end there. Crimmins' theory has problems arising both from the form and the content of his beliefs and attributions. First, as Schiffer argues, Crimmins' propositions are baroque entities that seem to contain information most believers would deny having. Most believers might admit believing that Hesperus is not Phosphorus, but deny having any beliefs - even upon being pressured - about the properties of the various modes of presentation of Hesperus or Phosphorus. And most attributors would deny making any attributions about the properties of the modes of presentation involved in the beliefs they attribute.

Crimmins' modes of presentation - "notions" - are themselves problematic. For they are inherently non-shareable, and as such, no individuals can possibly have the same beliefs. Moreover, in attributing beliefs, agents cannot report the actual notions believed, but must instead report that a believer believes a notion with certain properties. Just as with Soames, these properties are either linguistic or descriptive, e.g., the agent believes a proposition involving a notion associated with the word w, or the agent believes a proposition involving a notion associated with a certain description. And just as with Soames, neither of these properties can accurately

characterize the attributee's beliefs; the linguistic desrciptions are either unavailable (as in the case of cross-linguistic attribution, even in the case of causally connected languages) or the descriptions themselves are too coarse to individuate the attributee's beliefs (as in the case of most attributions involving Cicero and Tully). And finally, the account cannot handle "Paderewski" cases. Perhaps many of these problems could be resolved with an appropriate theory of words; but then, I think that we are really better off integrating this theory of words into an account of propositions and beliefs like the account offered in chapter 5.

iii) Pragmatic implications vs thick words

Again, the theory presented by Soames in chapter 4 bears some similarities to a few of the solutions invoked in the thesis of thick words. Both theories hold that in some cases, the real information to be conveyed (or intended to be conveyed) by a belief attribution is not its literal referent or semantic content, but rather some further (presumably pragmatically conveyed) content.

The difference in the theories is pretty significant, however. The theories differ both in what they take the form or structure of this content to be, and in what sorts of things they take the content to involve. On Soames' account, the pragmatically conveyed information is descriptive information,; i.e., of the form

> the thing referred to by the word 'P' is Q, or
> the thing with the property P is Q.

Soames' account has problems arising both from the form and the constituents of

his conveyed information. First, in the case where Soames invokes metalinguistic descriptive information, the conveyed information has the wrong modal value, and thus the inferring problem arises all over again (see appendix to ch 4). And in the cases where Soames invokes ordinary descriptive information, not only do similar modal problems arise, but all the old problems of descriptivism arise - particularly, the problem that attributors rarely have enough descriptive information to pick out which propositions are believed (e.g., most people don't have sufficient non-linguistic descriptive information to distinguish Cicero from Tully, yet they seem to have or attribute beliefs involving Cicero without having or attributing these beliefs involving Tully).

Moreover, without a theory of (thick) words, Soames' account does not have the resources to account for either the Paderewski cases or attributions among causally connected languages. The natural solution to these issues would be to adopt the theory of thick words as presented here. But once one adopts the theory of thick words, I think that the account of propositions given in chapter 5 of this work provides a much more elegant and intuitive account than Soames' descriptivism.

All that remains of this work is a brief discussion of the methodology, and the reliance on "Folk Linguistics."

After Words

> "I don't really enjoy metaphilosophy. I find philosophy itself quite difficult enough, thank you very much." - Paul Hovda

I agree entirely with the above sentiment. Especially in the current case, I find metaphilosophical speculation to be somewhat unhelpful. Nonetheless, I suppose a few final words on the matter are in order.

In this work, I have relied extensively on "folk-linguistic" intuitions about words and "folk- psychological" intuitions about beliefs and belief reports. For example, I have taken it as a datum that individuals may share beliefs; that belief reports such as 'A believes P' and 'A believes Q' may be true (or at least convey some truth) even where P and Q are clauses that differ only in containing co-referring predicates or names. I have taken it as a datum that homonyms are different words, and that words are created. From these data, I have drawn some fairly substantial conclusions about the ontology of words and beliefs. One might, I suppose, ask if this was quite fair. What really is the status of my data, and what really is the status of these theoretical entities?

I think the issue about the status of the intuitions can be resolved rather simply. First, if the folk-psychological intuitions about beliefs and belief reports have no evidentiary status, then, well, I do not really see what could possibly guide us in the philosophy of mind. For presumably, there would be no intuitive evidence *against* any theory either, and it is difficult to see how science would or could weigh in on the

question without invoking some sorts of intuitions. If, on the other hand, only the folk-linguistic intuitions about words have no evidentiary status, well, then, they cannot provide evidence against the theory of thick words either. In that case, the theory of beliefs involving thick words offered in Part I of this work stands or falls on its own. It succeeds just in case it does the best job explaining our intuitions about belief, and it does not matter if it conflicts with our intuitions about words.

One might also allow folk-psychological and folk-linguistic intuitions some evidentiary status, but take a certain deflationary approach. That is, one might say that the theory given here simply tells us what we are ultimately committed to by our common practices involving words and belief attribution. But two points can be made here, without going too deeply into the nature of philosophy and philosophical commitments. First, these are some pretty surprising commitments, and I think it is worth knowing what one's practices commit one to. Second, one might respond: what is really gained by this deflation? In the absence of offering another, demonstrably better, set of practices, why should we create some theory, and then add the caveat that 'but it does not really describe reality, just the practices we have of describing reality'?

Obviously, this is not a real response to a deflationarily-minded critic, nor is it a real defense of the evidentiary status of intuitions from folk-linguistics or folk-psychology. But such has not been my task. I leave the appropriate metaphilosophical response to people smarter, more patient, and more profound than I.

