

7-6-2017

Semantics, Meaning, Truth and Content: Disentangling Linguistic and Philosophical Approaches to the Study of Meaning

Toby Napoletano

University of Connecticut, tobynap@gmail.com

Follow this and additional works at: <https://opencommons.uconn.edu/dissertations>

Recommended Citation

Napoletano, Toby, "Semantics, Meaning, Truth and Content: Disentangling Linguistic and Philosophical Approaches to the Study of Meaning" (2017). *Doctoral Dissertations*. 1585.

<https://opencommons.uconn.edu/dissertations/1585>

Semantics, Meaning, Truth and Content: Disentangling Linguistic and Philosophical
Approaches to the Study of Meaning

Toby Napoletano, PhD

University of Connecticut, 2017

Abstract: The purpose of this dissertation is to clarify the relationship between two research programs engaged in the investigation of linguistic meaning. The first is the research program which we can think of as modern philosophy of language, broadly conceived. The second is semantics as it is pursued in contemporary linguistics, and in particular, semantics as it is pursued within generative linguistics.

It is often assumed that philosophers of language and semanticists in linguistics are working broadly within the same research program, addressing the same questions about language, its meaning, and its use. At least, it is assumed that the two research programs are continuous with one another, so that each places important constraints on the other. Philosophical theories of meaning—or of the meanings of some fragment of natural language—the thought goes, must square with the findings of our best linguistics, accommodating our best theories in syntax, semantics, language acquisition, and so on. Linguistic semantics, on the other hand, is typically taken to give “theories of meaning” for natural languages, or accounts of what our knowledge of meaning or “semantic competence” consists in. But just what sort of knowledge this is, and just what its object could be are arguably philosophical matters, and so philosophy of language plays a major role in determining the form that semantic theories must take if they are to be adequate qua semantic theories.

Consequently, notions that have been central to the philosophical study of meaning— notions of meaning, content, and truth—are also taken to play central roles in semantic theories. The dissertation argues that this view is mistaken. Semantic theories are not

Toby Napoletano – University of Connecticut, 2017

theories of meaning in any philosophically important sense, the semantic value of an expression does not even partly determine the content it is used to express, and facts about the truth-conditions and truth-conditional contributions of expressions do not play any explanatory role in truth-conditional semantics. The upshot is that the relations which notions of meaning, content and truth bear to linguistic semantics are more distant than is typically assumed, as are the relations between contemporary linguistics and the philosophy of language.

Semantics, Meaning, Truth and Content: Disentangling Linguistic and Philosophical
Approaches to the Study of Meaning

Toby Napoletano

B.A., University of Connecticut, Summa Cum Laude, **2007**

M.A., University of Connecticut, **2012**

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2017

Copyright by
Toby Napoletano

2017

ii

APPROVAL PAGE

Doctor of Philosophy Dissertation

Semantics, Meaning, Truth and Content: Disentangling Linguistic and Philosophical Approaches to the Study of Meaning

Presented by
Toby Napoletano, B.A., M.A.

Major Advisor _____

Lionel Shapiro

Associate Advisor _____

Michael P. Lynch

Associate Advisor _____

William G. Lycan

Acknowledgements

First, I must express my deep gratitude to Lionel Shapiro, whose generosity, encouragement, persistence, attention to detail, and insistence on clarity were all essential for the development and refinement of the ideas that appear in the dissertation. Thanks to Lionel, I have never suffered from a lack of critical feedback with which I needed to wrestle with, and the contribution this has made to the quality of my work cannot be overstated.

Gratitude is also due to Michael Lynch, William Lycan, Zoltán Szabó, and the members of the UConn LEM and Yale ELLMM City working groups for valuable feedback on earlier drafts of chapters.

Finally, I must thank the rest of the UConn Philosophy Department, its faculty and its graduate students for their support and their invaluable role in my development as a philosopher. Among the faculty—in addition to Lionel Shapiro and Michael Lynch—I am especially indebted to Paul Bloomfield, Don Baxter, Tom Bontly, and Crawford Elder. Among current and past graduate students, thanks to Benjamin Nelson, Jared Henderson, Alycia LaGuardia-Lobianco, Junyeol Kim, and especially Michael Hughes, Casey Johnson, Ross Vandegrift, and Hanna Gunn for their friendship, support, and continual enrichment of my intellectual life.

Introduction

The purpose of this dissertation is to clarify the relationship between two research programs engaged in the investigation of language and meaning. The first is the research program which we can think of as modern philosophy of language, broadly conceived. The second is contemporary linguistics, and in particular, the generative linguistics research program.

Within the philosophy of language, I include those theorists who study language roughly within the tradition that is typified by the work of figures like Frege, Wittgenstein, Russell, Carnap, Quine, Sellars, Davidson, Dummett, Montague, Lewis, Kaplan, to name just a few. There are, of course, distinct research programs to be recognized within the philosophy of language, but they are all engaged in answering at least some of the central questions of the philosophy of language. For instance:

“What is a language?”

“In virtue of what do we have our language?”

“How do we come to know the meanings of expressions in our language?”

“What sorts of things are meanings?”

“What does such-and-such (philosophically interesting) expression in such-and-so language mean?”

“In virtue of what does an expression mean what it does?”

“What are the desiderata that must be met by an adequate theory of meaning?”

“What is communication and how are we able to do it?”

There are also any number of questions which are, perhaps, conceptually less central to the philosophy of language than those on the above list, but which are nevertheless a focus of the philosophy of language, and which concern its relation to other areas of philosophy. Upon considering a theory of meaning, we might ask, for instance, what metaphysical, epistemological, or logical commitments we might incur (if any) were we to endorse it, or what other positions might be justified if that theory of meaning were true. One might think, for example, that endorsing a truth-conditional theory of meaning commits them to a realist metaphysics, or that endorsing a particular semantics for the English ‘knows’ will enable one to avoid skeptical paradoxes.

Central to the philosophy of language research program is the project called “semantics”. Semantics, to a first approximation, is that part of philosophy of language which is concerned with pairing the expressions of a language with their meanings.¹ The project is conceptually distinct from, but is obviously very tightly tied to questions about the nature of meaning, language, communication, and of the conditions of adequacy for a semantic theory. There is also a project called “semantics” which is central to contemporary linguistics. It, too, is conceived of as attempting to pair expressions with their meanings. Very often, linguistic semanticists employ the same formal tools and central concepts in their theories as do philosophical semanticists, and seem to offer similar sorts of explanations.

It is natural to wonder, then, whether there is any theoretically interesting difference between what I am calling ‘philosophical semantics’ and ‘linguistic semantics’. Perhaps the difference only reflects which building the semanticist does her work in. In that case, the findings of linguists would bear directly not only on the philosophy of language, but also on related subfields of philosophy. Likewise, findings in, e.g., metaphysics and

¹Even this formulation is philosophically loaded, since it suggests that there are entities that are meanings. Cf. Davidson (1967).

other subfields of philosophy would directly constrain linguistic theorizing.

In this dissertation, I argue that, despite appearances, there is considerable theoretical distance between linguistic semantics and philosophical semantics (and related philosophical subfields). In particular, I will argue that the relationships linguistic semantics bears to notions of meaning, content, and truth, are more distant than is typically assumed. I will not and can not defend any general thesis concerning the relationship between generative linguistics and the philosophy of language. The reason is that it is not clear there is currently anything to say about their relationship conceived at that level of generality that would be accurate and illuminating. Even if we hold fixed a particular linguistic research program, conceptions of philosophical pursuits are varied. Further, the relationship between the two research programs will depend on answers to linguistic and philosophical details that we have yet to work out.

My goal, instead, is to shed light on their relationship by examining a number of assumptions that are standardly made in philosophical and linguistic theorizing that link linguistic semantics tightly to the notions of meaning, content, and truth, which figure centrally in philosophical pursuits. I will argue, in each case, that these assumptions ought to be rejected, and that the relation between linguistic semantics and these philosophical notions is more distant than is standardly assumed.

I should also make clear my reasons for focusing on the generative linguistics research program as opposed to other linguistic frameworks. There are two main reasons for doing so. First, it is the linguistic framework that philosophers (and philosophers of language) most frequently reference when referencing empirical linguistic work.² The other reason, which explains philosophers' interest in the research program, is that it is a research program that has had considerable success, and is the most developed of

²For a quick sampling of recent philosophical work which references generative linguistics, see Horwich (2003), Collins (2005), Gross (2010), Ludlow (2011), Glanzberg (2014), Yalcin (2014), and Schiffer (2015).

the linguistic research programs. This might be less obvious in the case of semantics when compared to phonology and syntax, as semantics is still in a very early stage of development and semanticists are often less explicit about how their theories fit into linguistic inquiry more broadly. The most cited semantics textbook, however, Heim & Kratzer's *Semantics In Generative Grammar*, is meant to fit broadly within the generative linguistics framework.³ Further, even where it is not made explicit that a semantic theory is to be taken to fit within generative linguistics, it is typically taken to be directly continuous with syntax and phonology. But in those subfields of linguistics, the generative linguistics approach has become the “received view”.⁴ Thus, I think that it is charitable to interpret much of semantics as fitting broadly within generative linguistics, even when this is not made explicit. By focusing on generative linguistics, then, I intend to focus on the branch of linguistics that is typically taken to have the greatest bearing on philosophy, due in large part because of its prominence in linguistics.

The structure of the dissertation is as follows. In Chapter 1, I will give a sketch of the generative linguistics framework, and argue against the idea that meaning is the object of investigation of linguistic semantics. The idea that semantics, in this context, is a theory of meaning or a theory of speakers' knowledge of meaning, is nothing more than a useful but informal characterization of the research program. Notions of meaning and knowledge of meaning do not figure importantly in semantics, and so philosophical theories of meaning or knowledge of meaning do not obviously constrain linguistic semantics and vice versa.

Chapters 2 and 3 discuss the relation between linguistic semantics and the philosophical notion of communicative content. Chapter 2, focusing on a recent proposal by Schiffer, argues that semantic value—which figures centrally in semantic theories—

³See also Larson & Segal (1995).

⁴See Kenstowicz (1994, 1). The quote is a bit outdated, but applies just as well today.

should not be understood in terms of utterance content. I argue that if one understands semantic value in terms of content, the claims made by a semantic theory will be false, or else its explanations will suffer from problematic circularity. A broader lesson from this chapter is that our semantic competence should not be understood in terms of our communicative abilities.

Chapter 3 further explores the connection between semantic value and content. While it may be granted that semantic value is not identical to content, or that the former should not be understood in terms of the latter, it is a ubiquitous assumption in semantics that semantic value at least (partly) determines utterance content. While it is certainly true that our semantic competence is very important to our ability to communicate with one another, I will argue that no such determination relation obtains. I also sketch a view of the role that our semantic competence plays in communication which is consistent with the rejection of this determination principle.

In chapters 4 and 5, I consider the relationship between semantics in generative linguistics and truth-conditions, and hence truth. In Chapter 4, I argue for a particular conception of compositionality in the context of linguistic semantics. On the conception I argue for, the demand that semantic theories be compositional in no way forces one to truth-conditional semantics. In that case, a common a priori argument for thinking that semantics must be truth-conditional is undermined.

Even if there is no a priori argument that semantics must be truth-conditional one might give an a posteriori argument for this view on the grounds that truth-conditional semantics in generative linguistics has been the most successful approach to doing semantics. Chapter 5 argues that even if our best semantics is truth-conditional, it doesn't follow that truth and truth-conditions play a major role in semantics. The reason, I argue, is that reflection on our best work in truth-conditional semantics suggests that while truth-conditional theories give lots of good explanations of semantic

phenomena, truth-conditions do not actually do any explanatory work in the theories. Rather, the explanations appeal to properties of expressions other than their truth-conditions or truth-conditional contributions.

Contents

1	Meaning and Semantics in Generative Linguistics	1
1	Introduction	2
2	The Generative Linguistics Framework	3
2.1	Generative Linguistics in the Abstract	3
2.2	Semantics In Generative Linguistics	9
3	Conceptual Concerns for Generative Linguistics: Language, Knowledge, and Meaning	16
3.1	The Significance of Claims About What A Theory is About	17
3.2	Terminological Concerns	21
4	Do Semantic Theories in Generative Linguistics Constitute Theories of Meaning?	23
4.1	Semantics in Generative Linguistics and the Ordinary Concept MEANING	25
4.2	Semantics in Generative Linguistics and Conceptual Engineering	28
4.3	Conceptual Engineering and Concepts of Concepts	30
2	Why Semantic Value Should not be Understood In Terms of Utter- ance Content	34
1	Introduction	36

2	Schiffer’s Negative Arguments	38
2.1	The Insufficiency Argument	38
2.2	The Non-necessity Argument	41
3	Reasons Not To Bet On The Best Bet Theory	43
3.1	A Sketch of the BBT	44
3.2	Circularity In The BBT	51
4	Can The BBT Be Saved?	57
4.1	‘Literal meaning’ As ‘Compositionally derived meaning’	57
4.2	‘Literal meaning’ As ‘Explanatorily basic meaning’	58
4.3	What A Sentence Is “Standardly” Used To Mean	60
5	Conclusion	62
3	Why Semantic Value Does not Determine Content	65
1	Introduction	67
2	The Appeal of the Determination Principle	68
3	Background and Methodology	72
3.1	Semantics in the Generative Linguistics Framework	72
3.2	‘Semantic Value’	74
3.3	‘Expression’	79
4	‘Expressing content’	80
4.1	The Content Role, Broadly Speaking	81
4.2	Externalist Notions of Content-expression	82
4.3	What Is Expressed as Determined by Convention or Norm	85
4.4	Internalist Notions of Content	88
5	Determination	94
5.1	Metaphysical Determination	95

5.2	Epistemic Determination	97
5.3	Causal Determination	100
6	Does Semantic Value Determine Content?	104
6.1	The Problem of “Non-literal” Talk for (DET _{sys})	104
6.2	Problems of Anomaly, Ungrammaticality, and Micro-conventions	113
7	Objections and Replies	122
7.1	Restrictions to Utterances	122
7.2	Restricting the Principles to Utterances of Natural Language Sentences	131
7.3	Restricting the Principles to Normal Interpreters	134
7.4	Do The Arguments Undermine Pragmatics?	135
8	The Relationship Between Semantic Value and Content	136
8.1	Semantic Competence As A Perception-Like Ability	137
8.2	Mental Content and Utterance Content	144
4	Compositionality as Weak Supervenience: Why Compositionality Doesn’t Require Truth-Conditional Semantics	148
1	Introduction	150
2	Compositionality and Linguistic Explanations of Acquirability	152
3	Motivation For the Strong Supervenience Reading	158
4	(SS) and Linguistic Explanations of Acquirability	162
4.1	Compositionality _{SS} and Weak Systematicity _{PSY}	162
4.2	The Circularity of the C-M Approach	166
4.3	Why The Principle of Compositionality is Not (SS)	167
5	Compositionality as Weak Supervenience	172

6	Addendum: Compositionality and Non-Truth-Conditional Approaches to Semantics	180
5	How Important Are Truth-Conditions For Truth-Conditional Semantics?	185
1	Introduction	186
2	The Semantics of Gradable Adjectives and Degree Modifiers	192
2.1	The Basics of Truth-Conditional Semantics	192
2.2	Gradable Adjectives and Degree Modifiers	194
3	The Insufficiency Argument	197
3.1	Premise (i): Disquotational and Non-disquotational Analyses of Gradable Adjectives and Degree Modifiers	197
3.2	Premise (ii): Explanations of Semantic Anomaly and Entailment For Gradable Adjectives and Degree Modifiers	201
4	The Superfluity Argument	217
5	Generalizing the Arguments: Disquotation and Semantic Explanation .	230
6	Conclusions	239
	Bibliography	240

Chapter 1

Meaning and Semantics in Generative Linguistics

Chapter Overview: Part of what it is for a semantic theory to be *semantic* is for it to have to do with meaning as it is ordinarily understood. If a theory had nothing to do with our pre-theoretical conceptions of meaning, then it would be inappropriate to characterize it as being a semantic theory as opposed to some other sort of theory. Semantic theories in generative linguistics, I will argue, can be appropriately characterized as being semantic, and as being theories of meaning. However, the appropriate characterizations of semantics in generative linguistics as theories of meaning are appropriate only insofar as they are useful as informal characterizations. They should not be taken to have any more significance than that. And in particular, such characterizations should not be taken to have philosophical import. Consequently, one can endorse a semantic theory in generative linguistics without hereby incurring philosophical commitments about meaning.

This leaves open the question, however, as to whether we should take on the additional philosophical view that semantic theories in generative linguistics constitute

theories of meaning. Whether we should or not, I argue, depends on the relationship between semantics in generative linguistics and the ordinary concept of meaning. I then argue that given the nature of this relationship, we have reason to reject the idea that semantics in generative linguistics and philosophical theories of meaning are so tightly tied together. Whether this view is persuasive, however, will ultimately depend on a foundational metaphilosophical question concerning how we should individuate concepts.

1 Introduction

In this chapter, I want to discuss the relationship between semantics in generative linguistics and meaning. After all, it is not uncommon for such theories (and semantic theories generally) to be thought of as theories of meaning in some sense or other. It might be thought, therefore, that philosophical theories of meaning and semantics in generative linguistics are at least continuous with one another, and tightly constrain one another. The aim of this chapter is to argue against this view of the relationship between semantics in generative linguistics and philosophical theories of meaning.

My first task, then, will be to sketch the research program of generative linguistics, and then how semantics fits into it.¹ This is the task of §2. I will then argue in §3 that while it is informally useful to characterize semantics in generative linguistics as a theory of meaning, the characterization should not be taken to have any philosophical import concerning the relationship between semantics and meaning. Consequently, one can endorse a semantic theory in generative linguistics without thereby taking on any commitments about meaning. In §4, I give an argument for thinking that semantics does not constitute a theory of meaning in any philosophically committing sense.

¹I will be providing nothing more than a sketch of the generative linguistics framework here. For fuller presentations, see, e.g., Chomsky (1965, 1986, 2000) and Ludlow (2011).

Whether the argument should be convincing or not, I concede, depends on a foundational metaphilosophical issue—how one conceives of concepts in our metatheorizing. I conclude that, at least, we lack good reason to think that semantics in generative linguistics constitutes a theory of meaning.

2 The Generative Linguistics Framework

2.1 Generative Linguistics in the Abstract

I will first sketch the research program in its broadest outlines, and then narrow in on more informative characterizations. I will put off discussion of potential concerns with these characterizations until §3. I want to stress that the foregoing discussion of generative linguistics and the place of semantics in that broader framework is only a sketch. My aim in this section is simply to help familiarize the reader with the theoretical aims and methodology of generative linguistics, and to help orient generative linguistics in relation to other areas of inquiry. Further details about semantics and generative linguistics will become relevant in later chapters, and so I discuss them there.

Understood in its broadest and most abstract form, generative linguistics is a scientific research program which aims to investigate human natural language. It is not, therefore, concerned with language or communication systems in general, but only with those that can be adopted by (psychologically normal) humans. Further, ‘human language’, here, is not meant to include any system of communication that can be adopted by humans, but rather those languages that can be acquired as first languages by psychologically normal humans in the usual sorts of circumstances.² Consequently,

²‘Psychologically normal’ and ‘usual circumstances’ are, by necessity, vague. It is an open empirical question just what sorts of psychologically states would be abnormal in the relevant sense, and what kinds of conditions in early childhood must obtain in order for children to acquire their first language.

the languages of arithmetic, classical logic, and other formal and theoretical languages are not objects of study in generative linguistics. Unlike a child’s first language, these “artificial” languages are acquired under very different circumstances, typically later in life, and requiring explicit instruction and training. Now, just which languages are the natural languages—i.e. which can be learned by psychologically normal humans in normal circumstances—is not something we can know in advance of theorizing. This sort of situation is, however, the norm in scientific theorizing. We simply do not know, in advance of theorizing, just which phenomena will fall under the explanatory purview of a theory or research program. Our conception of the explananda of a theory becomes more fine-grained over the course of theorizing.

The extremely broad and abstract characterization of generative linguistics that I’ve given thus far does little to disambiguate it from distinct approaches to the study of language that might be pursued in linguistics, philosophy, or sociology, for instance. If one is to study human natural language, one has some important theoretical choices to make. How, for instance, should we conceive of a language? And what is it for a human to have some language as their language? Without answers to these questions, we can have little grasp of what kinds of evidence will bear on our theories and what sort of methodology would be appropriate in our investigation.

A bit less broadly, generative linguistics studies our “linguistic competence”, or our “knowledge of language”. This, for instance, is how Chomsky—the founder and one of the leading theorists in the research program—often characterizes the project. The characterization is reminiscent of how some philosophers of language have conceived of semantics—as a theory of our linguistic understanding or our “knowledge of meaning”.³ I will have more to say about this apparent overlap in later sections. But for now, we should just note that this characterization, while it does suggest a broadly psychological

³See, for instance, Davidson (1965), Dummett (1993a), and Lewis (1980).

sort of inquiry (in that it is concerned with the mental lives of speakers), still fails to satisfactorily pin down a coherent research program. What is this knowledge of language like? Is it explicit or implicit? Is it a skill or a kind of practical knowledge? What sorts of behaviors are explained by attributing this knowledge to speakers? And, again, what is a language, and in virtue of what does a speaker have knowledge of one language rather than another?

Here, if one wants to study linguistic competence, one must (once again) choose how this should be done. One might, as Lewis suggests, be interested in providing “a systematic restatement of our common knowledge about our practices of linguistic communication” (1980, 79). This knowledge, presumably, shapes our linguistic behavior, and results in community-wide patterns of linguistic behavior that call out for explanation. One might, for instance, be interested in the fact that English speakers know that, most often, if they utter to someone, “The dinosaurs are extinct”, then they will be taken to mean that the dinosaurs are extinct. They also know that, given their first utterance, it would be inappropriate to go on to utter “I have a pet dinosaur”, and that neither utterance is a good way to start a conversation with a stranger, and so on.

This sort of knowledge is essentially social, emphasizing the idea that language is a publicly shared tool that can be put to any number of ends. This is not, however, the way that linguistic competence is investigated in generative linguistics. In generative linguistics, one’s linguistic competence is not taken to consist in one’s knowledge (explicit or implicit) of a publicly shared language, nor does one’s having some language as their first language consist in their being a member of some linguistic community. Rather, one’s linguistic competence and one’s having some language both consist in a particular faculty of mind—the language faculty—being in a particular state. One’s linguistic competence, in this sense, will underlie one’s knowing that “Gilgamesh missed

often Herbology” is ungrammatical, or that “Mary is half tall” is anomalous, or that “The door is open but it’s closed” is a contradiction, and so on. This knowledge, it is supposed, is the result of one’s having knowledge of a complex of rules—phonological, syntactic, and semantic—and a large number of lexical items. These rules, together with the properties of the lexical items, determine (among other things) the syntactic structure of complex expressions (and so which complexes are grammatical), and the semantic “meaning” of the complex (and so which are meaningful at all).

All of this “knowledge” is taken to be “tacit knowledge”—a kind of unconscious knowledge that is not open to introspection but is discoverable by the methods employed by linguists in their theorizing. This body of tacit knowledge is what constitutes one’s linguistic competence in the relevant sense, and is the object of investigation of generative linguistics.⁴

One’s having this tacit knowledge, again, is simply a matter of one’s language faculty being in a particular state. The language faculty is taken to be a special faculty of mind possessed by humans which enables them to learn our first (spoken) language as we do—mostly in early childhood, with little instruction, and with limited exposure to linguistic stimuli.⁵ Typically, though not uncontroversially, the state of the language faculty at birth is taken to be such that it imposes constraints on the kinds of languages that can be learned by a child as a first language. Thus, it is supposed that there is certain, innate, tacit knowledge of language that guides a child’s acquisition of language, given their exposure to various linguistic stimuli.⁶

For the purposes of this chapter, issues concerning the innateness of language or the origins of the language faculty can be left aside. What is important, for current

⁴I will return to the issue of tacit knowledge and whether linguistic competence in the target sense should be characterized as a kind of knowledge in §3.

⁵See Guasti (2002) for a textbook presentation of language acquisition in generative linguistics.

⁶See Chomsky (1980, 1965) for arguments in support of the idea that the language faculty innately imposes certain on which languages can be learned as first languages. These constraints are taken to be what makes it possible for children to learn a language on the basis of limited linguistic stimuli.

purposes, is simply that generative linguistics studies a particular faculty of mind—plenty of room is left for disagreement about the details of that faculty. Generative linguistics is, thus, a branch of the cognitive sciences, on par with investigations into the faculties of mind that underlie our visual capacities, for instance.⁷ Theorizing within the major subfields of generative linguistics—phonology, syntax, semantics, and acquisition—is typically pursued at a very high level of abstraction, broadly within the framework of the computational theory of mind. Thus, while it is assumed that the language faculty is realized in the human brain, and that linguistics will be made continuous with neuroscience, linguistic theorizing is not couched within the vocabulary of neuroscience.⁸ The reason, in part, is that our knowledge of the brain and its organization is still quite limited, and we lack a firm grasp on how abstract structures of tacit knowledge are instantiated at the neural level. Linguistics and neuroscience do, however, constrain one another in principle.

If what it is to have one's language is for their language faculty to be in a particular state, how should we understand language itself? Languages, qua objects, are treated as abstract entities. These will consist of all of the objects posited by linguistic theories. It will include collections of phoneme-types, syntactic-property-types, and semantic-property-types—and rules (phonological, syntactic, and semantic) which determine how the various elements of the theory can be combined. It will also include a lexicon which will link simple expressions (i.e. complex phonological objects) with their syntactic and semantic properties.

Understanding language as abstract mathematical objects is not unique to generative linguistics. Lewis (1975), for instance, conceives of language this way as well. In Lewis' case, however, a language is something that is shared by an entire speech

⁷The work of Marr (1982), for instance, is cited by Chomsky (2000) and Jackendoff (1990) as being on a par with generative linguistics.

⁸Chomsky (2000, 25-26) stresses that we should not necessarily expect reduction of the linguistic to the neurological. It is enough that linguistics and neuroscience be made continuous with one another.

community. But if, in the context of generative linguistics, one has their language in virtue of their language faculty being in a particular state, then it is very unlikely that anyone actually shares a language with anyone else. After all, languages can differ not only with respect to phonological, syntactic, or semantic rules, but also with respect to the contents of the lexicon. Consequently, in generative linguistics, languages are not taken to be publicly shared objects, but as individuated at the individual level.⁹ Now, linguists *do* use terms like “English” or “Japanese”, and talk about various dialects of these languages. They should not be taken to be talking about public languages, however. Claims about English, for example, should be taken to be generalizations over the psychological states of speakers that we would informally characterize as being English speakers. Such generalizations will be justified insofar as the states of the language faculties of members of a population overlap. That there is a good deal of overlap is uncontroversial, as the very existence of typical human linguistic communities would seem to require this. Consequently, while languages, strictly, might be had by only one individual in the usual case, linguistics is not condemned to the study of the language of Mary and her neighbor John in isolation from one another.

To sum up thus far, then, generative linguistics studies human language by studying the human language faculty—a posit which is justified by the explanatory success of generative linguistics. The state of the language faculty of a speaker constitutes their linguistic competence in the target sense. Since the primary object of study is a faculty of mind, generative linguistics is a branch of the cognitive sciences, and its methodology, while distinctive in a number of ways—more on this in §2.2—accords with that of the cognitive sciences generally.

⁹Chomsky (1986) introduces the term “I-language” to refer to language as it is studied by generative linguistics.

2.2 Semantics In Generative Linguistics

I will now give a sketch the place of semantics in generative linguistics by outlining its explanatory aims, methodology, and by sketching the structure of semantic theories.

It is standardly thought that our linguistic competence (as studied in generative linguistics) is separable into at least three main kinds: the phonological, the syntactic, and the semantic.¹⁰ Our phonological competence, roughly, has to do with our ability to recognize certain sounds as being units of language—to create phonological representations in response to exposure to speech sounds and to generate these representations in the production of speech. Our syntactic competence is responsible for our ability to organize phonological representations into grammatical structures. Our semantic competence is responsible for our ability to assign meanings to those syntactic structures.