Bibliography

Bach, Kent 'Sometimes a Great Notion: A Critical Notice of Mark Crimmins' Talk About Belief', available at the authors website at http://userwww.sfsu.edu/~kbach/Crimminsreview.htm

Bealer, George "A solution to Frege's puzzle", in Philosophical Perspectives, vol 7 (1993)

Bealer, George 'Propositions', *Mind* #107 (1998)

Bealer, George 'Modal Epistemology and the Rationalist Renaissance', in Conceivability and Possibility, Oxford University Press, (2002)

Bealer, George 'An Inconsistency in Direct Reference Theory', *Journal of Philosophy*, volume 101, #11, (2004)

Carnap, Rudolf Meaning and Necessity, 2nd Edition (Midway Reprint), University of Chicago Press, (1988)

Chalmers, David 'On Sense and Intension', Philosophical Perspectives 16 (2002), also available at the authors website: http://consc.net/papers.html

Chalmers, David 'The Foundations of Two-Dimensional Semantics', forthcoming in Philosophical Studies (abridged), also available at the authors website: http://consc.net/papers.html

Church, Alonzo 'On Carnap's Analysis of statements of Assertion and Belief', *Analysis* 10, pp 298- 304 (1950)

Cresswell, M.J. Structured Meanings: The Semantics of Propositional Attitudes, MIT Press, (1985)

Crimmins, Marc Talk About Belief, The MIT Press, (1992)

Evans, G 'The Causal Theory of Names', in The Philosophy of Language, A.P. Martinich (ed) Oxford University Press (1985)

Forbes, Graeme "Objectual Attitudes", *Linguistics and Philosophy* #23 (2000)

Fromkin, Victoria, and Rodman, Robert. Introduction to Language, 6th edition, Dryden Press, 1997

Grice, H.P. 'Logic and Conversation' in Studies in the Ways of Words, Harvard University Press (Reprint Edition) (1991)

Kaplan, David 'Words', *Proceedings of the Aristotelean Society Supplementary Volume LXIV*, (1990)

Katz, Jerrold Languages and Other Abstract Objects, Rowman & Littlefield Publishing, (1980)

Katz, Jerrold The Metaphysics of Meaning, M.I.T. Press, 1990

Katz, Jerrold Realistic Rationalism, The MIT Press (Reprint Edition) (2000)

King, Jeffrey 'Structured Propositions and Complex Predicates', *Nous* 29, 1995

Kripke, Saul Naming and Necessity, Harvard University Press, (1982)

Kripke, Saul 'A Puzzle about Belief', reprinted in Propositions and Attitudes, (eds.) Salmon & Soames, Oxford University Press, (1988).

Jackson, Frank. From Metaphysics to Ethics: A Defense of Conceptual Analysis. Oxford University Press, 1998

Jerrold Levinson, 'What a Musical Work Is', *Journal of Philosophy* #77, 1980.
Locke, John An Essay Concerning Human Understanding, Ed. P.H. Nidditch, Oxford University Press, (1979)

Ludlow, Peter and Larsen, Richard 'Interpreted Logical Forms' *Synthese* 95 (1993)

Ludlow, Peter and Larsen, Richard 'Interpreted Logical Forms, Belief Attribution, and the Dynamic Lexicon' in The Pragmatics of Propositional Attitudes, Also available at the author's website at http://www-personal.umich.edu/~ludlow/ilf.rtf

Lyons, John Linguistic Semantics: An Introduction, Cambridge University Press, 1995

Mates, Benson 'Synonymity', In L. Linsky (Ed.), Semantics and the Philosophy of Language. University of Illinois Press. (1952)

May, Robert, and Fiengo, Ralph "Interpreted Logical Forms: A Critique", Reprint from *Rivista di Linguistica*, #8 1996, also available at the author's website http://kleene.ss.uci.edu/%7Ermay/ilf.pdf

Moffet, Mark "Knowing Facts and Believing Propositions: a Solution to the Problem of Doxastic Shift", Philosophical Studies 115, pp 81-97 (2003)

O'Grady, William. Contemporary Linguistics: An Introduction, 5th edition, Bedford/St. Martin's, 2004

Peirce, Collected Papers, vol 4, eds. C. Hartshorne and P. Weiss, Harvard University Press, (1931-1958)

Quine, WVO Methods of Logic, Fourth Edition, Harvard University Press, (1982)

Richard, Marc Propositional Attitudes Cambridge University Press, (1990)

Saeed, John. Semantics, 2nd edition, Blackwell Publishers, 2003

Salmon, Nathan Frege's Puzzle Ridgeview Publishing, (1991)

Saul, Jenny 'The Pragmatics of Attitude Ascription' *Philosophical Studies* #92 (1998)

Saul, Jenny 'The Road to Hell: Intentions and Propositional Attitude Ascriptions', *Mind and language*, #14 (1999)

Schiffer, Stephen "Belief Ascription" *Journal of Philosophy* #89, (1992)

Schiffer, Stephen "The Mode-of-Presentation Problem", in Propositional Attitudes, eds In Anderson and Owens, Stanford: CSLI (1990).

Segal, Gabriel "A preference for sense and reference", *Journal of Philosophy* #86, (1989)

Soames, Scott Beyond Rigidity Oxford University Press, (2002)

Soames, Scott 'Direct Reference, Propositional Attitudes and Semantic Content', *Philosophical Topics* 15, 47-87 (1987)

Stalnaker, Robert "Assertion", *Syntax and Semantics* 9, (1978)

Stalnaker, Robert "Mental Content and Linguistic Form", *Philosophical Studies* #58

(1990)

Stalnaker, Robert “Semantics for Belief” *Phil. Topics* 15 (1987A)

Stalnaker, Robert Inquiry The MIT Press (Reprint Edition), (1987B)