We can think of our semantic competence, then, or our “knowledge of meaning”, as being constituted by the state of a particular sub-faculty of the language faculty—for convenience, call it the “semantic faculty”. I suggest that we should think of semantics in a way that is analogous to the way I suggested we understand generative linguistics generally—semantics in generative linguistics is the study of the semantic faculty (or the semantic component of the human language faculty). The semantic faculty is responsible for our sensitivity to a range of semantic phenomena. Just which phenomena are the semantic phenomena, and just what their proper explanations are, is, again, an open and controversial matter. The following list, however, is at least

¹⁰I leave out pragmatics here not because it is somehow at odds with generative linguistics, but because our pragmatic competence is typically taken to lie outside of the language faculty proper. Our pragmatic competence—which (very roughly) underlies our knowledge of how to employ language in communication to achieve our various communicative ends—might consist in having knowledge of norms and conventions of the speech community. This sort of knowledge is less plausibly determined by the structure of some mental faculty, and is more plausibly determined by explicit or implicit beliefs that speakers employ in communication.

suggestive of the kinds of phenomena that have been of interest to semanticists.¹¹

Entailment-facts: Facts like the fact that, e.g., “Mary buttered the toast” follows from “Mary buttered the toast vigorously”, and “The door is closed” follows from “The door is not open”.

Semantic-anomaly-facts: Facts about which sentences are semantically well-formed and which are not—e.g., “Mary began the book” is semantically well-formed while (?)“Mary began the dictionary”, (??)“Mary began the rock”, and (*)“Mary is half tall” are not.¹²

Truth-condition-facts: Facts about which sentences are true in which circumstances.

Contradiction-facts: Facts like the fact that “John gave Mary flowers and John didn’t give Mary flowers” is a contradiction.

Synonymy-facts: Facts like the fact that “John gave Mary flowers” and “John gave flowers to Mary” mean the same thing.

Polysemy-facts: Facts like the fact that a single word ‘book’ can mean different things in different contexts—e.g., “Mary doesn’t believe the book” and “John sold the book to Mary”.¹³

Ambiguity-facts: Facts like the fact that utterances of “John wanted to marry a plumber” can sometimes have a *de re* meaning, and sometimes a *de dicto* meaning.

¹¹The list is by no means exhaustive and doesn’t represent any kind of consensus. Some might argue that certain items on the list can be given a syntactic explanation or pragmatic explanation, for instance. See Yalcin (2014) for a helpful discussion of the central semantic phenomena.

¹²See Pustejovsky (1995) and Kennedy & McNally (2005) for similar examples. A question mark before a sentence indicates marginal acceptability, double question marks indicate fairly strong unacceptability, and an asterisk indicates the strongest sort of unacceptability.

¹³The example comes from Pustejovsky (1998). In the one case, ‘book’ denotes an abstract body of information, while in the second case, it denotes a particular physical object. For many more such examples, see Pustejovsky (1995).

Productivity: The fact that speakers can understand and produce (any number of) novel complexes on the basis of grasping a relatively small number of simple lexical items.

Negative Interpretation-facts: Facts like the fact that “The Senator called the millionaire from Texas” can be read such that the millionaire is from Texas, that the call was placed from Texas, but not that the Senator is from Texas.¹⁴

Once again, to the extent that we have any sort of “knowledge” of these kinds of facts, the knowledge is tacit. Many of the above semantic phenomena are couched in theoretical terms, and so there’s no assumption that speakers explicitly recognize such facts, or that (in some cases) they even have the concepts to do so. Rather, this tacit knowledge is attributed to speakers in order to explain their linguistic behavior.¹⁵ For instance, speakers judge (1) and (2) to be unacceptable:

(1.1) *Gilgamesh failed often Herbology.¹⁶

(1.2) *Mary is half tall.

The observable linguistic behavior in this case is the judgments of English speakers that (1) and (2) are unacceptable in some sense. Speakers might respond to these sentences with a confused look or with the response that they make no sense. Linguists can also observe that tokens of (1) and (2) (and sentences with similar characteristics) simply aren’t produced by a particular population of speakers, despite the fact that

¹⁴This example comes from Pietroski (2006).

¹⁵Strictly, there’s no way to know, a priori, what explanatory relations will obtain in a scientific theory. Consequently, it’s an oversimplification to think that attribution of tacit knowledge of the listed semantic phenomena only helps to explain linguistic behavior (as opposed to say, other unobservable, theoretical sorts of facts). Nevertheless, such tacit knowledge is most often posited to explain speaker judgments and corpus data, so I focus on those here.

¹⁶This example comes from Adger (2003).

speakers might talk about Mary’s height or about Gilgamesh’s academic career.¹⁷ The linguist then posits that it is some aspect of English speakers’ linguistic competence that is responsible for the unacceptability of the sentences. The explanation for the unacceptability of sentences like (1), for instance, is taken to be syntactic. In other words, the best explanation for the unacceptability of (1) appeals to speakers’ tacit knowledge of syntactic rules and the syntactic properties of the words in (1). (1), according to syntactic theories of English, is ungrammatical, and it is supposed that speakers judge sentences that are ungrammatical in their language to be unacceptable. The explanation for the unacceptability of (2), however, is taken to be semantic. (2) is grammatical, but the meanings of ‘half’ and ‘tall’ are such that the former cannot modify the latter without unacceptability.¹⁸ Whether or not the explanation of some linguistic datum ends up being syntactic or semantic is a highly theoretical matter, settled in the course of engaging in linguistic research and theory construction.

It should be noted, too, that while speaker judgments are a very important source of data in generative linguistics, generative linguistics is not a theory of these judgments.¹⁹ Rather, it is supposed that the language faculty is partly causally responsible for these judgments. Since the language faculty is only partly causally responsible for these judgments, however, the evidence obtained from eliciting speaker judgments is highly defeasible. Speakers, for instance, judge very long or complicated sentences (e.g. so-called “garden path sentences”) to be unacceptable, despite the fact that they are syntactically and semantically well-formed. In those cases, insufficient attention, memory limitations, or other “performance” issues (i.e. causes that lie beyond the language faculty) might be the appropriate explanation for a particular judgment.

¹⁷While it is not uncommon for syntacticians and semanticists to appeal to corpus data, the most common source of evidence for syntax and semantics are speaker judgments. See Sprouse & Almeda (2012) for a defense of this reliance on judgments.

¹⁸The details of why this is so are discussed in Chapter 5.

¹⁹See, e.g., Chomsky (2000, 171-172).

Semantic explanations of semantic phenomena tend to appeal to the semantic values of the relevant expressions. Thus far, I have been talking loosely in terms of “meaning”, perhaps giving the impression that a notion of meaning plays an important role in semantics. But it does not. The concept MEANING is, of course, an important concept in the conceptual repertoire of our folk-descriptions of language and language use. As a rule, however, it is unreasonable to expect that folk-concepts will show up in our scientific theories of a given domain.²⁰ SEMANTIC VALUE (and not MEANING) is the fundamental explanatory notion in semantics. As with other such theoretical notions, we can understand SEMANTIC VALUE in terms of the various its role in semantic theories—i.e. in terms of the explanatory and evidentiary role its realizer plays in the theory.

What theoretical role do semantic values occupy, then?²¹ Broadly, they are posited in order to help explain semantic phenomena. Many of the semantic phenomena listed above, for instance, will be explained by appeal to the semantic value of the relevant expressions, or to relations between semantic values of the relevant expressions (in the case of entailment, for instance). Semantic values are also taken to be constrained by the principle of compositionality, which says, roughly, that the semantic value of a complex expression is determined by its syntactic structure, and by the semantic values of its constituents. This constraint is meant to help explain the productivity of natural language—the fact that speakers can understand (a huge number of) complex expressions that they have never encountered before.²² The idea, informally, is that if one knows the meaning of some stock of words, and knows how they are put together in some complex, then one can understand the complex. This is a bit simplistic,

²⁰See, e.g., Chomsky (1986, 1995). I will have more to say about MEANING and semantics in §3-§4.

²¹See Yalcin (2014) for a recent and illuminating discussion of the theoretical role occupied by semantic value in generative linguistics.

²²How, exactly, the principle of compositionality should be understood, and the role it plays in linguistic explanations, are taken up in detail in Chapter 4.

but it gestures at the structure of semantic theories. Roughly, the semantic theory attributes semantic value to simple expressions, and then via rules of composition—which are sensitive to syntactic structure of complex expressions and semantic values of their constituents—shows how the semantic value of a complex is arrived at. Thus, assuming that we think of the semantic faculty’s assignment of semantic value to an expression as being sufficient for a speaker’s understanding it, the compositionality of semantic value helps explain how it is that we can understand complex expressions that we have not encountered before.

The compositionality constraint also means that semantic value (and so semantics itself) is tightly tied to syntax. Indeed, theorizing in syntax and semantics do not proceed completely independently, since each heavily constrains the other. It is typically supposed, for instance, that what undergoes semantic composition is a syntactic structure—the object that is the result of syntactic processing. Consequently, puzzles that arise in the semantics of a given complex might require a different account of its syntactic structure, or vice versa.²³ Of course, the precise theoretical role of semantic value will be fleshed out only by engaging in semantic theorizing and will depend not only on the explanatory aims of semantics proper, but also its relation to other linguistic theories, and to theories of human psychology, generally. However, the demands to explain the central semantic explananda and to meet the compositionality constraint are the two main constraints that shape the form of semantic theorizing and the theoretical role of semantic value.

Different approaches to semantics have different proposals about the realizers of the semantic value role. Truth-conditional semanticists, for instance, think of the semantic value of an expression in terms of its truth-conditions (suitably understood) or its contribution to truth-conditions (in the case of sub-sentential expressions). An

²³Just what the connection is between syntax and semantics, however, is disputed. See, e.g., Culicover & Jackendoff (2005).

inferentialist, on the other hand, will understand semantic value in terms of the inferential role (suitably understood) of a given expression. A dynamic semanticist will understand the semantic value of an expression in terms of the way that speech acts involving that expression change a conversational context. And so on. While I will have more to say about potential realizers of the semantic value role in later chapters (especially Chapter 3), I want to maintain a broad and neutral perspective in this chapter.

One thing that I think ought to be stressed, however, is that while there is an expectation that facts about the semantic values of expressions will figure in fairly direct explanations of some of the potential semantic explananda I listed above, there is no guarantee that all of those phenomena will be explained in terms of semantic value. The list of phenomena I presented above are unified insofar as they are all pre-theoretically meaning-salient. They are phenomena that, were we to give folk-explanations for them, we would likely appeal to the meanings of the relevant expressions. Thus, the expectation that *some* portion of the phenomena be explained by appeal to semantic value is justified insofar as we think that, to be a *semantic* theory, the theory must make some contact with phenomena that we pre-theoretically think of as having to do with meaning. But we should not assume that all such phenomena are to be explained by appeal to semantic value or even by semantics at all. Our ordinary notion of meaning is very broad and messy, and is, very often, insensitive to any distinction between semantics and pragmatics, for instance. It is very plausible that we might find out in the course of our semantic investigations that not all of the phenomena ought to receive the same kind of explanation.

This, again, is the normal course of scientific theorizing—just what the central explananda are of a particular theory is something that we refine in our inquiry. It is, therefore, illegitimate to stipulate that any particular subset of the pre-theoretically

semantic phenomena are non-negotiable. To make such a stipulation would be an instance of what Chomsky (1995) calls “methodological dualism”—the placing of some (usually a priori) methodological constraint on linguistic theorizing that is illicit, and would be recognized as such in other scientific inquiries.

Consequently, I think it should also be made clear that there is no reason to stipulate, in advance, that there is only one realizer for the semantic value role. It’s not implausible that we might find several semantic properties of expressions, each of which figure differently in explanations of the pre-theoretically semantic phenomena.²⁴ Indeed, given the variety of phenomena that are pre-theoretically semantic, it seems to me that the more epistemically modest position is to expect that more than one sort of property will need to be appealed to in order to explain those phenomena.

3 Conceptual Concerns for Generative Linguistics: Language, Knowledge, and Meaning

In this section, I want to address some of the conceptual concerns that one might raise in objection to the characterizations of generative linguistics I gave in §2. For instance, given its focus on individual psychology, there might be concern that generative linguistics shouldn’t really be thought of as a theory of language, or that semantics shouldn’t be thought of as a theory of meaning. There might also be concern that facts about the state of one’s language faculty can’t ground knowledge states of any sort, and so generative linguistics shouldn’t be characterized in terms of linguistic competence

²⁴Whether or not there end up being multiple realizers for the semantic value role, however, will depend on just how one individuates that role. How a theorist individuates that role, it seems to me, is mostly a terminological decision, and so the difference between “monism” or “pluralism” about semantic value might not be particularly interesting. The point I am making might be better made by simply saying that there is no reason to suppose, a priori, that there will only be one property of expressions that is of central explanatory importance in semantics.

or knowledge, tacit or otherwise. I will argue in this section that to the extent that these concerns are justified, they are terminological concerns, and in no way impugn or constrain linguistic theorizing. The reason is just that these concerns arise for characterizations of the research program which are informal, and which are merely meant to gesture at the nature of the research program. I also argue that, understood this way, much of the choice of terminology is justified.

3.1 The Significance of Claims About What A Theory is About

What is generative linguistics about? Given my descriptions in §2, quite a few answers suggest themselves: human language, linguistic competence, the brain, the language faculty. Similarly, one might take semantics to be about meaning, semantic competence, or the semantic faculty, for instance.

Which of these is correct? The question, at least without further elaboration, is confused. In fact, it might be confused in a couple of different ways. The first, and perhaps more obvious way it is confused is that it would assume that the choices are exclusive. But this is clearly not the case. As I tried to emphasize throughout the discussion in §2, if one wants to study a subject as sprawling as human language, one will have to choose some particular way to do it. But it is clearly not the case that there is only one way to study language. The success of generative linguistics in no way impugns sociolinguistics, for instance, and indeed, they are not even competitors with one another, as they study different sorts of phenomena, both of which can be rightly characterized as being linguistic. “Language” is a term that is broad and messy enough that we should expect that there can be distinct, non-competing research programs which plausibly constitute inquiries into language. Generative linguistics can be said to study human language by studying linguistic competence, which it studies by investigating into the nature of the language faculty.

This brings us to the second way in which the question might be confused. Namely, one might mistakenly take claims about what generative linguistics is about to have philosophical import generally. If one views such claims this way, then they might take themselves to have substantive disagreements about the characterizations of generative linguistics I gave in §2. For example, someone who has some view about what meanings are, and who thinks that semantic values in generative linguistics are not meanings, might complain that, therefore, semantics in generative linguistics is not about meaning (or knowledge of meaning), as I have suggested it is. As I will make clear in §4, I am sympathetic to this conclusion, but I don't think it marks any substantive dispute with most characterizations of semantics in generative linguistics as being about meaning.

Now, I should be clear that I am not claiming that claims about what some theory is about are never intended as having philosophical import. My claim is that very often, they should not be taken this way, and many common characterizations of generative linguistics should not be taken this way. If we are asking about the ontological commitments of generative linguistics, then what it is about is just whatever the theory posits to explain a range of phenomena. It will be about semantic values, syntactic features, phonemes, and more broadly, the language faculty. Notions of language, meaning, competence, and tacit knowledge need not show up in the theory at all, and so endorsement of a theory in generative linguistics does not commit one to any views about language, meaning, or linguistic competence.²⁵ One might go on to identify meanings with semantic values, or think of languages along the lines of generative linguistics, but these are additional, philosophical commitments that are not obligatory in generative

²⁵A notion of truth shows up in truth-conditional semantics, of course, but there is no a priori reason for thinking that semantics in generative linguistics must be truth-conditional. It might turn out in the course of semantic theorizing that truth-conditional semantics is the best way to approach semantics, but this certainly wouldn't follow from a description of the subject matter of generative linguistics. I will have much more to say about the role of notions of truth in truth-conditional semantics in Chapter 5.

linguistics.²⁶

The sense in which generative linguistics is about language, meaning, or competence, then, is not philosophically committing (or at least, it is not mandatory that they be taken this way). Claims about what some theory or research program is about (and characterizations of a research program more generally), seem to me, generally have a different function. Namely, they are intended to give one a general sense of its theoretical aims, and of the methodology employed. Sometimes, a theorist must do this for an audience which lacks specialization in the field. In that case, one must employ ordinary, non-technical notions to give someone a sense of what the research program is like, and to help them locate it in relation to other forms of inquiry. To characterize generative linguistics as being about human language or about linguistic competence is to do just this. The first characterization indicates, for instance, that the research program is distinct from say, geology, and also from a study of communication systems generally (which might be non-human). The second indicates that the approach is broadly psychological, distinguishing it from theories which might, for instance, study the relation between marks or sounds in the world and various objects that humans might talk about using those sounds and marks. This indicates that the evidence for theories in generative linguistics will be like the evidence one finds in other psychological theories, and that it will employ a methodology appropriate to psychological theorizing, etc.

Characterizing generative linguistics as being about the human language faculty, of course, does the most to pin down its theoretical aims and methodology. It distinguishes generative linguistics from, for instance, certain philosophical approaches that have a claim to be inquiries into linguistic competence. Lewis (1980) takes the job of a grammar—i.e. a semantic theory for a particular language—to be to provide “part

²⁶These additional philosophical commitments are the subject of §4.

of a systematic restatement of our common knowledge about our practices of linguistic communication” (79). But he goes on to write that constructing a theory of linguistic competence that is “suited to fit into a psycholinguistic theory that goes beyond our common knowledge and explains the inner mechanisms that make our practice possible...is certainly not a goal I dare pursue” (81). Dummett (1993a), who, like Lewis, conceives of a philosophical theory of language as being a theory of linguistic competence, writes, “however psychology may evaluate it, Chomsky’s characterization of mastery of a language as unconscious knowledge contributes nothing to philosophical understanding” (xi).²⁷ Of course, to characterize generative linguistics as being about the language faculty is to characterize it in theoretical terms (“language faculty”). And so the characterization is only helpful once one is familiar with the term, and with the role the language faculty plays in explanations of various phenomena (where these phenomena, too, might need to be characterized in non-theoretical terms, at least to start).

In short, claims about what theories are about are typically made in order to give someone a sense of the theoretical aims and methodology of the theory. Consequently, the effectiveness of particular characterizations will depend on the background knowledge of one’s audience. For the sake of this dissertation, for instance, which assumes a philosophical audience, I will tend to characterize generative linguistics as being the study of the language faculty, as I think that this characterization is the least likely to lead to unnecessary confusion. It can also be read as having ontological import without bringing with it extra philosophical commitments about meaning, language, or linguistic competence. Now, again, one might be willing to take on such commitments, but they should not be taken to follow from informal characterizations of generative linguistics in terms of language, meaning, or linguistic competence. Consequently, any

²⁷See also Dummett (1993b, 1993c, 1993d).

concerns one might have with the characterizations of generative linguistics I gave in §2 should not be taken to constitute any concerns with the coherence of generative linguistics framework, but rather with terminological choices that are made in informal characterizations of that framework.

3.2 Terminological Concerns

Now, to claim that a dispute is terminological is not to claim that it is illegitimate. Even granting that characterizations of generative linguistics in terms of language, meaning, or competence are not meant to carry philosophical baggage, one might argue against the choice of terminology. Since I've argued that the usefulness of a particular characterization of a theory is heavily context-dependent, I am sympathetic to this complaint. I think it is inadequate and misleading, for instance, to simply characterize semantics in generative linguistics as being a theory of meaning, or a theory of semantic competence when writing for philosophers. One might argue that, in fact, the study of the semantic faculty constitutes a theory of semantic competence or a theory of meaning, but to simply characterize semantics as a theory of meaning or semantic competence fails to distinguish it from distinct, non-competing research programs. Such characterizations naturally lend themselves to confusion about the theoretical aims of a research program, and can give the illusion of overlap or competition where none exists.

I think we can say a bit more, however, about the baseline conditions of adequacy for an informal characterization of some theory. Why, for instance, is it appropriate, in some context, to think of semantics in generative linguistics as being a theory of meaning? As a really rough, crude start, we can say that the explananda of the theory are pre-theoretically meaning-salient. More particularly, it seems that were we to give explanations of the central explananda of semantics theory in pre-theoretical terms, we

would appeal to the folk notion of meaning to do so. Semantic value, in effect, occupies an explanatory role that overlaps to some extent with that of meaning in our ordinary conceptions of language. This is sufficient to establish a fairly strong conceptual overlap between the ordinary MEANING and semantics in generative linguistics. An informal characterization of semantics in generative linguistics as a theory of meaning, then, is an apt one in the right sort of circumstances.

Now, there might be concern that generative linguistics shouldn't be characterized as being about language because there simply isn't enough overlap between the domain of generative linguistics and our ordinary conceptions of language. After all, it seems plausible that we ordinarily conceive of language as being publicly shared, and as something we have in virtue of being a member of a speech community.

Likewise, one might object that "linguistic competence" and "knowledge of language" suggest that one can be incompetent with one's language, or have false (tacit) beliefs about it. But according to generative linguistics, this is simply not possible. The language the speaker is competent with is fully determined by the state of their language faculty. As the language faculty matures or changes, one does not become more or less competent with some language, but simply becomes "competent" with a new language. In other words, the very idea of "getting it wrong", about being mistaken about one's language in any way, makes no sense in the context of generative linguistics. This in no way impugns the coherence of generative linguistics, of course, because theorizing in generative linguistics is not beholden to our intuitions involving language or knowledge in the folk (or philosophical) senses (whatever those might be). Nevertheless, the informal characterizations of semantics as being a theory of semantic competence or tacit knowledge of meaning might be terminologically inappropriate.

In the current, philosophical context, I think that these concerns are legitimate. Indeed, Chomsky (1986) distinguishes between E-language and I-language, where I-

language is language as it is studied in generative linguistics, and E-language is language as a publicly shared entity. In some places, he also talks in terms of “cognizing” one’s language in place of talk of tacit knowledge.²⁸ In any case, I am happy to grant the terminological objections, and while I will, at times, use “language” and “tacit knowledge”, they should be taken as technical terms referring to the object of investigation in generative linguistics and the relation that speakers bear to states of the language faculty, respectively (except when otherwise specified).

It should be stressed that the tacit knowledge relation is not, therefore, something which needs to be defined in advance of theorizing. It merely gives us a way of talking about a speaker’s relation to the language they have in virtue of their language faculty being in a particular state. Just how we should understand that relation, more precisely, is a substantive question, but it is clearly not one that we can stipulate an answer to a priori.

4 Do Semantic Theories in Generative Linguistics Constitute Theories of Meaning?

I have thus far argued that one can accept a theory within generative linguistics—and informal characterizations of generative linguistics—without thereby incurring any philosophical commitments concerning language, meaning or knowledge of language and meaning. This leaves open the possibility, however, that one *should* think of generative linguistics as offering philosophical theories of language. And in particular, there is the possibility that one should adopt the philosophical claim that semantics in generative linguistics constitutes a theory of meaning. In this section, I want to offer

²⁸Though it ought to be noted that Chomsky (e.g., 2000, Ch. 3) does defend his use of “knowledge of language”, and argues that the notion of I-language overlaps substantially with our ordinary notion of language (Chomsky (1986, Ch. 2)).

some reasons to reject such a claim. While my arguments focus on the relationship between semantics and meaning, similar arguments apply to semantic competence and tacit knowledge of meaning.

One way to argue that semantics in generative linguistics does not constitute a theory of meaning or a theory of any philosophical import is just to argue from a prior philosophical view of what theories of meaning or theories of linguistic competence should look like. The quote from Dummett given in §3.1 suggests this angle. Dummett thinks of a theory of meaning as being a theory of speaker competence, but this competence cannot be characterized as purely unconscious knowledge—it must figure in rationalizing explanations of a speaker’s speech acts. The semantic phenomena I listed in §2.2 are not facts about speech acts or behaviors which call out for rationalizing explanations. Consequently, by Dummett’s lights, the subject matter of semantics in generative linguistics just doesn’t have philosophical import.

Similarly, if one thinks that a theory of meaning will be a theory of the relation that expressions bear to the things they refer to (for instance), one will not take semantics in generative linguistics to be a theory of meaning. Again, semantics in generative linguistics aims to describe the mechanism (the language faculty) which is fairly directly causally responsible speaker judgments of entailment, unacceptability, and so on. To describe this mechanism is not to describe reference relations and vice versa.

One might not have any of these views, however. And so I want to try to give some more general considerations against thinking of semantics in generative linguistics as constituting a philosophical theory of meaning, or as bearing directly on the philosophical study of meaning.

4.1 Semantics in Generative Linguistics and the Ordinary Concept MEANING

Philosophical theories of meaning are often disambiguated into two sorts—there are theories of meaning for particular languages and theories of what meaning consists in. The former kind of theory will state or “give” the meanings of the expressions of a given language, while the latter says what it is for a sentence to have whatever meaning it has. Both, however, are characterized in terms of the ordinary “meaning”. Crucially, this is not because it is convenient to informally characterize these theories in such terms, but because the theories are about—in the strongest sense—meaning. And importantly, their being about meaning in this sense is built into the official characterization of the project. In other words, it is a condition of adequacy of a philosophical theory of meaning that it be about meaning in the strongest sense. Even if one thinks that semantics in generative linguistics is about meaning in this sense, it is not a condition of adequacy on the theory that it be so.

But what does it take for a theory to be about meaning in this strong sense? It would seem that for a theory to be about meaning, it must bear a particularly strong relation to the ordinary concept MEANING.²⁹ The relation, presumably, must be stronger than is required for it to be appropriate to informally characterize a theory as being about meaning, for instance. One way a theory might satisfy this strong requirement is if it is just a theory of the ordinary concept MEANING.³⁰ Such a theory may or may not take the form of conceptual analysis in the traditional sense—i.e. a specification of necessary and sufficient conditions for something’s meaning ϕ —but

²⁹I focus on the concept MEANING, here, but similar remarks go for the English “meaning” and its counterparts in other languages.

³⁰See, e.g., Horwich (1998, 41n20; 2005, 175), who is explicit that his theory of meaning is a theory of our ordinary conceptions of meaning. Plausibly, his theory is also a theory of the property of meaning such-and-so the way we ordinarily conceive of it.

would take whatever form ends up being appropriate to the study of concepts. The project is both coherent and interesting, and seems to me to be of the sort that a social anthropologist might engage in, for instance. In the case of the ordinary concept MEANING, the aim would be to understand how a given population conceptualizes their linguistic practices. The methods appropriate to such an inquiry would also seem to fit within a broadly anthropological framework. The theorist will, presumably, be interested in the community's uses of "meaning", perhaps their intuitions about meaning, and so on.³¹

Endorsement of this first sort of theory need not commit one to any particular view of meanings, but only a view about how some particular community conceives of meaning. Or perhaps better, it commits one to a view about how a community conceptualizes their linguistic practice, irrespective of whether there is any referent for the community's concept MEANING.

It seems clear that semantics in generative linguistics does not provide an analysis of the ordinary MEANING. Speakers' speech behavior with the word "meaning" are only of interest to semanticists insofar as it is word that must be given a semantic value in a semantic theory of English. It is, thus, of no more interest to the semanticist than the English "cat" or "the", except insofar as it might end up posing interesting technical problems. Speakers' intuitions about "meaning" otherwise pose no constraint on semantic theorizing in generative linguistics.

There is another way in which semantics in generative linguistics might be a theory of meaning in the strongest sense, however. While it does not seek to study the concept MEANING, it might put the concept to use in explanations of linguistic phenomena.

³¹This sort of methodological approach is, of course, quite common in philosophy generally. And while it is less common for philosophers to claim to be interested in some ordinary concept, I think that the overlap in methods suggests that this is often the best way to interpret such theories. In that case, though, there is plenty of room for concern about the adequacy of philosophers' deployment of these methods. See, e.g., Chomsky (2000, Ch. 4) for similar concerns.

A theorist, relying on the theory of the concept MEANING to clarify that concept, might then use that concept to explain the kinds of semantic phenomena I listed in §2.2. This sort of theory, arguably, has a better claim to be a theory of meaning in the strongest sense than the first sort, since its focus would seem to be meaning itself, rather than the concept that refers to it. While the theory of MEANING might give one a sense of what sorts of explanations of linguistic behaviors meanings are supposed to participate in, the theory of meaning will give those explanations. What is explained in the second theory are those general linguistic behaviors, not merely the community's speech patterns concerning "meaning", for instance.

Now, it is *possible* that semantic theories in generative linguistics end up employing the ordinary MEANING in its explanations of various phenomena.³² This, however, would be no more than a coincidence, and an unlikely one at that. We do not expect, for instance, that our folk-physical concepts of force and momentum will end up playing a role in our best physics. Typically, as our theories in a particular domain mature, we find that our ordinary concepts are not suitable for explanatory roles in those theories. Consequently, we construct technical terms and concepts, which we can explicitly define, or at least we can go some way towards specifying the relations they bear to the rest of the theory.³³ So while it is possible that our ordinary concepts end up being suitable for use in our best scientific theories, the prospects of such an occurrence seem dim, and the track record of our scientific theories bears this out.

I think the prospects for MEANING are especially dim. It certainly admits of no easy definition, and seems to be an especially messy concept, as the philosophical literature surrounding the notion attests to. Thus, if any of our ordinary concepts end up playing important roles in our best scientific theories, MEANING would not seem to be one of them. And it should be noted that even if we grant that there is

³²This, for instance, seems to be the view of Horwich (2003).

³³See Chomsky (1995b, 2000) for extended discussion of these points.

some ordinary concept MEANING shared by our speech community—which will, itself, likely have very vague boundaries—the concept is an abstraction over a large group of individuals, many of whom will use “meaning” in conflicting ways.³⁴ Whatever such a concept ends up looking like, it seems very unlikely that it will have a claim to both being a concept that is shared (in some sense) by many members of a speech community, and being sufficiently clear, precise, and narrow enough to be suitable for use in theoretical explanations.

4.2 Semantics in Generative Linguistics and Conceptual Engineering

Another way that a theory might plausibly qualify as a theory of meaning is if it attempts to clarify, correct, or generally revise or replace the ordinary concept MEANING. Such a project will not merely be an attempt to provide an account of how a community conceives of their own linguistic practice, but will construct a concept which might not suffer from the same unclarity or defects (whatever those might be) of the ordinary MEANING. The project will, in effect, be an instance of “conceptual engineering”.³⁵ Such conceptual engineering projects are familiar in the philosophy of language from various attempts to “naturalize” semantic notions, for example.³⁶

For a theory with an engineered concept to count as a theory of meaning, the engineered concept must, at least, bear a strong resemblance to the ordinary MEANING. The theory might, for instance, need to deploy a concept which plays a role in explanations of linguistic phenomena that resembles the role that MEANING plays in our

³⁴See, e.g., Millikan (forthcoming) for arguments to the effect that there are no publicly shared concepts.

³⁵See Scharp (2013) for a useful discussion of conceptual engineering projects. See Cappelen (forthcoming) for further discussion of the importance of these projects in philosophy.

³⁶See, for example, Millikan (1984) and Dretske (1981, 1990). Similar sorts of projects are common in the philosophy of mind. See, for instance, Stich (1996) and Fodor (1987).

folk-explanations of those (or similar) phenomena.

But is this enough for a theory to be a theory of meaning in the strongest sense? It seems just as reasonable to think that if we engineer a new theoretical concept, SEMANTIC VALUE, for instance, and it plays a different role in explanations and in inferences, generally, then we have reason to distinguish SEMANTIC VALUE from MEANING. It would also seem reasonable to say something similar about a competing semantic theory which posits things called “meaning values”, which figure slightly differently in explanations of various semantic phenomena than do semantic values. To capture the fact that the two theories disagree about what sorts of semantic entities or properties exist, it seems reasonable to say that SEMANTIC VALUE and MEANING VALUE are distinct from one another, and that both concepts can be distinct from the ordinary MEANING.³⁷

Consequently, provided conceptual engineering projects engineer new concepts, and do not merely re-create ordinary ones, then it would seem that no conceptual engineering project will be a theory of meaning in the strongest sense (or any other phenomena construed in ordinary terms). That title will always belong to the theory which employs MEANING in the ordinary sense, or, to a lesser extent, the theory which is about MEANING itself. In that case, since (mature) scientific theories, generally, eschew ordinary concepts in favor of engineered, technical ones which are suited to play various theoretical roles, no successful scientific theory will be a theory of meaning in the strongest sense. The exception will be those scientific theories which study the ordinary MEANING, though these are probably best characterized as being theories

³⁷I echo a point from Chomsky (2000) here, who writes, “Structural phonologists postulated segments (phonemes) and phonetic features, with a certain collection of properties. Generative phonologists argued that no such entities exist, and that the actual elements have somewhat different properties. Suppose that one of these approaches looks correct (say, the latter). Were structural phonologists therefore referring all along to segments and in the sense of generative phonology? Surely not. They flatly denied that, and were right to do so” (43-44).

of MEANING in the strongest sense.³⁸ Since generative linguistics is a scientific research program which doesn't study MEANING, it is not a theory of meaning in any philosophically committing sense.

4.3 Conceptual Engineering and Concepts of Concepts

The preceding argument in §4.2 has far-reaching consequences for how we think of the relationship between ordinary and scientific concepts, and I do not, by any means, intend for it to be uncontroversial. Whether this sort of argument should be accepted or not depends crucially on how we individuate concepts. The argument presupposes that concepts are to be pretty finely individuated. In other words, given my framing, it is possible that the ordinary MEANING can be distinct from the theoretical SEMANTIC VALUE by virtue of the fact that, for instance, SEMANTIC VALUE bears a subtly different explanatory relation to various linguistic phenomena than does the ordinary MEANING. It is also possible, then, that while SEMANTIC VALUE has a referent, the ordinary MEANING does not (especially if we assume that what there is is dictated by our best scientific theories, for instance).

But on many views of concepts, such a possibility will be ruled out. For example, on views where concepts are to be individuated by their wide contents, so long as semantic values bear the right causal relation to uses of “meaning” and tokenings of beliefs involving MEANING, then SEMANTIC VALUE and MEANING might not be distinct. Or, even if, strictly, the concepts are distinct, it might be enough that SEMANTIC VALUE and MEANING have the same wide contents to think of semantics in generative linguistics as providing a theory of meaning in the strongest sense.³⁹ My

³⁸There can, of course, be a linguistic theory which employs the ordinary MEANING in the theory, but as I argued above, I think such an approach is very unlikely to be successful.

³⁹See, e.g., Kripke (1980), Putnam (1975), and Millikan (1984) for externalist views about (semantic and mental) content.

argument would also be rejected by those who differentiate concepts by their narrow contents, but do so in a more coarse-grained way, rejecting the kind of holism about conceptual content that my argument seems to assume.

How, then, should we think about concepts?⁴⁰ In the abstract, the question is similar to the question that has been in the background of this section, namely, “how should we think about meaning?” Here, again, we have two general options: we could investigate the ordinary concept CONCEPT (or construct a theory which employs that concept in various explanations), or we can engineer a concept, perhaps in the framework of contemporary psychology, which has a theoretical role which parallels that of the ordinary CONCEPT (again, assuming there is such an ordinary concept).

Some of the standard objections to fine-grained conceptions of concepts suggest that some ordinary notion of CONCEPT is the target. For instance, it could be objected that on a fine-grained understanding of concepts, communication is impossible because, strictly, nobody shares the same concepts—we only ever, strictly, talk past one another. For similar reasons, disagreement might become impossible. Similarly, competing theories of atoms, for example, would be about different things in the strongest sense of being about some phenomenon (and hence disagreement between theories is also illusory).⁴¹ These objections assume a view of communication and disagreement which—if it even is reflective of pre-theoretical ways of thinking of communication (which I think is contestable)—seems extremely simplistic and unlikely to be reflected in a sophisticated, scientific investigation of communication. Such an investigation, I would think, would need to reflect the messiness of communication. Successful commu-

⁴⁰The question might be better put, “How should we individuate concepts in the context of understanding our theories?” The question is better put this way because I see no tension between thinking of concepts as having narrow content and as having wide content. Wide content properties and narrow content properties surely both exist, and figure in different sorts of explanations. To the extent that there is any dispute here, it is over which gets called ‘content’, or how concepts should be individuated.

⁴¹See, e.g., Putnam (1988) and Fodor & Lepore (2002).

nication is not an all or nothing affair, requiring that we think of successful utterances as expressing something—a content—which is shared by speakers and hearers. It is enough that there be sufficient overlap between what a speaker intends to communicate and how a hearer interprets them to ensure that the hearer draws the sorts of inferences that the speaker wants them to draw, and so that they can coordinate with one another, generally. The objections only arise, then, if we suppose that the overly simplistic understanding of communication is the pre-theoretical way of thinking of communication, and that a theory of concepts is beholden to our ordinary CONCEPT and the role it plays in pre-theoretical understandings of communication.

To the extent that there even is something like an ordinary concept CONCEPT that we can plausibly investigate, it's unclear why that concept should play such a crucial role in our meta-theorizing. Indeed, it's not even clear that some close cousin of CONCEPT that figures in our best psychological theories should also dictate how we understand our theories, generally.⁴² Consequently, it's not clear that a single notion should do double-duty, figuring in broadly psychological explanations of behavior and also framing our understanding of our theories.

What sort of standards bear on the engineering of meta-theoretical notions? Broadly, it seems that—assuming some notion of a concept (related more or less to the ordinary CONCEPT) is necessary in our meta-theorizing—we would want to ask what sort of notion makes our theorizing go best. That, of course, is no easy question to answer. Concepts that bear close resemblances to ordinary ones might be preferred, since greater familiarity might mean that they are more easy and natural to use. However, clarity and precision are also important, and so unaltered ordinary notions will probably not suffice. I have no intention of settling the matter here. In my view, it is useful to individuate concepts in a fine-grained way. Consider, for instance, that given

⁴²Stich (1990, Ch. 1) makes similar remarks concerning the relation between the concept of reference, linguistics, and the rest of our theorizing.

how broad the ordinary concept MEANING is, many different sorts of theories will have claims to being theories of meaning in the strongest sense. For instance, if we individuate concepts in a more coarse-grained way, semantics in generative linguistics, pragmatics, and theories in sociolinguistics might all have a legitimate claim to being theories of meaning in the strongest sense. This might give the appearance that the theories are, in some sense, in competition, when they are clearly not. We could, of course, avoid this confusion by thinking of all of these theories as being part of a larger theory of meaning, and so think of semantic value (for instance) as being one aspect of meaning.

Alternatively, we could leave “meaning” out of it, except in cases where it is useful in informal characterizations of our theories. This, anyway, is my preference, since I don’t see what is gained by thinking of semantics in generative linguistics as a theory of meaning in any stronger sense.⁴³ The alternative approach, it seems to me, lends itself to greater confusion.

Absent convincing arguments one way or the other, however, it might be that the best one can do is to simply stipulate one notion of CONCEPT or other to avoid unnecessary confusion and dispute until this—in my view, rather fundamental—metaphilosophical issue is settled. In that case, while my argument against thinking of semantics in generative linguistics as being a theory of meaning in some philosophically committing sense begs the question against certain view of concepts, that might be the best that can be done. If, as a result, my argument is not persuasive, then opposing arguments shouldn’t be either. At best, then, we lack good reason to adopt any view about the status of semantics in generative linguistics as a theory of meaning.

⁴³Mutatis mutandis for other scientific theories and other phenomena characterized in pre-theoretical terms.

Chapter 2

Why Semantic Value Should not be Understood In Terms of Utterance Content

Chapter Overview: In this chapter, I begin to address the relationship between semantics in generative linguistics and the notion of content. The focus is on a recent paper by Schiffer (2015), which argues that truth-conditional semantics is a poor fit with generative linguistics. In particular, he thinks that it fails to explain speakers' abilities to understand the sentences of their language. In its place, he recommends his "Best Bet Theory"—a theory which aims to directly explain speakers' abilities to mean things by their utterances and know what others mean by their utterances. Facts about semantic values of expressions, on his view, are to be understood in terms of the content utterances of those expressions communicate.

I argue that Schiffer does not provide good reason to prefer the Best Bet Theory over truth-conditional semantics in the context of generative linguistics. The aim of this chapter is not primarily to defend truth-conditional semantics, however, but

rather to undermine Schiffer's grounds for thinking that his theory is superior to a truth-conditional approach—namely, that semantics must (very directly) explain our communicative abilities and consequently, that semantic value should be understood of the contents of utterances.

First, I address his negative arguments to the effect that truth-conditional semantics is inadequate because it fails to explain our abilities to mean things by our utterances and to know what others mean by their utterances. I argue that these arguments fail because they rely on placing illicit a priori constraints on semantic theorizing. Second, I argue that his insistence that semantics should directly explain our communicative abilities (and that the semantic value role be understood in terms of the contents of literal utterances) actually leads to a problematic circularity in the Best Bet Theory. Consequently, it is robbed of any potential explanatory power, and so is not even a viable competitor to truth-conditional semantics.

Importantly, the circularity arises for Schiffer's theory precisely because of the way it conceives of the explanatory role of semantic value—namely, in terms of facts about how we use expressions in utterances. Consequently, the argument generalizes beyond the Best Bet Theory to any semantic theory which understands semantic value in terms of the relation that expressions bear to the contents expressed by utterances of those expressions. I conclude that it is a mistake to conceive of semantics this way. The relationship between semantic value and content—hence semantics and theories of communication—is not so direct.¹

¹The material in this chapter appears in slightly altered form in Napoletano (2016).

1 Introduction

In his “Meaning and Formal Semantics in Generative Grammar”, Stephen Schiffer (2015) argues against the truth-conditional approach to natural language semantics, and in favor of his own “Best Bet Theory” (BBT).² In particular, he thinks truth-conditional semantics is a poor fit in the context of the Chomskyan, generative linguistics framework, where the primary object of study is a speaker’s linguistic *competence*—their ability to understand language. Given this description of the project, one of the primary explananda is taken to be what is often called the ‘productivity of natural language’—the fact that speakers of natural language are able to understand an indefinite number of sentences that they have never encountered before. The general outline of the explanation for this phenomenon given by generative linguistics is summed up in what Schiffer calls the “*Generative Grammar Hypothesis*” or “(GGH)”:

(GGH): The ability of a speaker of a natural language L to understand sentences of L requires her to have tacit knowledge of a generative grammar of L , that being a finitely specifiable theory of L that generates one or more syntactic structures for each sentence of L and interprets those structures both phonologically and semantically. (62)

(GGH) should not just be read as stating a mere necessary condition on the ability to understand the sentences of a language. Rather, the aim of linguistics is to say what this ability consists in. A correct semantic theory will specify the “tacit knowledge” which constitutes speakers’ semantic competence.³

²As Schiffer makes clear, ‘truth-conditional semantics’ is meant to be understood broadly, encompassing both the Montagovian, model-theoretic approach, and the Davidsonian approach. I will tend to use ‘truth-conditional semantics’ the same way, ignoring the distinctions between the two approaches.

³I place ‘tacit knowledge’ in quotation marks just to flag that it is an open empirical question just what this relation comes to. (GGH) is not meant to saddle generative linguistics with any substantial epistemological commitments concerning semantic competence. For detailed presentations of the generative linguistics framework, see Chomsky (1986) and Ludlow (2011).

Schiffer's claim, then, is that truth-conditional semantics fails to vindicate (GGH), and so is not a good fit with generative linguistics. One of its critical shortcomings, he argues, is that it fails to explain our communicative abilities—our abilities to mean things by our utterances, and to know what others mean by their utterances. These abilities, he thinks, are essential to the ability to understand sentences of a language. Thus, truth-conditional semantics cannot vindicate (GGH). The BBT, by contrast, has our communicative abilities as its target explananda. It explains these abilities by attributing to speakers tacit knowledge of what sentences can be used to mean in literal, unembedded utterances. I will argue, however, that the BBT does not offer a viable approach to doing semantics in generative linguistics. The problem, briefly, is that the appeal to literal utterances makes the theory explanatorily circular, and thus unfit for linguistic theorizing. Further, Schiffer's insistence that a semantic theory must explain our communicative abilities depends on the illicit assumption that generative linguistics is constrained by our pre-theoretical conception of what it is to be able to understand sentences. The upshot is that it's hard to see how communication-based theories like the BBT pose much of a threat to truth-conditional semantics in generative linguistics.

The structure of this paper is as follows. In section §2, I briefly consider his negative arguments against truth-conditional semantics and argue that they are not persuasive. I will argue in §3 that, in effect, even if Schiffer is right that our communicative abilities are at least partly constitutive of our ability to understand the sentences of our language, the BBT is a poor fit with generative linguistics, because it suffers from an explanatory circularity which is not easily avoided. I consider and respond to possible replies to this objection in §4 and then make some concluding remarks in §5.

2 Schiffer's Negative Arguments

Before raising objections to the BBT, I want to first make a few comments concerning Schiffer's arguments against the viability of truth-conditional semantics in generative linguistics. Schiffer argues that tacit knowledge of a truth-conditional semantic theory is neither necessary nor sufficient for the ability to understand the sentences of a language. Truth-conditional semantics, then, cannot vindicate (GGH), and does not comport with the explanatory aims of generative linguistics.

2.1 The Insufficiency Argument

Let's take the insufficiency argument first. Tacit knowledge of a truth-conditional semantics isn't sufficient for the ability to understand sentences, he argues, because that ability at least entails the abilities to understand what others mean by their utterances, and to mean things by our own utterances. But, he says, truth-conditional semantics, at best, explains our abilities to recognize truth-conditions and entailments, and falls short of explaining our communicative abilities.⁴ Even granting this last point, it's hard to see how this impugns truth-conditional semantics in generative linguistics. In short, the fact (if it is a fact) that the ability to understand sentences entails having communicative abilities, in the ordinary sense of 'understand', does not bear on a research program in cognitive psychology which seeks to describe the workings of a particular component of the human mind. It may be useful to informally characterize semantics in generative linguistics as giving part of an account of our ability to understand sentences, but—as is generally the case in naturalistic theorizing—theorizing is

⁴This argument comes out of the discussion in Schiffer (2015, 75-77). He cites Partee (2011) as an example of a truth-conditional theorist who grants that truth-conditional semantics should be thought of as explaining entailments and truth-conditions.

not constrained by our common-sense notions, whatever those come to.⁵

But then how should we read (GGH), which is construed in terms of “the ability to understand sentences”? It’s tempting to think that we need to first pin down what it is to be able to understand sentences, and then use (GGH) to constrain our linguistic theorizing. But this gets things backwards. The sense of “the ability understand sentences” which will be relevant to generative linguistics is an *a posteriori* matter, since we do not know ahead of theorizing just what phenomena are best explained by appeal to the language faculty, and which are better explained by appeal to extra-linguistic mechanisms. Nevertheless, (GGH), and talk of “the ability to understand sentences” give a useful informal characterization of the explananda of generative linguistics—psychological phenomena having to do with speakers’ competence with natural language. Again, we cannot simply enumerate, more precisely, what the relevant phenomena are ahead of theorizing, but they give a rough indication of the nature of the research program.

Likewise, semantics is often informally characterized as an account of our knowledge, or “tacit knowledge” of meaning. What this informal characterization does is suggest that what semantics is interested in is the component of the language faculty which is responsible for speakers’ competence with the pre-theoretically-meaning-salient aspects of language. Which phenomena are pre-theoretically meaning-salient? Traditionally, facts about synonymy, truth conditions, entailments, ambiguity, semantic anomaly, the abilities to interpret and communicate, etc. A semantic theory will likely be able to explain some subset of these phenomena, but we have no way of knowing *a priori* if a single semantic theory will explain all of them.⁶ Rather, the idea that a semantic theory can explain all of these kinds of facts would be a very bold empirical

⁵Thus, I think this is an instance of what Chomsky (1995) calls ‘methodological dualism’: placing constraints on theories of mind and language that everyone agrees are out of place in other naturalistic research programs.

⁶See Yalcin (2014) for relevant discussion.

hypothesis, amounting to the idea that they are all best explained by the workings of a single component of the language faculty. But it wouldn't be surprising at all if this weren't the case, and if it turned out that a variety of mechanisms were implicated in the explanations of these phenomena.⁷

The upshot is that like “the ability to understand sentences”, “tacit knowledge” and “meaning” are informal placeholders, useful for informally characterizing semantic inquiry.⁸ They should not, however, be taken to place serious *a priori* constraints on linguistic theorizing.⁹ Thus, the response from the truth-conditional theorist should just be that their theory is primarily concerned with that portion of the language faculty which is responsible for speaker's ability to recognize entailments and truth-conditions. These phenomena are certainly meaning-salient, and so truth-conditional semantics can rightly be called ‘semantics’.¹⁰ We do, of course, want some explanation for our communicative abilities, and truth-conditional theorists often do think that our tacit knowledge of truth-conditions plays an important role in such an explanation (though the role is more indirect than Schiffer allows for—more on this in §3). The point is just that we cannot reject a theory for failing to explain something it does not try to explain.

⁷The ongoing inquiry into the semantics/pragmatics distinction, e.g., suggests that it is not an easy task to say just which phenomena are semantic and which are pragmatic, for instance.

⁸The same sorts of considerations apply to the phrases “linguistic competence” and “semantic competence”.

⁹At least, they do not provide constraints of the sort that would adjudicate between, e.g., truth-conditional semantics and the BBT. They would, of course, rule out geological inquiries, for instance, as being semantic.

¹⁰That truth-conditional semantics does not explain our communicative abilities might be taken as evidence that the latter are explained by appeal to mechanisms which are not the truth-conditional theorist's target.

2.2 The Non-necessity Argument

Schiffer argues that tacit knowledge of a truth-conditional semantics isn't necessary for the ability to understand sentences roughly as follows. First, semantics should explain abilities of speakers that all competent speakers share. Thus, the "ability to understand sentences", in the relevant sense, is something that all competent speakers share. This premise seems fairly well-motivated, since linguistic theories of particular languages in generative linguistics are concerned primarily with the psychology of psychologically 'typical' speakers of those languages. Speakers with various aphasias and abnormalities, of course, are important sources of evidence, and are worthy objects of inquiry in their own right, but a theory of the "syntax/semantics of English", e.g., is an idealized account of the psychology of typical English speakers.¹¹ Schiffer claims, however, that competent speakers differ with respect to their abilities to recognize entailments, and that some competent speakers lack this ability altogether. Therefore, he concludes, the ability to recognize entailments is not necessary for the ability to understand sentences. Tacit knowledge of truth-conditional semantics, therefore, could not even partly constitute our ability to understand sentences, and so fails to vindicate (GGH).¹²

The claim that there are competent speakers that simply lack the ability to recognize entailments altogether strikes me as being very implausible. Would a speaker who did not recognize, for instance, that sentences like 'Ornette buttered the toast' are entailed by sentences like 'Ornette buttered the toast vigorously' understand these sentences? It wouldn't be a stretch to claim that such speakers are not competent with these kinds of constructions. A speaker who recognized no entailments of any English sentence would seemingly fail to resemble anything like a competent English speaker.

¹¹For convenience, I ignore the fact that various dialects of languages will differ from one another in theoretically important ways.

¹²The argument comes from Schiffer (2015, 75).

Putting aside intuitions about what it takes to be competent with a language, however, it is clear that *semanticists* take competent speakers of a language to be able to recognize not only entailments and ambiguities, but also things like infelicity, presupposition, and implicature as well. Now, it is true that speakers often will not directly judge that some sentence presupposes another, for instance. Indeed, they might simply lack a concept of presupposition altogether (and thus it might be better to say that speakers have a “tacit recognition” of these things). But semanticists, and linguists generally, do not proceed by eliciting speakers’ judgments about technical linguistic notions. Rather, they make use of, e.g., judgments of truth value and infelicity, translations, and various other sources of evidence, and posit that some sentence presupposes or entails another, for instance, in order to explain these sources of evidence.¹³ While Schiffer may have his doubts, he does not provide any reason for thinking that the established semantic methodology is defective or that its underlying assumptions are problematic. It’s difficult to see, then, how an intuition about linguistic competence could undermine the apparent success of the truth-conditional research program.

Now, Schiffer is surely right that speakers vary with respect to their abilities to recognize entailments, but this doesn’t threaten the idea that the ability is necessary for the ability to understand sentences of a language. Competent speakers, for example, vary with respect to their abilities to recognize grammaticality, and yet this does not threaten syntax as an account of speakers’ tacit knowledge of grammaticality (and the rules which determine grammaticality, etc.). As with syntax, an appeal to a competence-performance distinction might be relevant for the truth-conditional theorist here. They might, for instance, appeal to extra-semantic factors

¹³See, e.g., Matthewson (2004) for a general discussion of semantic methodology, including techniques for distinguishing entailments from implicatures utilizing speakers’ judgments of truth value and infelicity and other indirect sources of evidence. See also Ludlow (2011, Ch. 3) for an overview and defense of the role of speaker judgments in semantic theorizing. Thanks to an anonymous referee for suggesting that I include discussion on semantic methodology.

(memory/attention constraints, pragmatics, vocabulary limits, etc.) to explain why speakers only recognize a handful of a sentence's entailments, or struggle to recognize the entailments of certain kinds of sentences. Variations in speakers' abilities to recognize entailments, then, might be explained by differences in these extra-linguistic factors. Indeed, given that we would expect variation among competent speakers with respect to just about *any* linguistic ability, it's hard to see how this variation could be a special concern for truth-conditional semantics.

In sum, I don't think Schiffer's arguments against truth-conditional semantics should motivate us to abandon truth-conditional semantics in favor of his Best Bet Theory (BBT). They amount to the claim that in Schiffer's favored sense of 'understand', truth-conditional semantics does not vindicate (GGH). But we have no reason for thinking that a semantic theory, to be adequate, must vindicate (GGH) in that particular sense. Truth-conditional semantics can still vindicate perfectly reasonable readings of (GGH), on which the ability to understand sentences involves (at least) being able to recognize entailments. Truth-conditional theorists, then, shouldn't be terribly concerned by Schiffer's negative arguments.

3 Reasons Not To Bet On The Best Bet Theory

Schiffer's positive arguments for the BBT, however, would be troubling for the truth-conditional semanticist, since he argues that the BBT can explain everything that truth-conditional theories explain and more. This would make truth-conditional theories explanatorily otiose, and thus dispensable. I will argue here, however, that due to an illicit appeal to literality, the BBT suffers from an explanatory circularity which renders it unusable for theorizing in generative linguistics. Thus, whatever the demerits of truth-conditional semantics might be, the BBT is not a viable alternative.

3.1 A Sketch of the BBT

Before making this argument, however, I need to sketch Schiffer’s Best Bet Theory. Recall that for Schiffer, semantics in generative linguistics is tasked with giving an account of speakers’ abilities to understand sentences in their language. He thinks of this ability, in turn, as having tacit knowledge of the meanings of the sentences in their language.¹⁴ He also thinks, crucially, that possessing this tacit knowledge puts one in a position to understand what *speakers* mean by uttering those sentences with those meanings and to mean things by their utterances. So what are these meanings, and what is it for an expression of a speaker’s language to have the meaning it has, such that they can play the explanatory role that Schiffer requires of them?

The BBT differs from truth-conditional semantics both with respect to the nature of the meanings of expressions—i.e. the objects that a semantic theory assigns to expressions—and with respect to the relationship that obtains between an expression and its meaning. I will focus on the former difference first.

For Schiffer, sentence meanings are represented as ordered pairs consisting of a speech-act-type A and a propositional form Ψ , where we understand propositions to be more fine-grained than sets of possible worlds, as the truth-conditional semanticist might understand them. Following Schiffer’s own exposition of the BBT, we can think of propositions as structured, Russellian propositions—ordered pairs of the objects referred to by names, etc., and properties that are predicated of those objects.¹⁵ So, for instance, the BBT will represent the meaning of the English sentence “Is Hillary a politician?” as $\langle \text{Asking-whether}, \langle \text{Hillary, being a politician} \rangle \rangle$, and “She is a

¹⁴I use ‘meaning’ in place of ‘semantic value’ here and elsewhere to match Schiffer’s usage. ‘Meaning’ can be read as ‘semantic value’, throughout, however.

¹⁵Schiffer does not endorse the Russellian account of propositions, but adopts it for ease of exposition (2015, 80). The differences between his own view of propositions (2012, Ch. 2) and the Russellian view will not be relevant to my discussion, and so I will treat the BBT as making use of Russellian propositions.

politician” as <Meaning-that, < x_f , being a politician >>, “where ‘ x_f ’ holds the place for a female to whom the speaker refers in uttering the sentence” (2015, 81). Meanings of sub-sentential expressions can just be understood in terms of their contributions to the propositional forms associated with sentences.

The meanings assigned to sentences by the BBT, thus, differ from the extensions or intensions assigned to sentences by the truth-conditional theorist in two ways. First, meanings assign pairs of speech-act types and content forms to expressions, and second, the contents are not understood truth-conditionally.¹⁶ Now, Schiffer argues that what we communicate to one another by our utterances is more fine-grained than truth-conditional contents, and thus, tacit knowledge of the truth-conditional content of the uttered sentence won’t put one in a position to know what is meant by the utterance.¹⁷ This may be correct, but as I argued in §2, we haven’t been given good reason to think that this is a problem for truth-conditional semantics—it simply might not be its primary explanatory aim. However, it might be an advantage for the BBT if it can explain facts about our communicative abilities, in addition to the sorts of facts that are explained by truth-conditional semantics.

There is, of course, much to say about the differences between the BBT and truth-conditional semantics with respect to the nature of meanings, but for present purposes, I think the more interesting difference between the two theories has to do with the nature of the meaning relation—what it is for an expression to *have* its meaning. When

¹⁶There are linguists (e.g. Speas and Tenny (2003), Rizzi (1997)) who argue that information about the speech act type is encoded in syntax. In principle, there is no reason why a truth-conditional semanticist could not accept this view. However, they would still differ from the BBT theorist with respect to the meanings that they assign to the syntactic representation of a sentence. Thanks to an anonymous referee for raising this point.

¹⁷It is a well-known problem that knowledge (tacit or otherwise) of truth-conditions does not suffice for knowledge of meaning. See, e.g., Foster (2010) and Soames (1992). See Larson & Segal (1995), Heck (2007), and Kölbel (2001) for proposals to the effect that tacit knowledge of truth-conditions, in conjunction with certain interpretive dispositions, might suffice to be able to know what utterances mean. Other truth-conditional theorists will simply acknowledge that for *some* classes of expressions, meanings need to be understood in a way that is more fine-grained than truth-conditions.

a speaker has tacit knowledge of what one of their expressions mean, they tacitly know that this relation obtains between the expression and its meaning. In other words, we want to know what it is about <Asking-whether, <Hillary, being a politician>> that makes it the meaning of ‘Is she a politician?’ according to the BBT.

In fact, until we get some account of the meaning relation, we have no reason to think that the BBT is better positioned than truth-conditional semantics when it comes to explaining our communicative abilities and vindicating GGH. Merely pairing sentence-types with different sorts of objects does not tell us anything about the uses (i.e. tokens) of those sentences, even if the objects are pairs of speech-act types and content forms. Until we specify what the meaning relation consists in, we will not have any account of how speakers’ tacit knowledge of meaning figures in their communicative abilities.

What is the nature of the meaning relation on the BBT, then? Recall that the primary explananda of the BBT are our abilities to mean things by our utterances, and to understand what others mean by their utterances. In particular, our tacitly knowing that our expressions have the meanings they do must play a particular role in our abilities to mean what we do by our utterances, and to understand what others mean by them. Schiffer describes this role as follows:

“...at some point in the processing that takes [a speaker] from her perception of the utterance to her knowledge of what the speaker meant, the hearer (i) enters a state that constitutes her tacitly knowing that [a sentence] σ has [meaning-property] ϕ , after which (ii) that tacit knowledge is conjoined with other tacit or non-tacit knowledge she has, and (iii) from that combined state of knowledge she tacitly infers (no doubt via an “inference to the best explanation”) that the speaker meant such-and-such in uttering σ .” (78)

Schiffer argues that in order for tacit knowledge of meaning to play this role, sentence meanings must “constrain what speakers can mean in uttering sentences with those meanings, and constrain what can be meant narrowly enough so that uttering the sentence is an effective way to communicate propositions that satisfy the constraint” (79). Now, to be clear, the truth-conditional semanticist could and likely would endorse this picture of the role of our tacit semantic knowledge in communication. What is distinctive about the BBT is that it understands sentence meaning *in terms of* this role. What, then, is it for a sentence to have its meaning, such that tacit knowledge of a sentence’s meaning puts one in a position to be able to understand utterances of the sentence? According to Schiffer,

“What this suggests is that to know the meaning of a sentence σ is for there to be a type of speech-act A and a form of content Ψ such that one knows that in a literal and unembedded utterance of σ the speaker is performing an act of kind A whose content is of form Ψ . For example, knowing the meaning of ‘Is she a politician?’ requires knowing that in a literal and unembedded utterance of it the speaker would perform an act of asking-whether whose content was of the form x_f is a politician at t_u , where ‘ x_f ’ holds a place for a specification of the female to whom the speaker referred with his utterance of ‘she’ and ‘ t_u ’ holds a place for a specification of the time of utterance” (78).¹⁸

The thought, then, is that for a sentence to have its meaning $\langle A, \Psi \rangle$ is for it to be such that literal and unembedded utterances of that sentence are (or would be) used to perform speech acts of type A , having content of form Ψ . Sentence-meaning constrains speaker-meaning by constraining what speech acts a sentence can

¹⁸Similar quotes can be found throughout section 3 of Schiffer (2015).

be used to make in literal unembedded utterances of that sentence. Speakers' tacit knowledge of meaning consists in their tacitly knowing how sentences can be used in literal and unembedded utterances—i.e. tacitly knowing how sentence-meaning constrains usage—and they can exploit this tacit knowledge in their interpretation of both literal and non-literal utterances. Now, since the aim of semantics in generative linguistics is to describe this body of tacit semantic knowledge, the BBT—in pairing sentences with their meanings—will be making claims about how sentences can be used in literal and unembedded contexts. Schiffer writes,

“...if GGH is correct, then an internally represented generative grammar for English will issue in a theorem that pairs [‘She is a politician’] with [\langle meaning-that $\langle x_f$ is a politician \rangle], thereby encapsulating the information that in a literal and unembedded utterance of [‘She is a politician’] there is a female x (or so the speaker intends) such that in uttering [‘She is a politician’] the speaker means that x is a politician” (81).

To recap, then, we assume that speakers' tacitly knowing the meanings of the sentences of their language plays a crucial role in their being able to mean things by their utterances, and to know what others mean by their utterances. Schiffer's thought is that this is only possible if sentence-meaning sufficiently constrains speaker-meaning—i.e. uses of sentences. His proposal, then, is that a sentence S has its meaning $\langle A, \Psi \rangle$ in virtue of the fact that in a literal and unembedded utterance, S would be used to perform a speech act of kind A having content Ψ .¹⁹ Speakers' tacit knowledge of these facts is critical to their abilities to mean things by their utterances, and to understand what others mean by their utterances.

¹⁹See also Schiffer (2003, 112), where a sentence's having its meaning is a matter of it having its character*, where a character* is represented “by an ordered pair $\langle A, P \rangle$, where A is the kind of speech act that must be performed in a literal utterance of the sentence, and P is the kind of propositional content the speech act must have.”

Now, it is important to see why, according to the BBT, sentence-meaning constrains how sentences can be used in literal and unembedded contexts in particular. The restriction to literal and unembedded utterances is necessary if the BBT is to make true claims. Consider the restriction to unembedded contexts first. The embedded occurrence of “Hillary is a politician” in “Bernie said that Hillary is a politician”, for instance, is not used to perform a meaning-that Hillary is a politician. One can make such a report about what Bernie said without communicating (or meaning) that Hillary is a politician. Thus, if the BBT simply stated that “Hillary is a politician” is or would be used to mean that Hillary is a politician, it would simply make a false claim. It must be stressed here that the claims of the BBT are meant to imply that some sentence is used in such-and-such a way, and not other ways. If they did not have this implication, and the BBT just listed one of the ways a sentence is used, then there is no sense in which sentence-meaning provides any kind of constraint on usage, such that tacitly knowing sentence-meaning could play the role in interpretation that Schiffer requires.

Likewise for the restriction to literal utterances. In the right circumstances, one can utter “Is Hillary a politician?” not to ask whether Hillary has the property of being a politician, but whether she is trustworthy, for example. In general, there are not many limits to what we can use a given sentence to mean, given the right context, and enough background information. We might speak loosely, figuratively, or form an agreement with one another that we are to interpret certain utterances in a non-standard way. In many of these cases, it is simply not the case that we utter the sentence to mean what it literally means.²⁰ Thus, the restriction to literal utterances is also required if

²⁰Consider, for example, utterances like “The Broncos destroyed the Panthers”, which would typically be used to mean, roughly, that the Broncos defeated the Panthers in the contextually salient game by a wide margin. Neither the speaker nor hearer will seem to register at any level that the utterance also means that the Broncos organization literally destroyed the Panthers organization. Since in this context we are concerned with a kind of psychological theorizing, we can’t simply stipulate that the utterance also has the literal meaning.

the BBT is going to make true claims.

In the next section, I will argue that this restriction to literal utterances—while apparently necessary—is deeply problematic for the BBT. For now, though, I want to quickly argue that truth-conditional semantics understands the meaning relation differently from the BBT. The truth-conditional semanticist will agree that our tacit semantic knowledge plays an important role in our understanding what utterances mean, but she does not (or at least need not) identify an expression’s meaning what it does with the fact that it would be used to express such-and-such truth-conditional content in literal, unembedded utterances. Rather, for the truth-conditional semanticist in generative linguistics, the property *having such-and-so truth-conditional content* is a primitive semantic property—an expression’s having its truth-conditional content is not understood in terms of some relation the expression bears to uses of it, as is the case on the BBT.²¹

In fact, this shouldn’t be terribly surprising. Schiffer’s criticism is not merely that truth-conditional semantics fails to give an account of our communicative abilities, but that it has illicitly ignored these abilities, which should be among its primary explananda. After all, he argues, “Trying to theorize about sentence meaning without considering the very intimate connection between sentence-meaning and speaker-meaning is like trying to understand the properties of a hammer without knowing what hammers are for” (75).²²

This is not to say that the truth-conditional semanticist in generative linguistics

²¹This is not to deny that there will be the further question, “In virtue of what do expressions have the truth-conditional contents they have?” The point is that—except when we are providing a compositional explanation for the contents of complex expressions—this is a question which goes beyond semantic theorizing proper.

²²Schiffer also cites Heim & Kratzer (1998, 1), who claim that “to know the meaning of a sentence is to know its truth-conditions,” and Partee (1988, 49), who writes that “the real argument for compositional truth-theoretic semantics is not that language users can *understand* indefinitely many novel utterances, but rather that there are indefinitely many *semantic* facts to be explained, indefinitely many pieces of basic data about truth-conditions and entailment relations among sentences of any given natural language.”

sees *no* important connection between our tacit semantic knowledge and our ability to know what others mean by their utterances, e.g. But the route is more indirect, and the tacit semantic knowledge does not *consist in* tacitly knowing something about how sentences are used in literal utterances.

3.2 Circularity In The BBT

I argued in §3.1 that if the BBT is to make true claims about which speech acts sentences can be used to make, it needs to make claims about how sentences are used in *literal* utterances. I now want to argue that this restriction is problematic for the BBT, and that the properties that Schiffer thinks constitute sentences having such-and-so meaning cannot plausibly be the properties that semantic theories in generative linguistics should be concerned with. The BBT, it turns out, cannot vindicate any plausible reading of (GGH).

Consider again that on the BBT, for a sentence S to have $\langle A, \Psi \rangle$ as its meaning is for it to be such that in a literal, unembedded utterance of S , S would be used to perform speech act of type A with content of form Ψ . Thus, a sentence S 's having the meaning it has is a matter of it bearing a particular relation to the unembedded, literal utterances of it—roughly, *being able to be used to perform such-and-so contentful speech act when uttered in literal, unembedded utterances*.

But how are we to understand literality here? A standard way to think of literal utterances is as the utterances where a speaker speaks literally—where the speaker-meaning of the utterance is identical to the meaning of the sentence. (There are complications raised by indexicals here, which can be handled by invoking Schiffer's notion of semantic content.²³ I will be ignoring these complications since the discussion

²³The semantic content of an utterance of S is, roughly, any proposition communicated by an utterance of S which has the form of Ψ , where Ψ is the propositional form associated with S by its meaning (81-82). We can therefore think of the literal utterances of S , generally, as the utterances of

does not depend on addressing them.) The concern here is that this understanding of literalness would make an account of what it is to be the meaning of S circular. Again, to be a literal utterance is for the utterance to bear a certain relation to the meaning of the uttered sentence—*being an utterance of S which is used to mean the same thing as what S means*. Thus, what it is to be a literal utterance of S is understood in terms of what it is to be the meaning of S —*being the meaning of S utterance* is explanatorily prior to *being a literal utterance of S* . But what is it to be the meaning of S ? On the BBT, it is to bear a certain relation to the literal utterances of S . Thus, *being a literal utterance of S* is explanatorily prior to *being the meaning of S* and vice versa. The accounts of what it is to be the meaning of S , and what it is to be a literal utterance of S are, therefore, circular.

Another way to put the problem is to say that since constitution is an explanatory relation, *having meaning M* and *being the meaning of S* simply cannot consist in the properties that the BBT says they consist in.²⁴ Thus, the BBT does not attribute properties to sentences which can play the explanatory role required to satisfy GGH (or any other constraint for that matter). Therefore, if we want the fact that a sentence has the meaning it has to play any explanatory role in our semantic theories, we cannot think of what it is to have a meaning the way the BBT would require us to.

Now, even if one isn't convinced there is any viciousness to this circularity, it should be noted that having tacit knowledge of the meanings of sentences, on this view, would not be useful in interpreting utterances. On the proposal under consideration, a speaker's tacitly knowing that S means M will consist in tacitly knowing that a sentence S can be used—in unembedded utterances where S is used to speaker-mean

S where only contents of form Ψ are expressed.

²⁴As I argue in Chapter 4, something's having a particular property, in general, does not consist in its bearing a relation to things with that very property. In this case, a sentence S 's having a particular meaning M is (roughly) taken to be a matter of standing in a certain relation to uses of S with that very meaning M .

S's (sentence) meaning—to mean *M*. But a speaker with this sort of tacit knowledge will not be in a position to interpret utterances of *S*. This tacit knowledge, to be useful, would require that the speaker can identify which utterances are literal—which utterances are utterances where *S* is used to speaker-mean *S*'s sentence meaning. Knowing that *S* can be used to mean *M* in certain conditions will only help you know what an utterance of *S* means if you know that those conditions obtain. Typically, however, our being able to identify an utterance as a literal one is something which is dependent both on our knowledge of what a sentence means, and also our ability to antecedently grasp what is speaker-meant by the utterance. The proposal is problematic, then, for if our tacit knowledge of meaning is meant to explain our ability to grasp what others mean in uttering sentences, our utilizing this tacit knowledge in interpretation cannot depend on an antecedent ability to grasp what others mean by their utterances.

If it's not obvious that our ability to grasp whether an utterance is literal or not depends on our knowledge of the speaker-meaning of the utterance, consider how unusual it would be to be fail to know what a speaker meant in uttering *S*, yet succeed in knowing that they spoke literally.²⁵ We can, of course, imagine cases where we could know that someone is speaking literally without knowing what they or the uttered sentence meant—a trustworthy translator might assure us that John is speaking literally, for example. But ordinary communication is not like this, and what we are trying to explain is how we are able to communicate in the ordinary case.

Now, the BBT theorist might respond in the following way. They might note that it's not necessary that speakers be in a position to antecedently know that their interlocutors are speaking literally in order to come to know what they mean by their utterances. Rather, it is enough that speakers have a defeasible disposition to interpret

²⁵Note that we might have good reason to think that someone speaks *sincerely* without knowing what they or the uttered sentence meant, but the sincere/insincere distinction does not track the literal/non-literal distinction.

utterances literally. That is, supposing that the semantic component of their language faculty assigns M to S , then speakers will have a defeasible disposition to interpret utterances of S as having speaker-meaning M . In some cases, there will be good reason to think that the speaker's utterance does not mean M , and the interpreter will engage in additional pragmatic reasoning in order to interpret the utterance. Otherwise, provided the speaker speaks literally, the interpreter will come to correctly interpret the utterance. The disposition to interpret literally, in conjunction with speakers' tacit knowledge of how sentences can be used in literal and unembedded utterances suffices for their being able to know what others mean by their literal, unembedded utterances.

I don't think there is anything inherently problematic about this proposal. However, I think adopting it would undermine Schiffer's case against truth-conditional semantics, and the interest in the BBT, generally. Consider, first, that if we posit that interpreters have a defeasible disposition to interpret utterances literally, then there is little point in understanding the meaning relation as Schiffer does—as a relation between a sentence and its uses. Rather, we could just take the meaning relation as basic, so that speakers would just have tacit knowledge that S means M , for instance, and the disposition to interpret utterances of S as meaning M (because the language faculty pairs M with S). From the standpoint of explaining speakers' abilities to understand what others mean by their utterances, this proposal is equivalent to Schiffer's, on which tacit knowledge of meaning consists in tacit knowledge of how sentences can be used in literal and unembedded utterances.

But if we take the meaning relation to be basic, then the BBT starts to look a lot more like truth-conditional semantics. The truth-conditional semanticist, recall, also takes the meaning relation as basic, though they differ with respect to which objects they take to be meanings. One might think that since meanings, on the BBT, are more fine-grained than truth-conditions, that the BBT will thereby have a distinct advantage

over truth-conditional semantics. However, there are already truth-conditional views which purport to show how our tacit knowledge of truth-conditions—when combined with suitably complex dispositions to deploy this tacit knowledge in communication—suffice to explain how we mean things by our utterances, and come to know what others mean by their utterances.

Larson & Segal (1995, ch. 2), for instance, propose that our tacit semantic knowledge consists in tacit knowledge of a compositional T-theory, which compositionally computes T-sentences of the form ‘*S* is true iff *p*’, where ‘*p*’ is a translation of *S*. They posit that after the semantic component of the language faculty outputs such a T-sentence, there is a tacit inference (or arational “shift”) to the belief that *S* means that *p*. From here, they can suppose that there is a defeasible disposition in speakers to mean that *p* by their utterances of *S*, and to interpret others’ utterances of *S* accordingly.²⁶ Thus, tacit knowledge of truth-conditions can play the role in communication that is required of our tacit semantic knowledge without understanding the meaning relation as one that obtains between a sentence type and its uses.

Without additional argumentation against views like these, it’s hard to see what we gain by positing the more complex meanings countenanced by the BBT. The BBT requires them because it takes our tacit knowledge of meaning to play a more direct role in our communicative abilities (i.e., a role which is unmediated by dispositions to interpret utterances literally). They would thus be justified if the relation between our tacit semantic knowledge and our communicative abilities is as tight as is suggested by Schiffer. One can give up on this feature in order to avoid the circularity concerns raised above, but doing so undermines Schiffer’s criticism that truth-conditional semantics has illicitly ignored our communicative abilities, and thereby undercuts the motivation for the BBT altogether.

²⁶See Heck (2007) for a similar kind of proposal.

Let me recap the argument. Suppose that speakers have a defeasible disposition to interpret utterances literally, and that this disposition combines with their tacit knowledge of meaning in interpreting utterances. In that case, there is no advantage to understanding the meaning relation as a relation between sentences and its uses, as Schiffer does, rather than simply understanding the meaning relation as basic. However, the truth-conditional theorist also conceives of the meaning relation as being basic. In that case, the only important difference between the BBT and truth-conditional semantics has to do with what they take meanings to be. But, unless we have some argument against views like Larson & Segal's (and Schiffer does not provide one), or unless we assume that tacit knowledge of meaning figures more directly in explanations of our communicative abilities (i.e. we give up our initial supposition), we have no reason to posit the more fine-grained meanings countenanced by the BBT. As I've argued, however, if we do give up our initial supposition, then the circularity of the BBT renders the theory explanatorily empty.

Assuming the BBT-theorist can neither take the meaning-relation as basic nor appeal to interpretive dispositions to do much of the work in explaining our communicative abilities, the question is whether she can understand the meaning relation in a way which avoids the circularity I've pointed to in this section. In particular, is there some other, non-problematic notion of literality that Schiffer can appeal to? If not, is there some other way to think of how sentence-meaning constrains speaker-meaning which will comport with the explanatory aims of generative linguistics? In the following section, I consider some proposals along these lines, and argue that they do not help the BBT.

4 Can The BBT Be Saved?

4.1 ‘Literal meaning’ As ‘Compositionally derived meaning’

Sometimes, when we speak of literal meaning, we are talking about a sentence’s compositional meaning or the meaning that results when we combined the meanings of its constituents in the usual way. Thus, “John kicked the bucket”, understood compositionally, entails that there was some bucket that was kicked by John. On its usual, figurative reading, it just means that John died. Maybe, then, literal utterances are utterances where the speaker-meaning “conforms with” the meaning that results by combining the meanings of the sentence’s sub-sentential expressions in the usual way. On this understanding of literality, we avoid the worrisome circularity by appealing to the meanings of sub-sentential expressions instead of sentence meaning.

But there are problems with this view. First, one might think that the meanings of simple expressions are fixed by their contributions to the meanings of the sentences they appear in. In that case, sentence-meaning is metaphysically/explanatorily prior to simple-expression-meaning. This position is obviously incompatible with a view on which literality is prior to sentence-meaning, and is understood in terms of the meanings of simple expressions.

But regardless of metaphysical issues concerning the priority of the meanings of sentences or simple expressions, when it comes to linguistic theorizing, it seems plausible that sentence-meanings are *epistemically* prior to the meanings of simple expressions. That is, whether or not the meanings of simple expressions just are their contributions to sentence meaning or not, we (theorists) come to find out the semantic properties of simple expressions via their contributions to the meanings of the sentences they are part of. Thus, the theorist’s grasp of the latter is prior to our grasp of the former. But if we think of sentence-meanings the way the BBT does, then theorists will need to

know which utterances are the literal utterances to know which uses of a sentence actually constitute evidence for thinking it has a particular sentence-meaning. But to know which utterances these are, we need to first know the meanings of the simple expressions which compose the sentence. Thus, simple-expression-meaning is epistemically prior to sentence-meaning, which contradicts our starting assumption.²⁷

4.2 ‘Literal meaning’ As ‘Explanatorily basic meaning’

Here is another way we might understand literal usage: a use of S literally means that content Ψ when the fact that S is used to mean Ψ explains why S can be used to mean things other than Ψ . In other words, any given sentence S can be used by speakers to mean any number of things, but its having the literal use it has is what explains how speakers can use it to mean what they do. literal utterances are utterances where the speaker-meaning of an utterance of a sentence is this explanatorily basic speaker-meaning. So, sentence-meanings explain why the sentence has the literal usage it has, and its having the literal usage it has explains its non-literal uses. Literal usage, here, is not defined in terms of the meanings of sentences or expressions, but in terms of the various speaker-meanings that speakers can use S to express. This proposal thus avoids the circularity concerns raised above.

This understanding of literality fits well with analyses of implicatures and metaphors, for instance, for which it is often thought that we arrive at a speaker’s non-literal meaning via an inference from the literal meaning and some norms of conversation perhaps. The thought would be that while speakers infer speaker-meaning on the basis of literal meaning, the theorist works backwards: we notice that a sentence S is used to mean various things by speakers in different contexts, and we notice that certain uses are

²⁷It also seems as though simple expressions can have literal or figurative uses. For instance, ‘point’ in ‘I see your point’. It’s hard to see how we could understand the meanings of simple expressions except by reference to their contributions to the sentences they appear in.

basic, in that they can explain—in conjunction with various pragmatic principles—the various other uses. These explanatorily basic meanings are the literal meanings. This is enough for us to get a grasp on which utterances are the literal ones. From here, we can figure out the meanings of sentences.

An immediate concern with this approach is that it seems unlikely that for many sentences of natural language, there is some usage that explains all others. The approach may work for some implicatures and metaphors, but two speakers might, for instance, arbitrarily agree that for one night, any sentence with six syllables will be used to mean that the speaker wants a sandwich. One of them later utters, “The dog barked all night long” in order to mean that they want a sandwich. The interlocutor (perhaps after some reflection) realizes that the uttered sentence has six syllables, and thus that the speaker means that they want a sandwich. In this case, it is nothing about the literal meanings of uses of “The dog barked all night long” that explains why it is used to mean that the speaker wants a sandwich. Rather, it is just a prosodic feature of the sentence that figures in the explanation—i.e. that it has six syllables. Examples like these are easy to come by, and in principle, there is no limit to the kinds of facts that can explain particular uses of sentences.

Putting this concern aside, there is a deeper issue with the proposal. The problem is that if we specify literal utterances for a particular sentence in terms of a particular speaker-meaning, then sentence-meaning just won’t contribute to explanations of how it is that we can grasp speaker-meaning. Consider what it is to have a particular sentence-meaning, where we unpack ‘literality’ in the way under consideration here. A sentence S ’s having sentence-meaning $\langle \text{meaning-that, } \Psi \rangle$ will be a matter of it being such that if S is uttered in an unembedded context where it is used to mean its explanatorily basic content Ψ , then it can be used to mean Ψ . Thus, we posit that speakers tacitly know that in utterances where S is used to mean its explanatorily

basic content, it would be used to mean Ψ , and this explains how, in such utterances, they are able to know that an utterance of S means Ψ .

The problem is just that attributing this tacit knowledge to speakers does not at all explain how it is that we can grasp what is meant by an utterance because we don't have transparent access to whether an utterance is an "explanatorily basic" one or not. Without being able to recognize which utterances these are, the attributed tacit knowledge is inert. Further, we can't attribute to speakers an ability to recognize which utterances these are, since presumably, any such ability would require that they be able to grasp what S is being used to mean, and then be able to determine whether the use of S is the explanatorily basic one or not. But this would be to attribute to speakers the very ability the BBT is trying to explain in the first place—*viz.* the ability to know what uses of sentences mean.

4.3 What A Sentence Is "Standardly" Used To Mean

Another way to remove the circularity issues would just be to focus not on literal utterances, but on utterances where a sentence is used to mean what it is *standardly* used to mean.²⁸ This, in turn, can be understood in a couple of ways: either as what it is usually used to mean, or what it is conventionally used to mean. I take up these proposals in turn. First, we could take the standard utterances to just be those utterances where, in uttering S , a speaker means what speakers *usually* mean by uttering S . These utterances would be easy enough for theorists to identify (in principle), but as with the proposal in §5.2, it's not obvious that speakers typically have any kind of access to which utterances are the relevant ones. Thus, attributing tacit knowledge of what a sentence can be used to mean in utterances where it used in the usual way might not go very far towards explaining how we can grasp what

²⁸See, e.g., Alston (1999).

utterances of the sentence mean. But perhaps the explanation could go as follows: speakers tacitly presuppose that speakers use sentences in the usual way, and their tacit knowledge of meaning consists in their knowing that usually, utterances of *S* are used to mean Ψ . Thus, they come to tacitly infer that *S* is being used to mean Ψ , and for the utterances where *S* is used in the usual way, they come to have knowledge that it is being used to mean Ψ .²⁹ This avoids attributing to speakers a prior ability to grasp what is meant by utterances, and doesn't presuppose that they have access to which utterances are of the usual sort.³⁰

However, I think this proposal is in conflict with plausible, psychologically well-grounded views about the divide between semantics and pragmatics. It is, for instance, common to think of implicatures as being arrived at via processes and mechanisms which are independent of the semantic module, and are better thought of as being part of more general cognitive abilities that aren't tied specifically to the language faculty. But in many cases, it seems as though an implicature is present more often than not. Take scalar implicatures, for instance. An utterance of "Some of the students left", more often than not, implies that not all of the students left. Thus, the proposed semantic theory would be tasked with explaining phenomena that we have good, independent reason to think are extra-semantic. But if there are good grounds to think that grasping the implicature involves pragmatic mechanisms which go beyond the semantic module, then we have good reason to think that semantics shouldn't be concerned with explaining facts about what sentences are standardly used to mean.³¹

²⁹In non-literal utterances, the hearer will drop the presupposition that the context is a usual one (on the basis of various sorts of contextual evidence), and then arrive at the correct speaker-meaning (often enough, hopefully) via pragmatic processes.

³⁰I don't think the same sort of adjustment could be made for the proposal in §4.2. In that case, we would suppose that speakers tacitly assume that utterances of *S* mean the explanatorily basic use of *S*. But it seems as though the notion of explanatorily basic use is too theoretical to plausibly suppose that ordinary speakers have even a tacit conception of it.

³¹Even if one thinks that scalar implicatures are semantic phenomena, it wouldn't be difficult to come up with other examples that are even less plausibly semantic, or to imagine a human linguistic community which, most of the time, employed paradigmatically pragmatic processes with respect to

Another way to understand standard uses would just be as the uses where a sentence is used to mean what it *conventionally* means. Now, conventions do not obviously figure into semantic theories in generative linguistics since the latter are interested in states of the human language faculty, which likely assign semantic properties to expressions distinct from those that might be assigned by convention. This, however, is precisely what would allow the proposal to avoid circularity—conventional meaning is not to be understood in terms of semantic theories but some other theory of conventional meaning. I assume, though, that this is a non-option. Even if we thought that we needed a distinct theory of conventional meaning, it would be very odd to think that semantic theorizing in generative linguistics would be so completely beholden to such a theory.³²

5 Conclusion

The upshot of §3 and §4 is just that it's difficult to conceive of Schiffer's BBT in such a way that it could be a promising semantic research program in generative linguistics. The problem, again, is that if a semantic theory is going to directly explain facts about the uses of particular sentences, the uses need to be suitably restricted to certain sorts of utterances. On the surface, it would seem as though the relevant utterances would be the literal ones. These, after all, will presumably be the utterances where extra-semantic, pragmatic processes do not figure so heavily in a speaker's knowledge of speaker-meaning. In that case, we'd expect the attribution of tacit knowledge of a semantic theory to play a central role in an explanation of how a speaker comes to

uses of certain sentences of their language (e.g., metaphor).

³²One other desperate strategy might be to just abolish the literal/non-literal distinction, so that every use would, in effect, be treated as "literal". The result would be the positing of massive and widespread ambiguity. In that case, it wouldn't be the BBT that would explain our communicative abilities, but whatever it was that allowed us to disambiguate properly.

know what someone means by their utterance of a particular sentence.

As we have seen, however, the central explananda of a semantic theory cannot be facts about uses in literal utterances, or facts about our abilities to mean things by and understand literal utterances of sentences. This is not to say that a semantic theory cannot figure, *indirectly*, in explanations of these facts, but just that semantics should not be understood *in terms of* these facts or in terms of an explanatory relation to these facts.

An indirect, truth-conditional explanation of our ability to understand what utterances mean, e.g., could go as follows.³³ Upon hearing an utterance of *S*, the semantic component of a hearer's language faculty compositionally derives the truth-conditions of *S*. This is as far as the semantics goes. The tacit knowledge of truth-conditions is then accessed by other, extra-semantic faculties, which—in conjunction with other tacit and non-tacit information—make a tacit inference to the best explanation of what the speaker meant by uttering *S*. In many circumstances, the hearer will take the utterance to have the truth-conditions to be the same as those that are assigned to *S* by the semantic component of their language faculty. In circumstances where the speaker has reason to doubt that the utterance has those truth-conditions—because it conflicts with what the speaker believes about the speaker, or the utterance would violate certain pragmatic norms—the initial interpretation of the utterance is overridden, and a new interpretation is arrived at which, hopefully, is the intended meaning of the utterance. Facts about literal-use-understanding are thus explained via attribution of tacit knowledge of a semantic theory—the deliverances of which are still *true*—and facts about the way in which this tacit knowledge is put to use by other, extra-semantic faculties of mind.

This, in principle, would seem to be the beginnings of an explanation of how it is

³³I abstract away from details here, like the fact that the truth-conditional theory might not pair sentences with truth-conditions, but something sub-propositional.

that we are able to understand utterances of sentences of our language—indefinitely many, provided the semantic theory is compositional. Now, the explanation may not be the right one, and my aim here is not to provide any kind of general defense of truth-conditional semantics in generative linguistics.³⁴ My aim is, rather, to point out that truth-conditional semanticists do not, in fact, ignore the relationship between our tacit semantic knowledge and our communicative abilities altogether.

Schiffer's criticism is that facts about our communicative abilities do not figure centrally enough as explananda in truth-conditional semantics. But, as I have argued here, not only is it illicit to argue, *a priori*, that semantics must be primarily concerned with explaining our communicative abilities, it's not clear that such an approach is a promising way to pursue semantics in generative linguistics at all. Facts about the semantic values are not facts about how expressions can or would be used in utterances, nor should they be understood in terms of some relation they bear to those utterances.

³⁴See, e.g., Pietroski (2003) for a criticism of the viability of truth-conditional semantics in generative linguistics.

Chapter 3

Why Semantic Value Does not Determine Content

Chapter Overview: In Chapter 2, I argued that semantic value should not be understood in terms of the contents expressed by utterances. Thus, it leaves open the possibility that semantic value and content, nevertheless, bear a close relation to one another. Indeed, the standard assumption—which has the status of a guiding methodological principle in much of semantic and pragmatic theorizing—is that the semantic value of an expression determines the contents that utterances involving that expression express. This principle—the “determination principle”—ensures that there is a tight connection between semantics and theories of communication. After all, semantic theories are meant to be theories of our semantic competence, and our semantic competence surely plays an important role in our ability to communicate with one another using language.

In this chapter, I argue that while it is true that our semantic competence plays an important role in communication, no version of the determination principle obtains that could plausibly serve as a guiding methodological principle in generative linguistics.

Since the determination principle is multiply ambiguous, given the ambiguity of ‘semantic value’, ‘content’, and ‘determines’, the first task is to disambiguate these readings so that the ones that are most promising in the context of generative linguistics can be evaluated. I argue that on its most plausible readings, the determination principle expresses that a causal determination relation obtains between the assignment of semantic value to an uttered expression by (the semantic faculty of) an interpreter and their interpretation of the utterance. Such readings, in effect, express a psychologically necessary causal link between semantic processing and the process of interpretation.

I will argue, however, that the most plausible readings of the determination principle (in the context of generative linguistics) are false. I will argue, in particular, that cases of “non-literal” interpretation, interpretation of anomalous or ungrammatical speech, and the use of “micro-conventions” all present cases where the requisite causal connection between semantic value and utterance content does not obtain. The upshot is that the relationship between semantic value and utterance content is more distant than is typically supposed. And while semantic theories will bear an important relation to theories of communication and vice versa, the relation is more distant than is assumed—neither theory constrains the other as tightly as endorsement of the determination principle would suggest.

It does not follow from a rejection of the determination principle, however, that our semantic competence does not play an important role in communication. Consequently, I sketch what I think is a plausible account of this role which accommodates both the explanatory role of semantic value in linguistics and the complicated nature of communication, and which is consistent with the rejection of the determination principle.

1 Introduction

The idea that the semantic value of an expression (at least partly) determines the content it is used to express has the status of a guiding methodological principle in semantics, pragmatics, and theories of communication generally. Call this principle “the determination principle.” In this chapter, I argue against the “determination principle”. More specifically, I argue that no version of the principle is suitable as a guiding methodological principle in generative linguistics—each disambiguation is either false, trivial, or else out of place in linguistic theorizing.

Recent attempts at clarifying the relationship between semantic value and content have focused on distinguishing the two, arguing that we should not think of the semantic value of a sentence as being identical to the contents that the expression can be used to express.¹ It is, nevertheless, assumed that there is a tight connection between semantic value and content via commitment to some form of what I will call (following Rabern (2012)) “the determination principle” (DET):²

(DET) The semantic value of an expression e determines the content that uses of e express.

Out of context and without significant elaboration, (DET) is impossible to assess, as it is multiply ambiguous, depending on how one understands ‘semantic value’, ‘expression’, ‘determines’, and what it is to express a particular content. My aim, therefore, will not be to assess the truth of (DET) as it stands. Rather, I will consider the plausibility of various precisifications of (DET) in the context of contemporary linguistics, and in particular, in the generative linguistics framework. I will argue that in that con-

¹See, for instance, Yalcin (2014), Rabern (2012), Ninan (2010), Glanzberg (2014), and less recently, Lewis (1980).

²An exception here might be Yalcin (2014), whose discussion of the semantic value and content roles leave plenty of room to doubt that any determination relation obtains between the two. Commitment to (DET) is otherwise fairly ubiquitous in the philosophy of language.

text, (DET) should be taken to express a psychologically necessary causal connection between the semantic component of the language faculty and the process of utterance interpretation. I will also argue, however, that such a connection does not obtain, and so, in the context of linguistic semantics, there is no plausible disambiguation of the determination principle. Semantic value does not even partly determine content.

The structure of the paper is as follows. In section §2, I will discuss the underlying motivation for the determination principle. In §3, I will sketch the theoretical framework of generative linguistics and clarify how ‘semantic value’ and ‘expression’ are to be understood in that context. In §4, I will discuss various ways to interpret what it is for a use of an expression to express some content. I will argue that, for our purposes, a use of an expression expresses some content in virtue of an interpreter (or a speaker) being in a particular psychological state which constitutes their interpreting an utterance (in the case of a speaker, their own). In §5, I will consider a number of possible readings of ‘determines’. I argue that, in the relevant context, the determination should be read as a causal determination. In §6, I argue that the resulting versions of the determination principle, which are the most promising in the context of linguistic semantics, are, nevertheless, false. In §7, I respond to objections. Finally, in §8, I sketch an account of the relationship between semantic value and content which is consistent with my arguments against the determination principle, and which vindicates the original motivation for the principle.

2 The Appeal of the Determination Principle

The determination principle has the status of a guiding methodological principle in most semantic and pragmatic research programs. Consequently, the motivation for it does not come from linguistic arguments—arguments that rely on the findings of

some linguistic or semantic research program—but rather from prior conceptions of the place of semantics in a broader theory of language and communication, and the conceptual connections between the concepts of meaning, semantic value, content, and communication. I will have much more to say about semantic value and content later in the paper, but it will be useful to make a few very general remarks about their connections first in order to make clear why commitment to some form of the determination principle is so entrenched in most semantic and pragmatic pursuits.

Semantics, in linguistics and in many areas of philosophy, has been thought of as the branch of the study of language that is concerned with meaning, or with speakers’ “semantic competence” or “knowledge of meaning”.³ A semantic theory for a given language, it is supposed, is meant to pair the expressions of that language with its meanings. ‘Meaning’, however, very rarely shows up in semantic theories. It is instead replaced by the more technical ‘semantic value’. Semantic theories pair simple expressions with their semantic values, and if they are compositional, they show how the semantic values of complex expressions can be derived on the basis of the semantic values of the constituents of the complex (and their syntactic arrangement). Semantic values, essentially, play the role of meanings in semantic theories, and so they are very commonly simply identified with meanings.⁴

Contents, broadly, are typically taken to be what are expressed or communicated by our utterances. Contents, and the content-expression relation, therefore, have figured centrally in theories of communication. Given the *a priori* conceptual connections between ordinary conceptions of meaning and communication, it is not implausible

³Davidson (1965) and Dummett (1993b), notably, conceived of semantics as giving a theory of competence—an account of speakers’ knowledge of meaning. This is often how semantics in generative linguistics is characterized as well (e.g., Chomsky (1986), Ludlow (2011), Glanzberg (2014), and Schiffer (2015)), though I think it is less misleading to characterize it as the study of the semantic component of the language faculty. See Soames (1989) for arguments to the effect that semantics should not be thought of as a theory of competence.

⁴The identification of semantic value and meaning, at least informally, is still very common.

to assume a tight connection between semantic value and content. It is a platitude, for instance, that in the ordinary sense of ‘meaning’, the meanings of expressions are crucial to their usefulness in allowing us to achieve our various communicative aims. Understanding a language—knowing what its expressions mean—is tied tightly to being able to communicate with that language.⁵ Schiffer (2015), for instance, writes, “The *raison d’être* of natural language, or at least a huge part of it, is as a device for interpersonal communication, and whatever else understanding a sentence requires, it requires the ability to understand utterances of it, the ability to know what a speaker who utters the sentence means in uttering it” (77).

A natural starting point, then, is to simply identify semantic values and contents. In that case, semantic theories are theories of what is communicated by sentences in literal utterances.⁶ As I mentioned earlier, however, it is now common to distinguish between semantic value and expressed content. One common reason for doing so has to do with the fact that it is pretty universally demanded of semantic theories (and so semantic values) that they be compositional. That is, the semantic values of complexes should be determined solely by the semantic values of their constituents and by the syntactic arrangement of those constituents. However, what is expressed by sentences is not compositionally determined by syntax and the contents of constituent expressions.⁷ Rather than giving up on the compositionality of semantic value or on standard views of expressed content, theorists have opted to distinguish the two. Lewis (1980), for example, writes, “It would be a convenience, nothing more, if we could take

⁵I use the vague “tied tightly to” so as to remain neutral on just what the relation is between understanding a language and being able to communicate using it.

⁶This is similar to how Schiffer (2015, 2003) and Alston (1999), for instance, characterize semantics. The restriction to literal utterances is necessary in order to separate out what is communicated strictly in virtue of semantic processes, and what else might be communicated as the result of various pragmatic and other extra-semantic effects. See my (forthcoming) for an argument to the effect that this way of characterizing semantics is problematic.

⁷Lewis (1980), for instance, argues that the semantics of context-shifting expressions should push us to this view. See also Rabern (2012) and Ninan (2011).

the propositional content of a sentence in context as its semantic value. But we cannot. The propositional contents of sentences do not obey the compositional principle, therefore they are not semantic values” (95).⁸

The relationship between semantic value and content is thus more distant than identity. But theorists have by no means rejected the basic insight that the semantic values of expressions surely play an important role in communication. Lewis goes on to write, “It is enough that the semantic value of a sentence in context should somehow determine the assignment of propositional content” (93-94). This, of course, is another way of stating what I’ve called “the determination principle”, and it is prevalent in some form or other in most semantic research programs. Presumably, the thought is that semantics—and by extension semantic values—are interesting only insofar as they can be understood as playing a role in our theory of communication. It might be unclear just what the interest in semantics could be if it didn’t play such a role. So long as one thinks that the job of semantics is principally to specify the truth-conditions-at-a-context or propositional-contents-at-a-context of sentences in a language, some form of the determination principle seems mandatory if one distinguishes semantic value and content.⁹

I think it is right to assume that semantic value (in the context of generative linguistics) plays an important explanatory role in a theory of communication. However, in the context of generative linguistics, it should not be understood in terms of this role. In that case, we are much freer to wonder about the connection between semantics and theories of communication, and about the relation between semantic value and content. Indeed, as I will argue later on, we have good reason to reject any version of the deter-

⁸See also Glanzberg (2014) and Yalcin (2014).

⁹That the determination principle is so ubiquitous is a reflection of the fact that semantic value is still understood largely in terms of its role in the broader theory of communication. One of the aims of this chapter is to drive home the idea that semantics has an independent domain of explananda that need not be understood in terms of facts about communication. See also Yalcin (2014) on this point.

mination principle in the context of generative linguistics. Again, this is not to deny that semantic value plays an important role in communication, but to acknowledge that the relationship between semantic value and content is more complicated than the determination principle suggests.

3 Background and Methodology

In this section, I want to give a brief sketch of the generative linguistics framework that will serve as our backdrop for disambiguating and assessing the determination principle. I will also clarify how we should understand ‘semantic value’ and ‘expression’ in the context of semantics in generative linguistics.

3.1 Semantics in the Generative Linguistics Framework

Generative linguistics is a branch of cognitive science whose object of study is the human language faculty—that part of our brains which is responsible for much of what we might informally call our “linguistic competence”.¹⁰ Having a particular language, in the relevant sense, is not a matter of being a member of a particular speech community, but of being in a particular mental state. Language is studied, then, as an individual, psychological object, not a public object. I will use ‘generative linguistics’ here quite broadly, so as to be neutral on debates concerning, for instance, the extent to which linguistic competence is innate, or the correct way to conceive of the relationship between syntax and semantics. For our purposes, what matters is just that generative linguistics is a broadly psychological research program which studies a particular faculty of the human mind.

¹⁰See Chomsky (1965) for a classic elaboration of the generative linguistics framework. See also Chomsky (1986) and Ludlow (2011) for much fuller presentations than I can provide here.

Semantics, then, is the branch of generative linguistics which studies the “semantic” component of the language faculty—call it the “semantic faculty”—which underlies some portion of our meaning-specific linguistic abilities. These abilities include our abilities to recognize (certain) entailments, synonymy, polysemy, ambiguity, anomaly, truth-value, implicatures, what we mean by each others’ utterances, and so on. I say that the semantic faculty underlies “some portion” of these abilities because, while they all might plausibly be thought of as being pre-theoretically semantic phenomena—as having to do with “meaning” in the ordinary sense—there’s no guarantee that all of these phenomena fall neatly into a single linguistic kind, requiring the same sorts of explanations.¹¹ Indeed, we already recognize that the ability to recognize implicature, for instance, is at least partly extra-semantic, in that it requires knowledge of various conversational norms, and general world knowledge which is plausibly distinguished from linguistic competence. As is usual in scientific investigation generally, just what ends up being the proper domain of semantics is an *a posteriori* matter.¹²

It is worth stressing that semantics in generative linguistics is not to be seen merely as a component of a broader theory of communication. Semantics, here, has some autonomy, with its own subject matter and explananda, distinct from a theory of communication. It is not characterized, as Lewis (1980) characterizes it, as “part of a systematic restatement of our common knowledge about our practices of linguistic communication” (79). Semantics still is continuous with and thus constrained by other linguistic theories—syntax and pragmatics, for instance—but it is not understood in terms of those theories. Consequently, the determination principle will not have the *a priori* motivation that it enjoys in a semantic program like Lewis’.

¹¹See Yalcin (2014) for further elaboration on this point, and on the theoretical role that semantic value occupies in semantics.

¹²As Chomsky (1995, 2000) has emphasized, philosophers have often been guilty of failing to recognize this point in the context of linguistics. See also Johnson (2007).

3.2 ‘Semantic Value’

Having given a rough sketch of the semantic research program that will be in the background of our investigation of the determination principle, we can now say more about how we are to understand ‘semantic value’ and ‘expression’ as they occur in the determination principle. Like Yalcin (2014), I will understand ‘semantic value’ functionally, in terms of its rather broad and as yet vague theoretical role that I will sketch momentarily. Taking this approach will allow us to talk about semantics similarly broadly, so that we can maintain some (though not complete) neutrality with respect to various proposals for the realizers of the semantic value role.

‘Semantic value’, broadly, is a theoretical term denoting a property or set of properties of expressions which play a central role in explanations of semantic phenomena. Almost universally, they are taken to obey some version of the principle of compositionality. That is, the semantic values of complex expressions should be determined on the basis of the semantic values of their constituent expressions, the syntactic arrangement of those constituents, and some rules of composition.¹³ Just what the explanations of semantic phenomena look like in detail and what the explananda will be are, again, *a posteriori* matters which will be settled as we progress in our semantic investigations. However, since the object of inquiry is a psychological mechanism which, we suppose, helps to explain some number of our semantic abilities, we know that semantic values, whatever they are, should play a role in fairly proximate causal explanations of these abilities.¹⁴ We also know that an expression of a speaker (more on ‘expression’ shortly) will have its semantic value in virtue of the speaker being in a particular mental state.

¹³For discussion of the various disambiguations of the principle of compositionality and its place in linguistics, see Szabó (2000, 2012a, 2012b), Pagin & Westerståhl (2010a, 2010b), and Chapter 4 of this dissertation.

¹⁴Linguistic judgments, which serve as a primary source of data for linguistic theories, are explained slightly more distally, since they are manifestations of (or at least caused by) our linguistic competence, which is the primary subject matter of linguistics.

If a speaker's expression had such-and-so semantic value in virtue of some other sort of state of affairs obtaining, then attributing a particular semantic value to an expression would not be a matter of attributing any property to the semantic faculty, which is our object of study.

We cannot, therefore, be completely neutral with respect to the potential realizers of the semantic value role. For instance, without doing any linguistics, we know that an expression does not have its semantic value in virtue of its bearing some relation to an extra-mental reality. 'Dog' does not have its semantic value in virtue of bearing some relation to dogs, nor does 'Dogs bark' have its semantic value in virtue of the fact that it is true if and only if dogs bark (where 'in virtue of' is to be read as denoting a constitutive relation).¹⁵ I will have more to say about truth-conditional semantics shortly, but I should be clear that I do not intend to single it out. An expression's having its semantic value will also not be a matter of there being certain public conventions governing its use, or there being certain patterns of use for that expression that obtain in some wider linguistic community. Now, it could be that an expression's bearing a reference relation to some extra-mental objects or its being governed by certain public conventions figure *distally* in causal explanations of our linguistic abilities. But the semantic faculty is meant to figure in *proximate* causal explanations of our semantic abilities. Facts about reference and public conventions might play an important explanatory role in our final theory of language, but they would likely do so by figuring in explanations of why the states of the language faculty of some group are the way they are, not because they accurately describe the language faculty.

Now, I want to pause here and clarify why the foregoing discussion of semantic value should not be taken to have particularly controversial consequences for either

¹⁵See Chomsky (2000, 180-181) and Pietroski (2003) for arguments against the prospects of constructing a semantic theory in terms of reference.

linguists or philosophers. For linguists doing semantics, the concern might be that, given what I have said, truth-conditional semantics—plausibly the most popular and successful research program within linguistic semantics—must be false. Worse, I said that we can know this even without doing any semantic theorizing.

But truth-conditional semantics as it is pursued in generative linguistics is not plausibly interpreted as being primarily concerned with describing how words relate to the extra-mental world. Rather, the central hypothesis of truth-conditional semantics in this context is that the semantic faculty assigns truth-conditional contributions to simple expressions and compositionally derives the truth-conditions of sentences. The assignment of truth-conditions by the semantic faculty figures in explanations of judgments of entailment, truth-value, anomaly, and so on.¹⁶ This “internalistic” reading of truth-conditional semantics is, generally, the way to interpret truth-conditional work in generative linguistics.

If this is not obvious, consider the standard methodology in the relevant truth-conditional literature. One of the primary sources of evidence for truth-conditional semanticists are speakers’ linguistic judgments (including the theorist’s judgments when the language under investigation is their native language) concerning entailment, truth-value, contradiction, and so on. These judgments serve as a good source of evidence for their theories because it is hypothesized that the semantic faculty, which is the object of their investigation, is partly causally responsible for these judgments.¹⁷ The methodology is roughly the same as it is in syntactic theorizing, where the commitments of the generative linguistics framework tend to be more explicit. Theorists consider speakers’ (including their own) judgments of ungrammaticality or unacceptability and hypothe-

¹⁶Larson & Segal (1995) are explicit that this is how their truth-conditional semantics is to be interpreted. Heim & Kratzer (1998) are less explicit, but their methodology strongly suggests that they are to be interpreted this way as well.

¹⁷Other causal factors might be the syntactic faculty, “performance” factors like memory and attention, and various other extra-semantic mechanisms which go into the production of a linguistic judgment.

size about the nature of the syntactic component of the language faculty, which partly causally explains these judgments. Given this overlap in methodology, it would be surprising if the two sub-fields were studying entirely different objects.

Now, it is true that it is far more common for syntacticians to avail themselves of facts about language acquisition and aphasia, facts about other languages which shed light on the possible states of the language faculty, findings from nearby areas in cognitive science, and so on.¹⁸ The reason for this, I think, is that while syntax is still quite a young pursuit, semantics in generative linguistics is still in much more of an exploratory phase in its development, and so it is less clear than in syntax how it integrates with other areas of linguistics and psychology. In any case, it is also important to remember that syntax and semantics are thought to be continuous, and constrain one another rather tightly. This stance would not be justified if it turned out that syntacticians and semanticists were studying very different sorts of objects.

Finally, I will just note that it could very well be that there are perfectly good ways to study linguistic meaning that don't fit within the generative linguistics framework I've sketched. Nothing I've said about semantic value in generative linguistics would tell against these other semantic research programs. My claims only concern the realizers of the semantic value role in the generative linguistics framework.¹⁹

I don't think, then, that my characterization of the semantic value role should be taken to have any controversial consequences for linguistics. The foregoing remarks should also assuage concerns that I am impugning truth-conditional semantics generally, as it might be pursued in logic or in the philosophy of language. Insofar as these programs are distinct from the generative linguistics program I am concerned with, my

¹⁸Compare, for instance, a standard truth-conditional semantics textbook in Heim & Kratzer (1998) or von Stechow & Heim (2011) to a standard textbook in syntax in Adger (2003). It is not necessarily uncommon for semanticists to appeal to facts about other languages or acquisition in their work, but it is far less common than in syntax.

¹⁹In fact, it may turn out that there isn't a single semantic value role but several distinct roles which are geared towards explaining different subsets of the pretheoretically semantic phenomena.

remarks do not bear on them.

There is, however, one more lingering philosophical concern that is worth addressing, and that is the relationship between semantic value in generative linguistics and meaning. More specifically, the concern might be that if we are naturalists about meaning, and semantics in generative linguistics represents our most successful and promising semantic research program, then we ought to identify *having such-and-so meaning* with *having such-and-so semantic value*. But given what I've said about semantic value, it would follow that speakers' expressions mean what they do in virtue of their language faculty being in a particular state, and this is a controversial thesis. Perhaps worse, meaning would not attach to objects in a public language, but rather to mental entities that would be a part of one's language-at-a-time. This, of course, would run counter to many of the most popular philosophical theories of meaning and language, and so my sketch of the semantic value role would have serious consequences indeed.

This line of worry can be resisted at several steps. First, while it is common for theorists writing about semantics in generative linguistics associate semantic values with meanings, we are not compelled to do so. It is true that semantic theories have to do with "meaning" in the ordinary sense in that folk explanations of semantic explananda would typically appeal to meanings. However, this only tells us that it is useful to informally characterize these semantic theories as being theories of meaning. There is nothing which forces us to identify "meaning" in the ordinary sense with "semantic value" in the sense we are concerned with here. We may do so, of course, but the decision to do so is an extra, philosophical decision which is not forced on us no matter how successful the generative semantic research program turns out to be.

Second, as I noted above, there are surely many different ways to study language and linguistic meaning which do not focus on the language faculty. Many linguists

(sociolinguists, for instance) and philosophers study language as a publicly shared entity. Some of these research programs are sure to be very fruitful and successful, and will likely offer explanations of community-wide speech patterns by appealing to some meaning-like relation. If one feels pressure to locate some ordinary notion of meaning in a fruitful scientific research program but doesn't think that meaning in that sense is internalistic, it might be that one of these other projects is the better place to provide a home for meaning. This is the case even if semantics in the generative linguistics research program is currently the most well-developed extant semantic research program.

3.3 'Expression'

Let me now clarify how we should understand 'expression' in the determination principle, given our focus on generative linguistics. I have been speaking loosely about languages and expressions thus far. We can think of languages as abstract objects which consist of lexical items (which bundle together types of phonological properties with syntactic and semantic ones), phonological rules, syntactic rules, semantic rules, and whatever other elements of human language might be discovered in linguistic theorizing. Despite the fact that languages are abstract objects, generative linguistics is a psychological pursuit because it studies those languages which can be and are learned by humans as first languages.²⁰ Since speakers have their languages in virtue of their languages faculties being in particular states, human languages are studied via the study of the language faculty.

The standard approach in the philosophy of language is to think of expression tokens as being sounds or marks that have syntactic and semantic properties in a

²⁰See Chomsky (1986), Ludlow (2011), Collins (2004), and Alexander (1989) for more on the ontology of generative linguistics.

language.²¹ In the context of generative linguistics, however, expressions are mental entities which are assigned syntactic and semantic properties via the lexical entries that compose the lexicon. Since lexical entries are bundles of phonological, syntactic, and semantic features, it is natural to think of expressions as phonological entities—strings of phonemes. Phonemes, roughly, are composed of features which, together, constitute instructions to the articulatory system to make particular kinds of sounds.²² Thus, in this context, expressions are not sounds themselves. Expression tokens, then, can be thought of as the (usually complex) phonological—and hence psychological—entities that are tokens of phoneme-string-types that appear in the lexicon. The details will largely remain unimportant. What matters most is just that expression tokens are psychological entities that occur in individual language faculties. These are the bearers of semantic value in generative linguistics.

4 ‘Expressing content’

In this section, I will clarify how we should understand content-expression in the determination principle. I will first sketch, very broadly, the theoretical role that content is usually taken to occupy. I will then consider various proposals of what it is for uses of an expression to express to express such-and-so contents. Finally, I will argue that on the only plausible and interesting readings of the determination principle in the context of generative linguistics, a use of an expression expresses some content in virtue of the fact that a speaker arrives at an interpretation of that expression.

²¹See, e.g., Lewis (1975).

²²See, e.g., Kenstowicz (1994) for a textbook treatment of phonology in generative linguistics.

4.1 The Content Role, Broadly Speaking

The notion of content figures centrally in theories of mind and communication, around which any number of philosophical issues orbit. Contents are thought to be “what is said” or “what is expressed” by a speaker in making an assertion—they are the “objects of assertion”. Often, communication between speakers is thought of as being successful (at least along one dimension) when a communicatee recognizes an utterance (for example) as expressing such-and-so content, which content the communicator intended to express by making the utterance. That we can reliably produce, recognize, and coordinate content in this way is said to be essential for our ability to coordinate action with other members of the community. Contents thus figure in explanations of how our social practices function, and why they function the way they do. Contents also provide a theoretically significant dimension along which we can distinguish between linguistic acts of the same force. Similarly, they provide a theoretically significant dimension along which we can distinguish mental states involving the same propositional attitudes. This allows us to appeal to an agent’s beliefs, for example, in partial explanations of their behavior. For instance, Mary’s eating the apple is partly explained by her belief that the apple is nutritious—had she believed that it was poisonous, she presumably would have behaved differently.

For our purposes, it will not matter what kind of objects contents are. We can think of them, generally, as being propositions, and remain neutral on their nature—whether they are coarse-grained or fine-grained, whether they are to be understood as truth-conditions or inferential roles, and so on. The focus in this chapter will be on the content-expression relation—what it is to express such-and-so content. The reason for this focus is just that precisifying the content-expression relation will do more to narrow down various content roles than focusing on contents themselves. This should become clearer as we continue.

So far, I have only intended to offer a quick reminder of what sorts of theories and explanations contents are invoked in. None of what I said so far suffices to pick out a particular theoretical role for ‘expressed content’. And indeed, I think that the variety of realizers for that role that have been proposed suggest that it is a mistake to think there is a single content role to be realized. As there are a number of ways to study language, so too are there a number of ways to study communication, and there are various notions of expressed content which fulfill roles in these different sorts of research programs. Our task is to find the notions for which there might be a determination relation between semantic values and expressed contents.

4.2 Externalist Notions of Content-expression

As was the case with the semantic value role, we can rule out a lot of notions of content-expression as yielding plausible readings of the determination principle without doing any linguistics. We can do this merely by reflecting on whether or not they could plausibly play roles in a theory which would be continuous with semantics in generative linguistics such that semantic value might determine content in the relevant sense.

The first class of content-expression-notions that we can eliminate are so-called “externalist” notions. On these views, sentences express their contents in virtue of the sentence tokens standing in some relation to extra-mental reality. The problem with these views, generally, is that states of the semantic faculty and the relevant extra-mental states of affairs—even if they are not completely independent of one another—are insulated enough from one another that it’s implausible that any sort of determination relation obtains between them.

At best, the extra-mental states of affairs might be part of an explanation of why states of the semantic faculty are the way they are, but if one takes the determination relation to be an explanatory one, then the determination principle would get the

explanation backwards. In other words, facts about the contents of expressions might determine (in some sense) facts about their semantic values, but not vice versa.

I have not yet said much about determination (much more on this in §5), but we can just mention some of the relations that are variously considered to be determination relations for now. First, there is what we might call ‘metaphysical determination’—constitution or partial constitution. Then there is ‘epistemic determination’. Semantic value would determine content, in this sense, if hearers come to grasp the content of an uttered expression on the basis of their grasping the semantic value of the uttered expression. Determination might also be causal—an expression expressing the content it does might be causally dependent on its having its semantic value. Finally, semantic value might determine content in the weak sense that there is a function from the semantic value of an expression and features of context to the content it expresses. Without specifying some set of contextual features, this principle will be trivial, since presumably all contexts differ with respect to some of their features.²³ Consequently, we should think of the resulting principle as claiming that there is a function from the semantic value of expressions and some specifiable list of contextual features of utterances to the content expressed by those utterances.

My claim, then, is that semantic value does not determine, in any non-trivial sense, content externalistically understood. Suppose, for instance, that you have a causal view of reference for names or natural kind terms, and so think that the contents expressed by those terms (or the contents they contribute to sentences involving them) depends on some causal relationship that they bear to their referents.²⁴ In that case, it is obvious that an expression’s having such-and-such semantic value does not constitute or even partly constitute an expression’s expressing a particular content. What constitutes its expressing that content, again, is that tokens of that expression stand in a particular

²³I follow Lewis (1980) here.

²⁴See, for instance, Putnam (1975) or Kripke (1980).

causal relation to its referent. Facts about speakers' mental states are not even partly constitutive of content-expression facts.

Nor are speakers in a position to grasp that natural kind terms, for instance, express such-and-so content on the basis of their language faculty assigning to these terms a semantic value. If knowing what content some expression expresses requires knowing the causal history of tokens of the expression, such a grasp might only be possible after extensive scientific investigation. However, one might hold that a speaker can know what their expression refers to without knowing what constitutes its referring as it does.²⁵ In that case, some version of the resulting determination principle might be true (though it's not clear there's any reason to think that this is likely). I will postpone more general argumentation against epistemic determination principles until §5.2.

More obvious is the fact that a competent English speaker's language faculty assigning to 'water' some semantic value does not even partly causally explain why tokens of 'water' have the causal history they have.²⁶ Nor need there be any function from the semantic values of expressions to the contents they express, externalistically understood, in an utterance.

Similar remarks can be made about other externalist notions of content-expression. A teleosemanticist might hold, for instance, that an expression token expresses the contents it does in virtue of facts about previous tokens of that same expression type which explain why those tokens have been reproduced.²⁷ But it doesn't seem plausible that facts about the semantic faculties of present-day speakers could determine, in

²⁵Burge (1988), for instance, holds that one can know the contents of their own thoughts without knowing that the conditions which constitute the thought's having its particular contents obtain. See also Davidson (1984).

²⁶Facts about the mental states of the initial speaker (or speakers) in the initial "dubbing" event will be relevant to explaining why 'water' expresses the content (in the relevant sense) it does, but this does not hold for the rest of the speakers who are competent with 'water', and who are the targets of linguistic theorizing.

²⁷See, e.g., Millikan (1984) and Dretske (1981).

any sense, the kinds of facts that ground content-expression for a teleosemanticist. And likewise for other externalist theories. There doesn't seem to be any explanation running from semantic value to content. At best, there might be rather distal causal explanations running from content to semantic value.

To be clear, none of what I've said tells against externalist notions of semantic content. What it tells against is simply the idea that content understood that way bears a tight relationship to semantic value as it is understood in generative linguistics. In other words, in the context of generative linguistics, we shouldn't understand content-expression in the determination principle in an externalist way.

4.3 What Is Expressed as Determined by Convention or Norm

I think we can similarly ignore notions of content-expression on which expression tokens express some content or other in virtue of public conventions or norms which govern their use. I have in mind here Lewis' (1975) account, on which expressions in a public language have the meanings they have in virtue of certain conventions of trust and truthfulness obtaining in the population (7). The obtaining of these conventions, in turn, depend on the obtaining of community-wide regularities concerning the production of sentence-tokens. Meanings, together with context, determine what content is expressed by an utterance.

What is the relation between these content-expression facts, which depend on broad patterns of behavior obtaining in a particular speech-community, and semantic value as we have understood it in the context of generative linguistics? It is difficult to know without knowing more about the role of the semantic faculty (and the language faculty more generally) in the production of the kinds of behaviors which constitute the obtaining of a convention in Lewis' sense. As I have mentioned, the semantic faculty is taken to be partly directly responsible for linguistic judgments of anomaly, entailment,

contradiction, and so on. It is not taken to be directly responsible for dispositions to speak truly, or expectations that others will speak truly, for instance.

One immediate problem with any determination principle in which we take content-expression to depend on community-wide conventions is that conventions are meant to govern public language expressions—types of sounds and marks—not expressions in the sense that I sketched in §3.3. The determination principle, as it is typically stated (as in (DET)), would need to be complicated, since it doesn't allow for this ambiguity. Presumably, the idea would be that the conventions that govern public language expressions somehow depend on the semantic faculties of the members of the linguistic community being in particular states. The dependence, presumably, is not metaphysical, since the obtaining of conventions does not consist (even partly) in speakers' language faculties being in particular states. Most likely, the dependence would be causal, so that the assignment of semantic values to phonological entities is a causal determinant of public language expressions being governed by particular conventions. More specifically: the assigning of a particular semantic value m to a phonological entity e causally determines the obtaining of a convention C which constitutes a mark or sound-type E having content M , where, presumably, tokenings of E cause tokenings of e .²⁸

There are a couple of questions here. First, does the obtaining of Lewisian conventions, generally, depend on speakers having something like a language faculty as posited by generative linguistics? I think there is good reason to doubt this. True, the obtaining of those conventions would surely require a great deal of psychological sophistication, and perhaps even some kind of internal representation of language. But it would seem as though internal representations of language that differ from the sort

²⁸The relation between sound or mark-types and the relevant phonological constructions is massively oversimplified here for the sake of discussion.

posited in generative linguistics could support those conventions as well.²⁹ In that case, semantic value-facts do not causally determine the relevant content-facts.

But our focus is not on the connection between the language faculty and Lewisian conventions generally, but on the connection between semantic value and content-expression in particular. And so the relevant question, for our purposes, is whether for every phonological entity-public expression pair e - E (where tokenings of E cause tokenings of e), the convention which constitutes E 's having content M would not obtain unless e has semantic value m (for the speakers that make up the relevant community). This, presumably, is what it would be for the semantic values of expressions (in the internal sense), generally, to causally determine the content (as fixed by convention) of public language expressions. Without knowing more about the relationship between the semantic faculty and behavioral patterns which help fix conventions, it is difficult to be certain about the proposal. But we can at least say that given the large gap between the explananda of semantics in generative linguistics and the facts that constitute the obtaining of a Lewisian convention, it seems very unlikely. At least, we have no good reason to think that such a relation will obtain, and it seems unreasonable as a default assumption.

None of this is to deny that facts about the language faculty play an important role in shaping our linguistic conventions. They surely do, and acknowledging this is consistent with denying the relevant determination principle. The determination principle is meant to ensure a close connection between semantics and a theory of communication, such that each very tightly constrains the other. Indeed, each lexical entry posited by semantics would be constrained by the conventions that govern the

²⁹Yalcin (2014) writes, "I doubt that an agent which had the behavioral dispositions characteristic of the Lewisian linguistic conventions of some population of speakers must thereby be in a language-cognizing state of the sort [that is posited by generative linguistics]" (41). Lewis (1975, 22) agrees that a population's having a language does not *metaphysically* depend on their internally representing a grammar for that language, though he is open to the hypothesis that having a language does causally depend on having an internal representations of of the language.

relevant public language expression, and vice versa. If we understand semantics in the context of generative linguistics, however, and the theory of communication is one on which what is communicated is fixed by convention, then one ought to reject the idea that there is such a tight connection.

Indeed, without going through each case, I think we have good reason to ignore accounts of content-expression on which the content expressed by an expression (in context) depends on community-wide patterns of behavior, or norms which explain those behaviors. We can ignore, for instance, the inferentialism of Sellars (e.g. 1951, 1969) or Brandom (1994), or the “social externalism” of Burge (1979). Indeed, given the rather large explanatory gap between semantics in generative linguistics, and facts about the broad patterns of behaviors of speech communities, I think we can put aside any version of the determination principle on which content-expression has public language expressions as one of its relata. It is possible that some of these versions of the determination principle are *true*, but none of them would be suitable as a guiding methodological principle in linguistics, as the determination principle is meant to be.

4.4 Internalist Notions of Content

What we need, then, in order to eliminate the explanatory gap between semantic value and content, is an internalist notion of content-expression. On this sort of view, an expression expresses its content in virtue of facts about a speaker’s mental state. Consequently, content-expression-facts, internalistically conceived this way, are not meant to directly explain community-wide speech patterns or facts about the proliferation of expression-tokens in a community, but rather facts about the behaviors of individual speakers. What we get is a theory of communication which seeks to explain speakers’ individual communicative behaviors. Such a theory might seek to explain, for instance, why Mary rejected John’s claim, why she came to have certain beliefs after John made

his utterance, or why she found John's utterance puzzling or infelicitous. By positing that utterances are associated with particular contents, the theory can contribute to explanations of these kinds of facts.

This sort of theory will have a much closer connection to semantics in generative linguistics than theories which rest on social or externalist views of content-expression. In fact, construction of such a theory is arguably the job of pragmatics, which is taken to be in direct contact with semantics. Semantics, and the attribution of semantic values to phonological entities, helps to causally explain speaker judgments of anomaly, entailment, and the like. The attribution of content to utterances figures in causal explanations of speaker judgments of conversational infelicity, presupposition failure, and perhaps the ability of speakers to co-ordinate their behaviors in some complicated task. There is at least hope, then, that an internalist reading of content-expression might vindicate some version of the determination principle.

Indeed, a more general argument can be brought to bear against versions of the determination principle that depend on externalist or social views of content-expression like those considered in §5.1 and §5.2. The argument is just that, given the sort of theory of communication we have in mind—one that gives proximate causal explanations of the communicative behaviors of individuals—the relevant notion of content-expression employed in such a theory ought to be determined by the mental states of individuals. Externalist and social notions of content-expression need not be explanatory *irrelevant*, but the explanations of communicative behavior that they would figure in are simply more distal, and thus not the kinds of explanations that a theory contiguous with generative semantics will be trying to give. Consequently, even if some externalist or social reading of the determination principle turned out to be true, it would still not be useful as a guiding methodological principle in generative linguistics.³⁰

³⁰Similar remarks about the explanatory aims of generative linguistics and non-relative, external or public notions of content, can be found in Chomsky (2000, Ch.2).

How, then, should we think of content-expression, more specifically, in the kind of theory under discussion here? In other words, in virtue of what does an utterance express the contents it expresses? If facts about the contents of utterances are meant to help explain speakers' reactions to utterances they encounter, or to help explain why they make the utterances they make, what matters is how a speaker interprets utterances (including their own). Content-expression will not be understood in terms of any notion of what is expressed, or what is said, beyond what each interlocutor takes the speaker to be saying by their utterance. Or at least, no such notion will be necessary for giving the proximate, causal explanations of individual speaker behavior that are intended, since such notions will supervene on more than just the mental states of the individual speakers.

It will be worth slowing down here to clarify the relationship between two notions of locutionary content that will be relevant to the theory. The first is the one which will play the content-role in the theory of communication of the sort I've just described. It will, consequently, figure directly in the relevant theoretical explanations of individuals' communicative behaviors. It will also be relativized to individual interlocutors, so that the content expressed by a single utterance will vary from individual to individual. After all, content-expression is meant to depend simply on how individuals interpret a particular utterance. There is no sense in asking what content an utterance expressed simpliciter in the relevant sense of content-expression. This is the theoretical notion of utterance content or content-expression, and since my focus is on the linguistic theories of communication, I will simply use 'utterance content' and 'content-expression' to pick out these notions unless otherwise specified. Likewise, I will use 'interpretation' to pick out whatever theoretical relation it is that speakers stand in to utterances that explains the relevant communicative behavior.

But there is also a folk notion of locutionary content, or 'what is said' that might

have relevance to the kind of theory of communication under discussion. I suggested that we understand content-expression as being determined by what interlocutors take an utterance to express or to say. But the relevant notion of taking an utterance to express or say such-and-so is a folk notion, and it is one that we employ in folk explanations of various sorts. Taking (or interpreting), I will presume for convenience, is akin to believing, so that to take an utterance to express such-and-so is to believe that it does so. However, when an interlocutor believes of an utterance that it expresses some content, they are not employing the theoretical notion of content-expression that I sketched above. Rather, they are employing a folk-notion of content-expression. Folk-content-expression—assuming for the sake of convenience that there is only one such relation—is unlike the theoretical kind of content-expression in that it is not plausibly relativized to individual interlocutors. Rather, it seems that, much of the time (though not always), the content that is folk-expressed by an utterance is a matter of the speaker’s intentions, perhaps in conjunction with various conventions governing the relevant utterances. Consequently, there *are* non-relativized facts of the matter about what is folk-expressed by an utterance.

My proposal, then, is that theoretical content-expression can be understood in terms of something like the state of taking an utterance to folk-express such-and-such. In that case, the non-relativized, folk notion of content-expression will play a role in explanations in the theory of communication I’m imagining as well. However, it wouldn’t figure as directly in theoretical explanations as the theoretical notion of content-expression. The notion of folk-content-expression, itself, is not a theoretical concept designed to play a particular explanatory role, as that of content-expression is. Rather, it makes its way into the theory because the theory appeals to facts about speakers’ mental states in explanation—facts about what speakers take an utterance to folk-express. Facts about what is folk-expressed will not appear in the theory except

when embedded in a “taking” or “interpreting” context.

I should be clear that it would be methodologically inappropriate to make an a priori identification of facts about content-expression in the theoretical sense with states of taking an utterance to folk-express such-and-so. It seems to me that the latter states would be more or less suitable in filling the explanatory role carved out for utterance-content. However, this is nothing more than a first guess, and I only mean to informally gesture at the kind of state that might be able to play the relevant role. In the end, facts about content-expression may end up differing from states of taking utterances to folk-express such-and-so in subtle ways. They will, nevertheless, supervene solely on the states of individuals, and thus be relative to individuals.

I should also stress that the idea that content-expression, in the current theoretical context, is relative to an interpreter, should not be a controversial thesis. It is a platitude that two interlocutors might interpret an utterance differently. If we are trying to explain what each does in response to that utterance, what will matter is how each interprets it (and perhaps how each takes the other to have interpreted it, and so on). Further questions about which interpretation is the correct one are completely irrelevant to the theorist, and would be akin to a syntactician asking which reading of an ambiguous sentence is the correct one. We might also appeal to a notion of common ground along the lines of Stalnaker (2014) in order to explain various phenomena having to do with conversational kinematics. Once again, however, we will have to recast the notion of common ground so that it is relativized to interlocutors. What matters is what each speaker *takes* there to be in common ground. These kinds of facts will do better at explaining their conversational behavior than a common ground which results from taking the intersection of some portion of their mental states.

It should also be noted that the bearers of contents will be *utterances* of expressions, not expressions, or expressions-in-context. Because what matters for our explanations

of communicative behavior are hearers' interpretations of utterances, and because these interpretations often depend on pragmatic processing involving, for instance, beliefs about the speaker, an utterance might be interpreted in any number of ways by different interpreters in a single context.³¹ It is difficult to see what adding an additional notion of what is said or expressed, which attaches to sentences or sentences-at-a-context, would add to the explanations of communicative behavior.³²

The internalistic notion of content I have been sketching is still quite vague. In particular, I have not said much about the mental state that constitutes "interpreting an utterance". I have not said, for instance, whether the interpretation is a matter of attributing particular truth-conditional content to it, of taking oneself to be able to make certain inferences from it, and so on. While I have given a broad sketch of the theoretical role of an internalistic notion of content, and while I have put some constraints on the content-expression relation, determining just what that constitutes that relation is the job of the empirical theory of communication which I have been alluding to.

We can say a bit more about the mental states that constitute utterance-interpretation in the relevant sense, however. Namely, they differ from the mental states that constitute the assignment of semantic value to an expression by one's semantic faculty in that the contents of the former states are far more open to introspection. Much of what is called our 'linguistic competence' is what linguistic theories are attempting to describe, and is completely inaccessible to introspection. Our folk explanations of why we judge a string to be ungrammatical, for instance, will not resemble anything like the explanation given by our syntactic theories. We can seemingly introspect on how we

³¹If we think of contexts in such a fine-grained way that we include all facts that might bear on the interpretation of an utterance, then 'expression-at-a-context' and 'utterance' will be practically equivalent.

³²That is, unless we just think of an expression as expressing some content just in case utterances of that expression express those contents. This sort of proposal will be taken up and rejected in §6.

interpret utterances, however, and the folk explanations we give of our communicative behaviors that appeal to facts about how we interpret utterances might not depart too much from the explanations the theorist might give.³³

The sketch of the internalist notion of content-expression can otherwise remain vague. What matters, for our purposes, is to have a general sense of what kind of relation that will be, and to have a similarly general sense of the kinds of explanations that facts about the contents of utterances will figure into.

5 Determination

The most plausible versions of the determination principle in the context of generative linguistics express, roughly, the idea that there is a tight relationship between semantic processing (narrowly construed as the processing that the semantic faculty is responsible for) and the process of interpreting what an utterance expresses. More specifically, the idea is that the assignment of semantic value to an expression by the semantic faculty of a hearer determines (in a sense to be elaborated) their interpretation of that expression—what they take the utterance of the expression to express.

We can see, then, how the determination principle might tie semantics closely to a theory of communication, such that each heavily constrains the other. The principle of compositionality states that the semantic values of complex expressions can be derived from the semantic values of their constituents, their syntactic arrangement, and a handful of composition rules. As a guiding methodological principle, it constrains the theorist's assignments of both semantic value and syntactic features to both simple and complex expressions. The determination principle entails that a determination relation obtains between the semantic value of an uttered sentence and the content

³³Schiffer (2015), for instance, thinks there is a very important connection between ordinary indirect discourse speech, and theoretical explanations of speaker behavior.

assigned to that utterance by a hearer. As a guiding methodological principle, then, the determination principle constrains assignments of both semantic value and content by the theorist. Given the determination principle and the principle of compositionality, the theorist's assignment of content to an utterance will constrain the assignment of semantic value to simple expressions as well, and vice versa.

As was mentioned in §4, however, there are a number of relations that are considered determination relations, and it's not clear which sense is intended in the determination principle. In this section, I will first argue against versions of the principle that result from thinking about the relevant determination as being metaphysical or epistemic. I will go on to argue that on the most plausible disambiguations of the determination principle, the relevant determination is causal.³⁴ Finally, I will specify what I take to be two of the most promising precisifications of the determination principle.

5.1 Metaphysical Determination

The first sense in which semantic value might determine content is metaphysical—the assignment of semantic value to an uttered expression might constitute or partly constitute the assignment of an interpretation of the relevant utterance.

We can reject the idea that the assignment of semantic value to an uttered expression fully constitutes an interpretation of the relevant utterance on several grounds. First, the content expressed by an utterance is typically taken to depend at least on the context of the utterance, in addition to the uttered expression's semantic value.

³⁴I will ignore the functional reading of 'determines' the rest of the way. Again, there is a worry of triviality unless one specifies the particular features of contexts that work together with semantic value to determine content. If one does specify such features, then given that the determination principle is meant to be stating a relationship between semantic processing and the process of interpretation, it would be very surprising if a functional determination relation obtained, and a corresponding causal one did not. I will be assuming, going forward then, that if a functional reading obtains then some causal reading does. In any case, some of the objections I raise to the causal readings will apply just as well to a functional one.

Second, while the semantic value of complex expressions is determined compositionally, the contents they express is not. Finally, we can just note the differences in theoretical role of semantic values and contents. Semantic values, again, are posited to explain anomaly, contradiction, entailment, truth-value, and so on, and speakers' judgments of acceptability, truth-value, etc. Contents are posited to explain, broadly, more communicative and conversational behaviors—the denial of some assertion, presupposition failure, judgments of conversational infelicity, and so on. While it is possible that the semantic value and expressed content theoretical roles could have the same realizers (or realizers which stand in a constitution relation), we ought not simply assume so, and we have good grounds for doubting that this will be the case.³⁵

The assignment of semantic value to an uttered expression might still *partly* constitute an interpretation of the relevant utterance, however. In other words, (DET_{MP}) might be true:

(DET_{MP}) An uttered sentence *S*'s being assigned semantic value *m* by one's semantic faculty partly constitutes their interpreting the utterance of *S* as expressing content *c*.

Since having a semantic value and expressing a content, in this context, are just a matter of a hearer (or speaker) or speaker being in the right kind of mental state, we can identify an uttered sentence's having its semantic value with the uttered sentence being assigned a semantic value by one's semantic faculty. Likewise, we can identify an utterance of a sentence expressing content *c* with a hearer or speaker's interpreting the utterance as expressing *c*.

The arguments given above do not decisively tell against (DET_{MP}), but reflection on the differing theoretical roles of semantic value and content should give us reason

³⁵See also Yalcin (2014) on this point.

to be skeptical about endorsing it. For now, though, I will leave it on the table. More decisive arguments against it will be given in §6, where I will provide reason for thinking that an utterance can express a content even if the interpreter assigns no semantic value to the uttered sentence.

5.2 Epistemic Determination

Sometimes, determination is understood as an epistemic relation. For instance, sometimes the principle of compositionality is understood as the idea that speakers grasp the semantic value of complexes on the basis of grasping the semantic values of the complex's constituents and their syntactic arrangement. With respect to the determination principle, then, the thought would be that a speaker grasps the content of an utterance on the basis of grasping the semantic value of the uttered sentence. And importantly, grasp of semantic value is meant to play a justificatory—not merely causal—role in the speaker's grasp of the content of an utterance. Call the resulting determination principle '(DET_{EP})':

(DET_{EP}) Speakers/hearers grasp the contents of utterances on the basis of their grasping the semantic value of the uttered sentence.

I think that there are good grounds for rejecting the idea that speakers grasp the contents of utterances even partly on the basis of grasping the semantic value of the uttered sentence. First, we need to be careful with how we are understanding “grasping”. While it might be plausible to think of hearers as grasping what is said or expressed by an utterance in an ordinary sense of ‘grasp’, speakers do not stand in the ordinary grasping relation to semantic values. We seem to be able to reflect, often enough, on our interpretations of utterances, and perhaps use these in explanations of our own behaviors. The states that constitute our semantic processing are, by contrast,

not open to this sort of reflection. While it is an open, empirical question just what sort of states those are, it is uncontroversial that coming to know the semantic values of our own expressions (including sentences) requires engaging in the kinds of scientific investigation that semanticists are engaged in when they do semantics.

One might object here that any competent English speaker knows (or at least can come to know on reflection) the truth-conditions of their own sentences. So if one identifies the semantic value of a sentence with its truth-conditions (as truth-conditional semanticists do), then speakers should be able to grasp, in the ordinary sense, the semantic values of their sentences (if not the sub-sentential expressions). However, this objection misunderstands the nature of explanations in truth-conditional semantics and also underestimates the complexity of the specifications of truth-conditions that truth-conditional semanticists posit. First, in truth-conditional semantics, the truth-conditions that are specified by the semantic theories are theoretical posits which are meant to figure in explanations of speakers' judgments of anomaly, entailment, contradiction, truth-value, and so on. Semanticists do not take speakers' judgments of truth-conditions as evidence for a particular proposal about the truth-conditions of sentences (except insofar as they might reflect judgments of entailment or truth-value, which will be a source of defeasible evidence for thinking a sentence has certain truth-conditions.). Second, the specifications of truth-conditions that one finds in truth-conditional semantics are complicated, involving mathematical and other technical concepts that speakers need not, and typically will not possess. Consequently, speakers are typically not in a position to be able to reflect on truth-conditions as they are specified by the semanticist.

If (DET_{EP}) requires that speakers be able to grasp the semantic values of their expressions via introspection so that they can use facts about semantic values as evidence for a particular interpretation of an utterance, we have good reason to reject

it. No such access is available, and the resulting picture of the psychology of linguistic interpretation is unrealistically deliberate and over-intellectualized.

We need not read (DET_{EP}) that way, though. We can read it, instead, as simply claiming that an uttered sentence's being assigned a particular semantic value makes it likely that—or reliably indicates that—the utterance expresses a particular content. Consequently, a hearer is justified in interpreting the utterance accordingly, despite not being able to cite the uttered sentence having its semantic value as a reason for doing so.

However, talk of justification of the interpretation of an utterance, in this context, is out of place, given that there is no sense in which the hearer can fail to “get it right”, from the theorist's perspective. Let me elaborate. It is true that interpreters generally take it that they can fail and succeed in their interpretations. Indeed, this is typically understood as a matter of interpreting a speaker as the speaker intended (though not always). And indeed, it might be true that, generally, a hearer assigning a particular semantic value to an uttered sentence is good evidence that a speaker intends to express a particular content. However, content, in the sense that we have are interested in, does not supervene on facts about the intentions of speakers.³⁶ Rather, an utterance's expressing a particular content is constituted by (or perhaps identical with) an interpreter's interpretation itself. Whether or not the interpreter interprets a speaker in the way that the speaker intends is an additional question, and one which does not need to be answered in order to give explanations of the relevant linguistic behaviors.³⁷

³⁶Note that a speaker might interpret their own utterance differently from the way in which they intended their utterance to come across. This can happen when a speaker is intending to communicate some content, but recognizes that their utterance fails to capture what they are trying to say, for example.

³⁷One might rightly point out that whether or not an interpreter interprets as the speaker intends to be interpreted will be relevant to whether or not a given proposition makes it into common ground. Facts about common ground, however, construed the way Stalnaker (2014) construes them are not relevant for the proximate explanations of behavior that I have been focusing on. What will be

Finally, (DET_{EP}) simply would not be a suitable constraint to place on linguistic theorizing, even if it turned out to be true. The reason is just that linguistic theory, as a branch of cognitive science, is a theory which seeks to describe the human language faculty, which is partly responsible for the shape of our linguistic behaviors and practices. It does not seek to explain—at least within the confines of semantics and the sort of theory of communication that I have been discussing—why such behaviors or practices are *justified*.³⁸ (DET_{EP}) would require that linguistic theories assign semantic values to expressions that justify speaker interpretations of utterances, and assign contents to utterances on the basis of what interpretation a speaker would be justified in arriving at, given the uttered sentence’s semantic value. (DET_{EP}) would, consequently, force constraints on linguistic theorizing which go beyond its explanatory aims.

5.3 Causal Determination

The final sense of determination we will consider is causal determination. As was the case when we considered metaphysical determination, we can rule out the idea that determination is full determination. The mental states that constitute the assignment of semantic value to an expression will not fully causally determine the mental states that constitute interpretation of an utterance for reasons of context-sensitivity—the interpreting of an utterance in a particular way also causally depends on the states which constitute grasping particular features of context. In other words, we need to assess (DET_{causal}) :

(DET_{causal}) : A speaker’s/hearers interpreting an utterance of S as expressing content relevant, again, will not be what is in common ground, per se, but roughly, what each speaker *takes* to be in common ground.

³⁸I qualified that this is not the aim of the linguistic theories that I have been talking about because an explanation for why some practice (e.g.) is justified might be part of a broad explanation of why our practices persist, or why they have taken their distinctive shapes. This sort of explanation can be seen in the work of Millikan (1984), for instance, and could be construed as being part of linguistics, broadly conceived.

c is partly causally determined by their assigning semantic value m to S .

Unlike (DET_{EP}), this is the sort of determination principle that would seem to fit within a psychological research program. What DET_{causal} suggests is that there is a strong causal link between semantic processing (narrowly construed) and the process of interpreting an utterance. Some such link is surely well-motivated. As I suggested in §2, even if the immediate product of semantic processing does not constitute an interpretation of an utterance, it surely plays an important role in our abilities to interpret and to communicate, generally. At least, that there is some important link between our semantic competence and our communicative abilities is a well-motivated guiding assumption.

Our question, however, is whether the proposed causal link between semantic processing and utterance interpretation constitutes a determination relation which would put major constraints on theorizing on semantics and communication. Before we can answer this question, however, we need to clarify what ‘partial causal determination’ would come to in this context. It would not be particularly interesting if all that was required is that an uttered sentence’s having semantic value m is one of the things that *can* cause a speaker to interpret the sentence as expressing content c . Nor is it sufficient that a sentence S ’s having m is merely a common or usual partial cause of utterances of S expressing c . In that case, the semantic value of an expression would still not constrain the content it is used to express—expressed content could, in some cases, vary completely independently of semantic value.

At the very least, then, it would seem that the content expressed by an utterance of S should causally *depend* (for lack of a better word) on S having semantic value m . In other words, it should be the case that for any utterance U of a sentence S , U ’s expressing content c is partly causally depends on S ’s having m . Further, since the determination principle is meant to be a guiding methodological principle in linguistics,

it should hold as a matter of psychological necessity. That is, for all adult humans with typical psychology, the process of interpretation is causally dependent on their semantic processing in some way.³⁹

But how should we characterize this dependence? I don't think we can require that the dependence relation be counterfactual dependence. After all, it might be that two sentences S_1 and S_2 , with distinct semantic values m_1 and m_2 , express the same content c_1 at some context C . At least, we don't know enough about what sort of property will realize the *having such-and-so semantic value* role to rule this possibility out. I think, then, we should think of the dependence as just requiring that, as a matter of psychological necessity, interpretations of utterances of sentences are partly causally explained by the assignment of semantic value to those uttered sentences. More explicitly, I think we should interpret the determination principle as (DET_{ling}) , the "linguistic determination principle":

(DET_{ling}) : For all psychologically normal humans, for any utterance U , if U is an utterance of a sentence S , and U expresses content c , then there is a semantic value m such that S has m and U 's expressing c is partly causally explained by the fact that S has m .

Keeping in mind that, in this context, a sentence's expressing some content and having some semantic value are determined by the particulars of one's semantic processing and their process of utterance interpretation, (DET_{ling}) expresses a psychologically necessary causal connection between the two.

Now, there is a case to be made here that (DET_{ling}) is not strong enough to support the determination principle's usual role as a guiding methodological principle. More

³⁹The principle of compositionality—another guiding methodological principle in linguistics—for instance, is taken to hold as a matter of psychological necessity in Szabó (2000) and Napolitano (2015).

specifically, it would not seem to be strong enough to justify the practice of pairing particular contents with expression-semantic value pair-types. The reason is that, while it ensures a law-like causal connection between occasions of utterance interpretation and occasions of semantic value assignment, it does not ensure that there is any kind of systematic causal relationship between the two. In other words, semantic value might not *constrain* content in any real sense. It might be, instead, that while assignment of semantic value is always a partial cause of one's interpreting an utterance in a particular way, the causal relationship is highly complex, involving the intervention of any number of other variables. This would be the case if, for instance, the contents of utterances can be sensitive to a potentially open-ended number of features of context, as radical contextualists hold. When we pair content-types with sentence-semantic value pairs, the idea is typically that utterances of the relevant sentence (which has its semantic value invariably) express such-and-so content-type. But if semantic value of the uttered sentence is just one of many properties of utterances that can determine the content of an utterance, then we need to complicate how we think of utterance types. We cannot simply pair contents with utterances of tokens of a particular sentence-type, but rather utterances of tokens of a sentence-type at a particular context-type. If it turns out that there are any number of features of context which can affect the interpretation of utterances, however, such pairings would seem to be of little theoretical interest.⁴⁰

But I have not argued that utterances are radically context-sensitive in this way. Consequently, I think we should countenance another version of the determination principle (DET_{sys}):

(DET_{sys}): For all psychologically normal humans, for any utterance U , if U is an utterance of sentence S , S has semantic value m , and U has contextual features $\phi_1 \dots \phi_n$, then U expresses content c and U 's expressing c is caused by U 's being

⁴⁰This point will be expounded on in §6.1.

an utterance of S , S 's having m , and U 's having contextual features $\phi_1 \dots \phi_n$.

What (DET_{sys}) suggests is that there is a systematic causal relationship between a sentence having its semantic value and utterances of that sentence expressing a particular content, given certain contextual parameters. If the relevant contextual parameters are not open-ended, and can be specified in a theory, then this would justify the practice of pairing triplets consisting of a sentence-type, semantic value, and the relevant contextual parameters, with a content that is expressed by utterances of the relevant sentence in the relevant contexts.

6 Does Semantic Value Determine Content?

The relevant question, then, is whether our semantic processing is involved in our interpretation of utterances in the way suggested by (DET_{ling}) or (DET_{sys}) . The question, obviously, is an empirical one, but I think that given some standard assumptions about semantics and natural language, we have good reason to think that the connection between semantic processing and interpretation is looser. I will argue in §6.1 that certain kinds of “non-literal” talk suggest that (DET_{sys}) is unsuitable as a guiding methodological principle. In §6.2, I will argue that the interpretation of utterances of anomalous sentences, ungrammatical strings, and micro-conventions suggest that both (DET_{sys}) and (DET_{ling}) are false.

6.1 The Problem of “Non-literal” Talk for (DET_{sys})

I want to argue first that the insights of so-called “radical contextualists” give us reason to reject (DET_{sys}) as a guiding methodological principle in linguistics.

Radical contextualists have argued that language exhibits much more context-sensitivity than is often countenanced by semanticists. They emphasize that the phe-

nomenon, is, in fact, fairly ubiquitous, and that the contents of utterances can depend on any number of features of context. Furthermore, the contents expressed by utterances of sentences can depend on features of context in ways that are not explained by the context-sensitivity of the sentence constituents.⁴¹

Here are some of the classic examples:

(3.1) The ham sandwich wants coffee.

(3.2) The leaf is green.

(3.3) Mary has had enough.

The intuitive meaning of an utterance of (1), made by a member of the wait-staff to a co-worker, is that the customer who ordered a ham sandwich wants coffee (or wants us to bring him a cup of hot coffee, etc.). In another context, we might take (1) to express that the customer who always eats ham sandwiches wants to purchase coffee beans. Spoken by a child, we might take (1) to express a variety of category mistakes, depending on the context—ham sandwiches don't have desires, nor do toy sandwiches. The child, however, interprets her utterance as expressing a truth. She conceptualizes ham sandwiches or ham-sandwich-toys in such a way that they can want coffee.

Given the right sorts of contexts, we find that English sentences exhibit a great deal of flexibility—the same sentence can be used in different contexts to express very different contents.⁴² Furthermore, what content is expressed can seemingly depend on any number of features of context, the background beliefs of an interpreter, the interpreter's beliefs about the speaker, their beliefs about the speaker's beliefs about the interpreter, and so on.

⁴¹See, e.g., Recanati (2002), Carston (1988).

⁴²I use English only as an example. I take it that it would be very surprising if any communication system (not to mention natural language) was incapable of this kind of flexibility.

However, it is a standard assumption that semantic values of expressions remain somewhat fixed. In other words, an interpreter hears the same ‘ham’ with its usual semantic value, whether (1) is uttered by a member of the wait-staff or by a child. Non-semantic information combines with semantic value in some potentially complex way, and an interpretation is arrived at. This is problematic for (DET_{sys}) , since the examples suggest that the causal relationship between semantic value and content is complex and irregular, so that the assignment of the same semantic value to an expression can contribute to any number of differing interpretations, depending on the context. Indeed, the potential for context-sensitivity is even more radical in our current, internalistic context, since an interpreter might have any number of strange beliefs about the speaker, and consequently interpret them in strange, unreasonable ways. Content-expression, as we are understanding it here, does not depend on interpretations being reasonable or usual.

Supposing there are a finite number of contextual features of utterances which, when combined with the semantic value of the uttered sentence, suffice to causally determine the interpretation of the utterance, (DET_{sys}) might be true. However, the number of contextual features will be so great that (DET_{sys}) will be largely uninteresting. The practice of pairing expression-semantic value-pairs with contents would be practically impossible and unilluminating, since it would require individuating utterances not merely by the uttered sentence, but also by any number of contextual features of the utterance which might cause it to express the contents it expresses. Since an interpreter might bring any piece of information to bear on the question of what a speaker means by their utterance, it would, in effect, require a theory of human practical reasoning, and no such theory is forthcoming.⁴³

This basic insight about the interpretation of utterances makes (DET_{sys}) unsuitable

⁴³See Chomsky (2000, Ch.2), and also Davidson (1986).

as a guiding methodological principle, even if true. I will argue in the coming sections that there are reasons to think it is false, but first, I will defend the argument I've just made from radical context-sensitivity to the conclusion that (DET_{sys}) is unsuitable as a guiding methodological principle.

Let me consider some immediate objections that might arise to the argument I've made against (DET_{sys}). First, it ought to be stressed that the argument does not presuppose radical contextualism. Radical contextualism is typically construed as a thesis about the truth-conditions of sentences or the “semantic content” of utterances, and these are a matter of a sentence's meaning or semantic value. The debate, then, is about how we conceive of semantic values and thus, how we should conceive of semantic theories. The argument I've given, however, does not assume anything about the truth-conditions of sentences or their semantic contributions to utterance content. Rather, it makes use of the insight that serves as part of the motivation for radical contextualism—namely, that the intuitive truth-conditions of (or propositions expressed by) *utterances* are radically context-dependent. And strictly, what I require is a much weaker claim—namely, that the *interpretation* of utterances by interpreters is radically context-sensitive, varying from individual to individual, depending on any number of their background beliefs.

Nobody, including the opponents of radical contextualists, deny that speaker meaning, for instance, is radically context-sensitive. But the content expressed by an utterance, as I've understood it here, plays a theoretical role similar to that which speaker meaning is typically taken to play. Though strictly, expressed content in my sense is more like “what interpreters take speaker meaning to be”. These kinds of beliefs, uncontroversially, will vary wildly depending on any number of beliefs of the interpreter.⁴⁴ The dispute between radical contextualists and their opponents, again, has little to do

⁴⁴See Cappelen & Lepore (2005, 187), e.g.

with content-expression in the sense relevant here.

Now, I do think that there is an objection to my argument against (DET_{sys}) that might be raised here by semantic minimalists like Cappelen & Lepore (2005). They might argue that in assessing the determination principle, I've focused on the wrong notion of content-expression, and that using their notion of "semantic content", the determination principle can be rescued. On their view, the semantic content of (1) is just that the relevant ham sandwich wants coffee. Speakers may interpret the utterance of (1) as meaning that a particular customer wants coffee, or wants to be given coffee, but this is not the semantic content of the utterance. How speakers arrive at their final interpretation of the utterance, they agree, will be an extremely complicated, unsystematic affair about which their might not be much to say that is of theoretical interest in linguistics.⁴⁵

The obvious question to ask is whether or not semantic content in Cappelen & Lepore's sense—call this "minimal content"—belongs in the kind of linguistic theories that are under discussion here. Carston (1988, 40) argues, for instance, that such a notion of semantic content has no psychological reality. In that case, it would clearly not figure importantly within the generative linguistics framework. But Cappelen & Lepore argue that, in fact, minimal content is psychologically real. On their view, the proposition that is the minimal content of the utterance is the one which roughly, one can grasp even if one knows nothing about the speaker, audience, or context of utterance. It is that proposition which "the speaker can expect the audience to grasp (and expect the audience to expect the speaker to expect them to grasp) even if they have mistaken or incomplete communication-relevant information" (184). While grasping the semantic content typically falls well short of grasping the proposition that the speaker might be most concerned with sharing with their audience, minimal

⁴⁵See Cappelen & Lepore (2005, Ch. 13).

content is “our minimal defense against confusion, misunderstanding, and it is that which guarantees communication across contexts of utterance” (2005, 184-185). Thus, we might not be able to grasp what a speaker intends to convey by uttering (1) if we are unfamiliar with the details of the context, but we (and they) at least know that they are saying something about a ham sandwich and its desire for coffee.⁴⁶ We can then apply whatever information we have about the speaker and the context of utterance to figure out what else the speaker means to convey.

I don’t think there’s any reason to deny that minimal content might be psychologically real—the minimal content of (1) may, very often, play the role in interpretations of (1) that Cappelen & Lepore describe. The question is whether the role is one which fits into the kind of theory that will be continuous with semantics in generative linguistics. If it does, then we can ask whether some version of the determination principle is vindicated. I now want to briefly argue that even if minimal content does have a home in the sort of theory under discussion, no resulting version of the determination would be vindicated. I will then offer reason to think that, in fact, minimal content has no such home in the first place.

While minimal content will, in many cases, serve as the kind of “fallback” content that Cappelen & Lepore describe, such content will very often play no role in the kinds of explanations that we expect of content. For instance, when the member of the wait-staff utters (1) to a co-worker, the co-worker does not take the speaker to be saying anything about ham sandwiches. The proposition that the ham sandwich wants coffee does not enter into common ground, they do not come to believe that proposition, nor do they come to believe that the speaker believes it, and so on. Minimal content, much of the time, will not figure in explanations of communicative behavior.

⁴⁶Importantly, Cappelen & Lepore hold that an utterance can express any number of propositions, so they don’t need to deny that the utterance of (1) might also say something about a particular customer, for instance.

Cappelen & Lepore maintain, however, that the minimal content “of a sentence S is the proposition that all utterances of S express (when we adjust for or keep stable the semantic values of the obvious context sensitive expressions in S)” (3). I think it is clear, then, that they are conceiving of the theoretical role of minimal content differently, such that it is not meant to figure in the kinds of explanations that content, in the sense I have been focusing on, figures in. In fact, the way they describe it, it would seem that it might be a candidate realizer for the semantic value role. If the semantic value role were realized by minimal content, however, this would be no help to (DET_{sys}). Cappelen & Lepore agree that examples like the ones given in this section suggest that the relationship between minimal content and content in our sense is extremely complicated and unsystematic.

However, it is also clear that Cappelen & Lepore do not intend minimal content to play the role of semantic value (143-145). We need an additional category of semantic value, at least to account for (by their lights, the limited set of) context-sensitive expressions like indexicals. Minimal content, unlike semantic value, is not constrained by compositionality. Minimal content is the proposition that results once we fill in the semantic value with the relevant features of context that any context-sensitive expressions might require. What we have, then, is a proposal for a new theoretical role for minimal content to occupy—one that sits, in effect, “between” semantic value and content in the order of explanation. Grasp of the semantic value of an uttered sentence results in the grasp of the minimal content, and the grasping of minimal content figures in the process of interpreting utterances. In that case, perhaps (DET_{sys}) should be understood as claiming that there is a systematic causal relationship between the assignment of semantic value and the grasping of minimal content.

If minimal content is not the content expressed by an utterance in the sense we have been interested in, however, it’s difficult to make sense of how it is “grasped” in

the interpretation of every utterance. But this is what would be required for minimal content to be of interest in the internalist theories of semantics and communication under consideration. In light of Cappelen & Lepore's comments about minimal content, perhaps we should understand speakers' grasping minimal content as consisting in their having the expectation that fellow speakers will interpret an utterance as at least expressing that minimal content. Since minimal content is a technical notion, however, and since interlocutors do not expect of each other that they will at least interpret an utterance as expressing the minimal content in the more full-blooded sense (i.e. they will not add the minimal proposition to common ground, and so forth), we need to unpack the expectation further. Perhaps the expectation is something like an expectation that interlocutors will add the minimal content to common ground in the absence of further contextual information needed to determine what is expressed by the utterance in the more full-blooded sense. In that case, speakers could "grasp" minimal content even if their so doing does not play any explicit or implicit role in interpreting utterances (if both parties recognize plenty of the context-dependent information that goes into interpretation, and recognize that each other recognize this information).

I think it is extremely implausible that speakers have such stable expectations with respect to the interpretation of utterances, as the examples considered in this section and the next section should suggest. But even if they do have such expectations, it's not clear why these expectations play such an important role that they require a special theoretical role to occupy. If interlocutors have expectations about how they ought to interpret utterances in the absence of information about the context of utterance, then these expectations would just seem to constitute just another piece of information that goes into the interpretation process. It doesn't seem as though that sort of information needs to be privileged more than, say, an expectation for utterances of a sentence involving 'the x', that some particular x be something which is salient to

the interpreter.

Finally, it's not clear that Cappelen & Lepore think that speakers actually *do* have these expectations with respect to every utterance. Rather, they write that minimal content is the proposition that speakers and hearers *can* expect one another to grasp in the absence of relevant contextual information. But what is the force of the 'can' here? The obvious reading is an epistemic one, such that speakers would be reasonable to have the relevant expectations. But as I argued in §5.2, theoretical roles which are understood partly in terms of epistemic constraints are out of place in generative linguistics.

There doesn't seem to be any explanatory need for a notion of what is said or of content-expression such that all utterances of (1) express something about ham sandwiches. Indeed, the best justification in favor of the notion just is some version of the determination principle, and in particular, something along the lines of (DET_{sys}). But, of course, the assumption of (DET_{sys}) in linguistic theorizing is precisely what I am questioning via reflection on the theoretical roles occupied by semantic value and content-expression. Note too that while Cappelen & Lepore do not identify minimal content with semantic value, I am not ruling out that the *semantic value* of (1) might be the proposition that the contextually salient ham sandwich wants coffee. The argument I have made against (DET_{sys}) would still stand in that case, since the contents of utterances could still vary widely from context to context.⁴⁷

It perhaps ought to be stressed here that the notion of content expression that I have been focusing on is also at odds with the notion employed by radical contextualists. Recanati (2002) explicitly denies, for instance, that his notion of 'what is said' (i.e.

⁴⁷Some background assumptions here are that natural language is not intractably ambiguous, and that an expression has a single semantic value. One could propose either massive ambiguity, or propose that an expression can have many semantic values corresponding to the various contents that might be expressed by utterances of the relevant expressions. In either case, however, the task of assigning semantic values to expressions would become intractable and uninteresting. This, presumably, is why these options are not typically considered to be live ones.

expressed content) is identical to interpreters' understanding of what is said. This latter view of what is said is, roughly, the view of content-expression that I argued for in §4. On his view, what is said is determined by how a “normal” speaker would understand the utterance, where the normal speaker has access to quite a lot of information about the context of utterance.

This raises a few objections that need to be addressed. What the objections suggest is that the determination principle should be restricted to, for instance, literal utterances, normal speakers, or to particular languages. If we restrict (DET_{sys}) to literal utterances, for instance, then the non-literal utterances of (1) that I took to be problematic for (DET_{sys}) are meant to be excluded. The argument I've made here against (DET_{sys}) , then, would not work. I will take up these objections in §7.

For now, I conclude that (DET_{sys}) , even if true, is not suitable as a guiding methodological principle in linguistics.

6.2 Problems of Anomaly, Ungrammaticality, and Micro-conventions

The argument in §6.1 leaves open the possibility that (DET_{sys}) is true, even if it is unsuitable as a guiding methodological principle. It also leaves open the possibility that (DET_{ling}) is true. After all, it might be that there is no systematic causal connection between semantic value and content, but so long as facts about the semantic values of sentences are partly causally responsible for interpreters interpreting utterances the way they do, (DET_{ling}) is not threatened. It is worth stressing, again, that (DET_{ling}) would not justify the practice of pairing sentence-semantic value pairs with the contents expressed by those utterances. Given the radical context-sensitivity of content-expression, there just isn't much sense in talking about “the” content expressed by utterances of some sentence.

Nevertheless, (DET_{ling}) would place a substantial constraint on semantics and the

theory of communication. It requires, in effect, that the process of interpretation be causally dependent on our semantic processing—the assignment of semantic value is part of a complex interaction of an interpreter’s beliefs about the speaker, the context, etc., which results in an interpretation of an utterance. The effect on linguistic theorizing would be that a full explanation of any interpretation of an utterance would require appealing to the semantic value of the uttered sentence, and then an account of how the assignment of semantic value interacts with the other relevant mental states of an interpreter to yield that interpretation. Now, it is true that we will want a general account of how assignment of semantic value interacts with other mental states of an interpreter in the process of communication. But (DET_{ling}) would make the job a bit harder, since it rules out the possibility that, sometimes, assignment of semantic value is not part of the process of interpretation. Consequently, semantics will be constrained by every exotic case of utterance interpretation, since it will be required that semantics assigns a semantic value to any interpreted sentence, and that it fit with a general account of how assignment of semantic value interacts with other mental states of an interpreter to result in an interpretation.

This, by itself, is no objection to (DET_{ling}). There might, after all, be such a constraint on semantics. I want to argue now, however, that we have good reason to reject (DET_{ling}), given other standard assumptions given the place of semantic processing in relation to interpretation and other sorts of linguistic processing. My strategy will be to present a range of cases in which (DET_{ling}) conflicts with these assumptions. I will conclude that if forced to choose between giving up all of these assumptions and giving up on (DET_{ling}), we ought to give up on (DET_{ling}).

6.2.1 Interpretation of Anomalous Speech

Consider sentences like the following:⁴⁸

(3.4) *Mary is completely tall

(3.5) *John realized the answer for hours.

(3.6) *Colorless green ideas sleep furiously.

These are cases of “semantic anomaly”—they are typically judged as being unacceptable, but the explanation for their unacceptability is semantic, and not syntactic or pragmatic. The unacceptability of (4) is explained by Kennedy & McNally (2005) by the fact that *completely* is a degree modifier which can only modify a gradable adjective which is associated with a scale that is closed on the upper end. Informally, there is no maximum height value, and so there could be no tallest thing. Things—boxes and jars for instance—can be completely empty, and so *completely* modifies *empty* without unacceptability. The unacceptability of (5) is explained by the fact that *realize* denotes an event type that is known as an “achievement”—roughly, a kind of event which is more or less instantaneous. The modifier *for hours*, however, typically needs to combine with an adjective denoting a (temporary) state as in *sad for hours* or *full for hours*, or an activity that can go on for some duration as in *swam for hours* or *read for hours*.⁴⁹

Now, the full explanation of the unacceptability of instances of semantic anomaly typically goes further than just acknowledging constraints on the combination of certain expressions without unacceptability. For instance, given Kennedy & McNally’s

⁴⁸For examples of anomaly involving gradable and degree modifiers like that in (4), see Kennedy & McNally (2005). See Jackendoff (2010) for discussion of examples like (5). (6) comes from Chomsky (1957). See also Pustejovsky (1995) for a wide variety of cases of anomaly.

⁴⁹The state must be a temporary one because of the oddness of **tall for hours* or **smart for hours*. See, e.g., Jackendoff (2010), and Vendler (1957) for further discussions of these kinds of examples.

semantics for gradable adjectives and degree modifiers, the semantic value of (4) is simply undefined. The fuller explanation, then, is either that expressions with undefined semantic values are judged to be unacceptable or that the semantic values of *completely* and *tall* simply fail to compose, and so no semantic value for (4) is computed, resulting in unacceptability. For Jackendoff (2010, Ch. 6), the unacceptability of sentences like (5) is explained by the fact that the semantic values of *realized* and *for hours* cannot compose, and when this happens, the result is unacceptability.

In other words, a common explanation for the unacceptability of cases of semantic anomaly is that anomalous sentences fail to get an assignment of semantic value at all. Despite this, utterances of anomalous sentences can be interpreted. In many cases, we might *not* be able to interpret utterances of these sentences. But what matters for our purposes, is that they *can* be interpreted in various contexts. A child that utters (4), for instance, might plausibly be taken to simply mean that Mary is very tall. There is nothing that prevents us from interpreting her utterance that way. In such a case, since, by hypothesis, the sentence has no semantic value, there is no fact about its semantic value that can partly causally explain the interpretation of the utterance.

Let me address what I think are two unpromising responses here. A first response is that perhaps the right way to understand anomaly is that it results in unacceptability because the semantic value of anomalous expressions is undefined. But to have an undefined semantic value is still to have a semantic value, and it might be that the fact that the semantic value of (4) is undefined partly explains why we interpret the utterance the way we do. This is a logical possibility, but given that all semantically anomalous expressions would receive the same, undefined semantic value, and given that different utterances of anomalous sentences might receive any number of interpretations, it seems implausible that a single semantic value could plausibly figure in causal explanations of such a wide variety of interpretations.

Another response is that what explains the unacceptability of anomaly, generally, is just that anomalous sentences are necessary falsehoods, and necessary falsehoods are, generally, unacceptable.⁵⁰ In reply, we should note, first, that the best semantic theories in linguistics do not give this kind of explanation of anomaly, but give the kinds of explanations that I sketched above. Second, appealing to necessary falsehood as an explanation for the unacceptability of anomaly is simply implausible. While (4)-(6) might plausibly be taken to be necessary falsehoods, many necessary falsehoods are not anomalous.⁵¹ For example, $2+2=5$ and *water consists in nitrogen* are, for many people, necessary falsehoods, but they are not anomalous, and do not exhibit the unacceptability of (4)-(6).

Here is a more promising response. While it cannot be that the semantic values of the sentences in (4)-(6) play a role in the interpretation of utterances of (4)-(6), surely the semantic values of some of the sub-sentential components do play such a role. For instance, when a child utters (4), and we take her to be saying something about Mary's height, we presumably do so, in part, because of the semantic value of *Mary*, the semantic value of *tall*, and so on. So while the best accounts of semantic anomaly might conflict with (DET_{ling}), they do not conflict with a weaker principle which would still ensure a causal role for semantic processing in all interpretation. Call this principle '(DET_{subsent})':

(DET_{subsent}): For all psychologically normal humans, for any utterance U , if U is an utterance of a sentence S , and U expresses content c , then there is a semantic value m and an expression e such that S has e as a constituent, e has m , and U 's

⁵⁰Fodor and Lepore (1998) claim, for example, that instances of semantic anomaly are to be explained either by appeal to syntax, necessary falsehood, or pragmatic unacceptability. See, however, Pustejovsky (1998).

⁵¹Note that it is not obvious that (4)-(6) are, in fact, necessary falsehoods. If their semantic values are undefined, or if they have no semantic value at all, then they would not be necessary falsehoods. Instead, they would presumably fail to be truth-apt.

expressing c is partly causally explained by the fact that e has m .

I will argue in §6.2.3 that (DET_{exp}) is false as well. For now, let me just note that to fall back on it is to give up on (DET_{ling}) and to acknowledge that the connection between semantic value assignment and utterance interpretation is much messier and complicated than is typically assumed.

The upshot is just that endorsement of (DET_{ling}) seems to be in conflict with standard ways of explaining semantic anomaly.

6.2.2 Interpretation of Ungrammatical Speech

It is a standard assumption in linguistics that syntax and semantics bear a tight connection to one another. In particular, the syntactic arrangement of the simple expressions that compose a complex expression dictate how the semantic value of the complex is calculated. Consider, for instance, the common composition rule Functional Application:⁵²

Functional Application (FA): If α is a branching node with β and γ as daughters, and $\llbracket \beta \rrbracket \in \llbracket \gamma \rrbracket$, then $\llbracket \alpha \rrbracket = \llbracket \gamma \rrbracket(\llbracket \beta \rrbracket)$.

The properties of being a branching node, being a daughter of a branching node, and so on, are syntactic properties. Composition rules like Functional Application determine how the semantic values of complexes are calculated on the basis of the semantic values of the constituent expressions, and the syntactic structure of the complex. Consequently, if a syntactic structure is not assigned to a complex expression, the complex will not be able to undergo semantic composition, and its semantic value will not be calculated.

⁵²I take this rule from Heim & Kratzer (1998).

In syntax, however, a standard explanation of what it is to be an ungrammatical complex is to be such that it fails to generate a syntactic structure that can be interpreted by the semantic system. That is, given the syntactic properties of the simple expressions involved, and given a stock of syntactic rules, the syntactic faculty is not able to produce an object which the semantic systems can take as an input to semantic processing.⁵³ The complex, then, is assigned no syntactic structure that can be interpreted by the semantic faculty, and so semantic composition will not take place.⁵⁴

Here are a few examples of ungrammatical sentences from Adger (2003):

(3.7) *Anson demonized.

(3.8) *Gilgamesh failed often Herbology.

(3.9) *Enkidu wanted to live and Ishtar tried to do.

By hypothesis, the complexes in (7)-(9) have no semantic value. Nevertheless, it would not necessarily be difficult to interpret utterances of any of (7)-(9) in the right context. A plausible interpretation of (8), for instance, would just be that Gilgamesh often failed Herbology. (9) might seem more difficult to interpret, but in the right context, it need not be. We might, for instance, know a child who has trouble with auxiliaries like the English *do*. Indeed, their errors might be quite predictable, so that their utterances of ungrammatical strings are nevertheless easy to interpret. Their utterance of (9), for instance, might just be interpreted as expressing that Enkidu wanted to live and Ishtar tried to live.

In that case, though, facts about the semantic value of the sentence could not figure in the causal explanation of our coming to interpret the utterances the way we did. In

⁵³See Adger (2006, 143).

⁵⁴The standard assumption tends to be the syntactic processing takes place first, and then semantic processing operates on the resulting syntactic structure. However, the main point here is compatible with so-called “direct compositionality”, the thesis that syntactic and semantic composition happen simultaneously.

other words, (DET_{ling}) is in conflict with either standard accounts of ungrammaticality or with standard accounts of the relation between syntax and semantics.

Jackendoff (2010), for instance, argues that semantic and syntactic processing result in distinct syntactic and semantic structures, and that they are built up in parallel.⁵⁵ In that case, it's possible that a complex could have a semantic value even if it weren't grammatical. This is a minority position, however, and one which has far-reaching consequences for syntax and semantics. It should be noted, also, that taking on Jackendoff's account of the relationship between syntax and semantics does not help with the argument in §6.2.1 that (DET_{ling}) conflicts with standard explanations of anomaly.

As was the case in §6.2.1, the argument in this section does not impugn ($\text{DET}_{subsent}$). After all, it seems likely that in the examples of the interpretation of ungrammatical utterances I considered, facts about the semantic values of the sub-sentential expressions of the utterance partly causally explain the proposed interpretations.

The argument in the next section, however, will tell against both (DET_{ling}) and ($\text{DET}_{subsent}$).

6.2.3 Codes and Micro-conventions

Thus far, I have considered examples where we have theoretical reasons to think that an uttered complex lacks a semantic value, and yet interpretation of the utterance can proceed easily enough. It did seem, however, that the semantic value of *some* of the uttered expressions played a causal role in interpretation. But not all cases of interpretation are like this.

Consider the following example. Suppose you and a friend are going to a party together. Your friend doesn't want to stay out late, but for whatever reason, they find it incredibly uncomfortable to mention to anyone (besides you) that they want to leave

⁵⁵See also Jackendoff & Culicover (2005).

any social gathering. So they tell you that if they utter a sentence with only three syllables, then that is their way of signaling to you that they are ready to leave the party. Later on that night, they say to you, out of the blue, “Jim likes dogs.” You interpret them as saying that they are ready to leave the party. You alert everyone else that it is time for you and your friend to leave, and you leave.

You and your friend, essentially, set up a kind of “micro-convention”, or an impromptu code by which to interpret certain utterances. Using this code, you interpret them as saying that they are ready to leave the party—you come to believe that they want to leave the party, that they want to leave the party becomes common ground, etc. If we are interested in explaining your behavior in relation to the utterance, then in order to do that, we posit that the utterance expressed that they want to leave the party. That Jim likes dogs does not (or at least need not) enter into the common ground, you do not come to believe it, and so on.

This is a case, then, where your interpretation of the utterance doesn’t seem to depend at all on any facts about the semantic values of the uttered expressions. Rather, the reason you interpret as you do is because of the antecedent agreement you make with your friend and because of certain prosodic features of the sentence “Jim likes dogs”—namely, that it has three syllables. In principle, there is no limit to the kinds of non-semantic features of an utterance that might suffice to determine the way an utterance is interpreted. Two speakers might set up a micro-convention on which certain utterances are to be interpreted solely on the basis of their syntactic features, their phonetic features, the pitch of the utterance, the volume, the temperature at the time and location of utterance, the time of day, and so on. It seems very implausible in these kinds of cases that the semantic values of any of the uttered expressions have any bearing on the interpretation, even when the utterances are utterances of sentences which have semantic values (though they needn’t be).

If the examples I've described are psychological possibilities—and it seems clear that they are—then they would suggest that (DET_{ling}) , $(\text{DET}_{subsent})$, and (DET_{sys}) are all false. There are cases where interpretation of an utterance does not causally depend, at all, on facts about uttered expressions having their semantic values (even if the uttered sentence has a semantic value). As these versions of the determination principle were the last hope for justifying the determination principle's role in linguistics, I conclude that there is no version of the determination principle that can serve as a guiding methodological principle. Each version either fails to appropriately constrain linguistic theorizing in any interesting way, or else is false.

The upshot is just that semantic information is just one kind of information that can go into the interpretation of the utterance, and that in some cases, the semantic information is either unavailable, or is not utilized in interpretation. This does not mean that this semantic information is not tremendously important in our communicative lives. Indeed, it seems likely that most of the time, assignment of semantic value goes a long way towards determining how an utterance is interpreted. However, to acknowledge the importance of the assignment of semantic value in interpretation is not to provide support for the idea that semantic value *determines* content in the relevant sense.

7 Objections and Replies

7.1 Restrictions to Utterances

One of the most glaring objections is the following: some of the examples I relied on to argue against (DET_{sys}) (DET_{ling}) and $(\text{DET}_{subsent})$ appealed to non-literal utterances or appealed to code languages which are, therefore, not natural languages. The determination principle, the objection continues, ought to be restricted to literal utterances

of natural language sentences. The examples relying on (1)-(3), for instance, would not be a problem for (DET_{sys}) if we restrict it to literal utterances, since the examples were examples of non-literal utterances. The suggestion, then, is that we can amend (DET_{sys} as follows):

(DET_{sys}*): For all psychologically normal humans, for all natural languages L , for any utterance U , if U is a literal utterance of sentence S , S is a sentence of L , S has semantic value m , and U has contextual features $\phi_1 \dots \phi_n$, then U expresses content c and U 's expressing c is caused by U 's being an utterance of S , S 's having m , and U 's having contextual features $\phi_1 \dots \phi_n$.

Likewise, it might be argued, the examples of micro-conventions in §6.2.3 are not, strictly, utterances of natural language, but of impromptu code-languages established by the speakers involved. Consequently, they threaten neither (DET_{ling}) nor (DET_{subsent}) if we amend those principles appropriately:

(DET_{ling}*): For all psychologically normal humans, for all natural languages L , for any utterance U , if U is a literal utterance of a sentence S , S is a sentence of L , and U expresses content c , then there is a semantic value m such that S has m and U 's expressing c is partly causally explained by the fact that S has m .

(DET_{exp}*): For all psychologically normal humans, for all natural languages L , for any utterance U , if U is a literal utterance of a sentence S , S is a sentence of L , and U expresses content c , then there is a semantic value m and an expression e such that S has e as a constituent, e has m , and U 's expressing c is partly causally explained by the fact that e has m .

It could be argued further that ungrammatical and anomalous complexes are not sentences of natural language either, and so I have not given any reason to doubt these three versions of the determination principle.

I will break this objection into two suggestions: the first, to restrict the principles to particular kinds of utterances—literal utterances, sincere utterances, and so on—and second, to restrict the principles to utterances of sentences in a particular language.

7.1.1 Restricting the Principles to Literal Utterances

Let's first consider (DET_{sys}*):

(DET_{sys}*): For all psychologically normal humans, for all natural languages L , for any utterance U , if U is a literal utterance of sentence S , S is a sentence of L , S has semantic value m , and U has contextual features $\phi_1 \dots \phi_n$, then U expresses content c and U 's expressing c is caused by U 's being an utterance of S , S 's having m , and U 's having contextual features $\phi_1 \dots \phi_n$.

This principle escapes the argument that was made against the viability of (DET_{sys}) as a guiding methodological principle in §6.1 because, it can be argued, the examples demonstrated the flexibility and context-sensitivity of communication. But according to (DET_{sys}), the determination principle isn't concerned with those cases, because it only holds for literal utterances.

But what is a literal utterance? Typically, a literal utterance is just taken to be an utterance of a sentence where what is expressed is the meaning of the sentence. "Meaning", as I've mentioned in Ch. 1, is a folk term that will not show up in linguistic theorizing, except in informal characterizations of linguistic research programs. The objection, then, is not intending to require some additional theory of meaning in the folk sense of "meaning", but is thinking of meanings as semantic values, as is not uncommon to do.

For this suggestion to even get off the ground, we need to assume, then, that whatever ends up playing the semantic value role is something that can be communicated.

To simplify the discussion, then, we can suppose for the sake of argument that semantic values are particular propositions.⁵⁶ Literal utterances, then, are utterances in which the proposition expressed is the proposition that is the semantic value of the uttered sentence. Keeping in mind the way we are understanding expressed content, the literal-ity of an utterance will be relativized to an interpreter. What we can say, then, is that an utterance will be literal for an interpreter just in case they interpret the utterance as expressing the proposition that is the semantic value of the uttered sentence. For the sake of readability, I will leave this relativization to interpreter implicit when I say that an utterance expresses some content or other.

Unpacking the rest of our terminology, we can understand (DET_{sys}^*) as $(\text{DET}_{sys}^* *)$:

$(\text{DET}_{sys}^* *)$: For all psychologically normal humans, for all natural languages L , for any utterance U , if U is an utterance of sentence S , S is a sentence of L , U expresses content c , S has semantic value m , $m = c$, and U has contextual features $\phi_1 \dots \phi_n$, then U expresses content c and U 's expressing c is caused by U 's being an utterance of S , S 's having m , and U 's having contextual features $\phi_1 \dots \phi_n$.

The examples of non-literal talk that were problematic for (DET_{sys}) are not a problem for the amended version of those principles since in those cases, the proposition that is the semantic value of the uttered sentence is not identical to the proposition expressed by the utterance. But this should give $(\text{DET}_{sys}^* *)$ a ring of triviality. If a literal utterance is one in which the content expressed, c , is identical to the uttered sentence's semantic value, then it is trivial that such an utterance will express c . It is

⁵⁶A natural thought here would be to understand these propositions as the ones which are expressed via literal utterances. But this, as I argued in Ch. 2, results in problematic circularity. For the sake of argument, here, I will ignore these problems by supposing that semantic value need not be understood in terms of literal utterances.

not trivial, however, that the sentence's having some proposition as its semantic value is a partial cause of utterances of that sentence expressing that proposition. ($\text{DET}_{\text{sys}}^{**}$), then, is not trivial, even if it does include some redundancy.

Examples of certain micro-conventions, however, will still be problematic for ($\text{DET}_{\text{sys}}^{**}$). We can modify the original example from §6.2.3 so that instead of it being the case that three syllable utterances are to be interpreted as meaning that your friend is ready to leave the party, eight syllable utterances are to be interpreted this way. At some point, your friend utters (discreetly so no one else overhears) the eight syllable sentence, "I'm ready to leave the party". Now, you might interpret this utterance in the normal way, without relying on the micro-convention you establish with your friend. On the other hand, you might not. You might be sick of the party yourself, so that you are just paying attention to the number of syllables that your friend's utterances have. Finally they utter a sentence that has eight syllables, and so you interpret them as saying that they are ready to leave the party.

The utterance of "I'm ready to leave the party" is plausibly literal, since you interpret the utterance as expressing the proposition that your friend is ready to leave the party, which, we suppose, is the proposition that is the semantic value of the sentence as well (once we account for the indexical 'I').⁵⁷ However, it seems plausible—or at least psychologically possible—that the fact that you assign some proposition as semantic value to the uttered sentence does not play any causal role in the interpretation process. This is so whether we think of a sentence's having its semantic value as being occurrent—as being the result of a process which occurs upon hearing an utterance—or as simply requiring competence with the relevant sentence.

The possibility of these kinds of micro-convention-based interpretations are prob-

⁵⁷I am ignoring complexities that arise because of indexicals and other context-sensitive expressions here. The example could be altered in such a way that no (traditionally) context-sensitive expressions are involved.

lematic for (DET_{ling}^*) and (DET_{exp}^*) as well. Again, the utterances are literal, but your interpretation of the utterance would seem to be explained, completely, by your recognition that the utterance has eight syllables, and by your belief about the obtaining of the relevant micro-convention. At least, such cases seem *possible*, and that is all that is required here.

Even without appealing to these kinds of examples, I think there are good reasons to think that the restricted determination principles under consideration here are not suitable as guiding methodological principles. Consider, first, (DET_{sys}^{**}) . Even if it is true, it is not particularly interesting. It does not tell us anything about the relationship between semantic processing and the process of interpretation, generally. It tells us, in effect, that sometimes—namely when utterances are literal—semantic processing plays a causal role in interpretation. Given this, I don't think it's plausible even to read (DET_{sys}^{**}) as a potential disambiguation of the determination principle. Likewise for (DET_{ling}^*) and (DET_{exp}^*) .

Further, I think there should be concern that avoiding the micro-convention example might require trivializing (DET_{sys}^{**}) , either because one stipulates the causal connection between semantic value and expressed content in literal utterances, or stipulates that examples like the one I gave above are psychologically impossible. Likewise for (DET_{ling}^*) , which becomes practically equivalent to (DET_{sys}^{**}) once it is restricted to literal utterances. Similar remarks go for (DET_{exp}^*) , assuming that it is to be read in such a way that a sentence can be a constituent of itself.

Now, one might suggest that we shouldn't understand literal utterances in terms of the content expressed by the utterance and the semantic value of the uttered sentence. Perhaps, instead, literal utterances should be understood as sincere utterances. Indeed, sometimes this restriction is made in place of a restriction to literal utterances. But there's nothing insincere about any of the examples of non-literal utterances used in

§6.1, nor the examples involving micro-conventions. Consequently, this response fails to deal with the original objection.

7.1.2 Restricting the Principles to ‘Typical’ Utterances

Another suggestion might be that the determination principle is not meant to apply to all utterances, but just typical utterances. In other words, it might be that the determination principles can accommodate some counterexamples because, most of the time, semantic value determines content in one of the relevant senses.

Here I want to repeat and expand upon a response I made in the previous section. The problem with the restriction to “typical” utterances is that the resulting principles would seem to be uninteresting from the standpoint of generative linguistics. Generative linguistics, again, seeks to describe the nature of the mechanisms causally responsible for our linguistic behavior. It would be an interesting finding, for instance, if (DET_{ling}) turned out to be true. It would suggest that there is a psychologically necessary causal connection between one’s semantic processing and the process of interpretation. This might have interesting ramifications for theorizing about the details of linguistic processing. Consider, for example, the principle of compositionality, which says that, as a matter of psychological necessity, the semantic values of complex expressions are determined by the syntactic structure of the complex, and the semantic values of the constituents of the complex.⁵⁸ Compositionality lends plausibility to the common view that syntactic processing occurs *before* semantic processing, because semantic processing operates on the output of syntactic processing.⁵⁹ Likewise, (DET_{ling}) might suggest similar thoughts about the relationship between semantic processing and interpretation. The process of interpretation might need the output of semantic processing

⁵⁸For arguments to the effect that compositionality should hold as a matter of psychological necessity, see Szabó (2000) and Ch. 4 of this dissertation.

⁵⁹This kind of view is implicit in textbooks like Heim & Kratzer (1998) and Adger (2003).

as a starting point, explaining the causal dependence suggested by (DET_{ling}).

But if we restrict the determination principles to typical utterances, it's hard to see how they would bear on the investigation of the nature of language mechanisms. The problem, basically, is that what is typical of utterances and their interpretation depends heavily on factors that have little to do with the connection between semantic processing and interpretation. Let's suppose, for instance, that typical utterances are non-metaphorical, and that metaphor, generally, has gone out of fashion. There might, at some point, be an explicit movement to bring back metaphor and metaphorical interpretation so that it becomes inappropriate or perhaps just embarrassing not to speak metaphorically most of the time. Assuming that the semantic values of the uttered sentences stay largely constant, the relationship between interpretations of utterances and semantic values of uttered sentences would seem to vary greatly, depending on whether or not metaphor is in fashion. But it is implausible to think that structure of linguistic processing, or the various mechanisms involved in this processing, change based on what is in fashion.

Obviously, the prevalence or absence of metaphor is just one way that what is typical of utterances and their interpretations might be significantly altered. There could also be shifts towards literal or non-literal talk, generally, or towards the adoption of codes and micro-conventions. Such shifts would likely have explanations that appeal to broader societal and cultural facts, and would not be explained by a shift in the general role of semantic processing in utterance interpretation. Basically, the restricted determination principles are simply false, since what is typical of utterances and interpretation is not typical as a matter of psychological necessity (i.e. there would be nothing psychologically abnormal about speakers if much of what is atypical now became typical). If we drop the requirement that the principles hold as a matter of psychological necessity, then the resulting principles lose the status of methodological

principle, and simply become data to be explained (and only in part by appeal to facts about the language faculty).

More generally, facts about how we *typically* use language to communicate do not seem to obviously bear directly on facts about the nature of the mechanisms causally responsible for our linguistic behaviors. This is not to say that such facts have *no* bearing on an investigation into the nature of our linguistic competence, since facts about our linguistic competence surely figure in explanations of what is typical communicative behavior. But, as I mentioned above, much of the explanation will reside outside of generative linguistics. Likewise, it would be odd, in the context of the study of our visual faculties, to take facts about what we typically use our visual faculties to see as illuminating the nature of those faculties. While we are concerned about what happens in the typical case, the atypical cases—visual illusions, e.g.—are often even more revelatory of the nature of our psychology.

In general, then, I think that the strategy of restricting the principles to literal or typical utterances fails. First, the restrictions don't obviously deal with the problematic examples that motivated the restrictions in the first place. Second, and more importantly, such restrictions make the resulting principles strange in the context of a theory which tries to understand the mechanisms involved in language processing generally. There's no reason to think that certain kinds of utterances or interpretations are privileged with respect to how revelatory they are of the relationship between semantic processing and interpretation.

7.2 Restricting the Principles to Utterances of Natural Language Sentences

It might be thought that there is another way to restrict the kinds of utterances under the scope of the determination principles which avoids the worries just raised. In particular, it might be argued that the examples I've used against (DET_{ling}), (DET_{subsent}), and (DET_{sys}) relied on examples of utterances which were not utterances of natural language sentences. Anomalous and ungrammatical strings might not qualify as natural language sentences, and micro-conventions are, strictly, utterances of expressions of impromptu code languages. Generative linguistics, however, does not aim to study all languages that might be used as communication systems, but rather those languages that can be acquired by psychologically normal children in the usual kinds of circumstances—the natural languages.

Let me first address the restriction as it applies to examples of the interpretation of anomalous or ungrammatical speech. I don't think the objection can be used to shield (DET_{ling}) from examples of the interpretation of anomalous speech. 'Sentence' is typically taken to be a syntactic, and not a semantic term, so 'Mary is completely tall, by virtue of its being a grammatical string, is a sentence. Nevertheless, it plausibly lacks semantic value and can be interpreted.

But what about ungrammatical strings like 'Gilgamesh failed often Herbology'? I don't know if there is any standard view of whether ungrammatical strings can be sentences.⁶⁰ Let's suppose, however, that being a sentence is a syntactic property which applies only to certain grammatical strings, and that ungrammatical strings lack this property. We should ask whether the restriction to natural language sentences is appropriate. First, let's suppose that only grammatical natural language strings

⁶⁰This is complicated by the fact that 'sentence' often only seems to appear informally in discussions of syntax.

receive semantic values. In that case, the restricted (DET_{ling}) would, in effect, say that whenever an utterance is an utterance of a sentence that has a semantic value, the fact that the uttered sentence is assigned that semantic value is a partial cause of the interpreter interpreting the utterance the way they do. Expressed content is not causally dependent on semantic value full stop, but it is causally dependent on semantic value when the uttered sentence has a semantic value.

My first reaction here is to say that the resulting principle doesn't qualify as a disambiguation of the determination principle. The idea behind the determination principle is to vindicate the idea that our semantic competence plays some necessary role in our ability to communicate—that the latter depends somehow on the former. But the resulting principle amounts to saying, again, that *sometimes* (most of the time, even) interpretation is partly caused by assignment of semantic value. This falls short of vindicating the idea that a dependence or determination relation obtains between expressed content and semantic value.

My second reaction is to reply that the restricted (DET_{ling}), ($\text{DET}_{subsent}$), and (DET_{sys}) are still false because of the micro-convention examples. In all of the examples given, the utterances were of natural language sentences, and so have a semantic value, but the assignment of semantic value did not play any causal role in the interpretation of the utterance. It might be replied here that even though the utterances in the micro-convention cases were homophonous with English sentences (and so natural language sentences), it was an utterance of an expression which belongs to an impromptu code language, and so is not a natural language expression. However, this reply requires equivocating on 'expression' and 'sentence'. So far, we have been thinking of expressions as phonological entities that are paired with syntactic and semantic properties in the lexicon. Sentences are just phonological structures from which suitably complex syntactic structures can be built. In all of the micro-convention examples I've given,

the uttered sentence would qualify as a natural language sentence on this understanding of ‘expression’ and ‘sentence’. There is nothing preventing the construction of a grammatical syntactic structure which is then assigned a semantic value in the mind of the interpreter. There is similarly nothing preventing the interpreter from interpreting the utterance entirely on the basis of some antecedently agreed upon feature of the utterance or context of utterance.

The objection requires that we think of ‘Jim likes dogs’, in the original example, as being an expression in an artificial language that was created when you and your friend set up the relevant micro-convention. Presumably, we are to think of this utterance of ‘Jim likes dogs’ this way because, informally, the sentence has its meaning in virtue of the micro-convention you create, not because of something about your competence with English. But this is just shorthand for saying that you interpret the utterance in an unusual way. In particular, you interpret the utterance on the basis of features of the utterance other than the semantic value of the uttered sentence. This is no reason to think that the expression uttered is not a sentence of English, however, even if we assume that expressions are individuated by their syntactic properties and semantic values, in addition to their phonological properties. One must, in addition to these extra assumptions about expressions, adopt some version of the determination principle (or the identification of semantic value and expressed content). This assumption is clearly not available in the present context.

As a final attempt, one might claim that when the micro-convention is established, a new lexical entry for ‘Jim likes dogs’ is created in your lexicon. It is syntactically simple and just means that your friend is ready to leave the party. In that case, we might think of your friend’s utterance as an utterance of a natural language expression, but a syntactically simple one. In that case, there’s reason to think that, perhaps, the uttered expression’s having the semantic value assigned to it in your lexicon partly

causes you to interpret it as you do. The problem with this proposal is that the relevant micro-convention applied to any three-syllabled utterance. And it is extremely implausible that upon making your agreement, a massive number of lexical entries were constructed, one for each three syllabled expression that could be constructed using your phonological resources.

7.3 Restricting the Principles to Normal Interpreters

Recanati's (2002) view of expressed content in radical contextualism suggests another way to restrict the principles. Perhaps instead of restricting them to particular kinds of utterances, we ought to restrict them to "normal" interpreters. "A normal interpreter knows which sentence was uttered, knows the meaning of that sentence, knows the relevant contextual facts (who is being pointed to, and so on). Ordinary users of the language *are* normal interpreters, in most situations" (19-20).

First, it ought to be stressed that while the principles should be restricted to psychologically normal language users—since it is their psychology that is primarily under investigation—this use of 'psychologically normal interpreter' does not coincide with Recanati's 'normal interpreter'. An interpreter who, for whatever reason, has some false beliefs about the context of utterance or the speaker, whether they ought to have these beliefs or not, is not thereby psychologically abnormal. We do not expect to find that their language faculty is much different from another interpreter of the same community. Being a normal interpreter in Recanati's sense is not something that comes with the basic linguistic competence that generative linguistics is concerned with, but is an achievement that comes from gathering and employing a wide variety of information about the speaker and the context. The differences between an interpreter with false contextual information and the one with correct contextual information might simply lie at the level of explicit beliefs about the relevant speaker. These differences

are not of particular interest for generative linguistics.

It is worth noting that part of Recanati's motivation for thinking of expressed content this way is to capture the "objectivity" of what is communicated (19). In other words, he takes it to be important that the theory be able to reflect that speakers get it wrong sometimes—that they take an utterance to express some proposition that the utterance does not in fact express. But as I argued in §4, a theory of communication that is continuous with generative linguistics will simply not reflect the possibility that speakers sometimes "get it wrong". This sort of objectivity will be of no theoretical import in the sorts of theories that are under discussion here. We can, of course, acknowledge that interpreters can interpret speakers in a way that departs from how the speaker interprets their own utterance, and that interpreters sometimes interpret speakers in a way that is unreasonable, but these kinds of observations are neither the central explananda of the theory, nor do they immediately figure in explanations of those explananda.

7.4 Do The Arguments Undermine Pragmatics?

A larger concern with my arguments might be something like the following. The arguments stress the messiness and complexity of communication, and the complexity of the relationship between our semantic competence and our communicative abilities. But what does this say for theories of pragmatics which seek to study our communicative abilities and how we employ them in conversation? Are such theories doomed from the start?

The short answer is 'no'. What I have argued is that the relationship between assignment value and utterance interpretation is very complicated, and that the attempt to pair sentence-semantic value-pairs with utterance contents in the context of generative linguistics is confused. Consequently, there is nothing wrong with making claims

about, for instance, how utterances of particular sentences are typically interpreted by interpreters of a given speech community. These facts will often be of great interest to those studying various sociological, socio-linguistic, and cultural trends (for example). Generative linguistics, however, does not study these trends, and so they would seem to bear on that research program only tangentially at best.

Additionally, while I've argued that the relationship between assignment of semantic value and interpretation of utterances is complicated and unsystematic, it does not follow that there's nothing systematic to say about, say, the kinematics of conversation. Even if the relationship between semantic processing and interpretation is unsystematic, it could be that, in conversation, interpretations of utterances often systematically depend on the interpretations of previous utterances, for example. There might also be systematic patterns of behaviors which are causally downstream of utterance interpretation—patterns of acceptance and rejection, and so on, which are amenable to systematic theorizing. These are the kinds of patterns of behavior that are the purview of the active and successful research programs in pragmatics.

8 The Relationship Between Semantic Value and Content

I've argued that no version of the determination principle can serve as a guiding methodological principle in generative linguistics. The various disambiguations of the principle are either too trivial to meaningfully constrain linguistic theories, or else they are false and conflict with other standard theoretical assumptions in linguistics. This, of course, is not to deny that our semantic competence plays an important role in communication. This assumption is what was meant to motivate the determination principle in the first place. Nothing I've said rules out the likely possibility that, for

instance, *most of the time*, the assignment of semantic value to an uttered sentence partly causes interpreters to interpret utterances of that sentence as they do.

But if semantic value is somewhat independent from utterance content, as I have been suggesting, it would be nice to say a bit more about their relationship, and about the role of our semantic competence in communication. In this section, I'd like to sketch a possibility which would help flesh out this role.

8.1 Semantic Competence As A Perception-Like Ability

As a starting point, consider the communicative role of assignment of semantic value if we identify the semantic value and expressed-content roles. On that view, upon hearing an utterance, an interpreter compositionally assigns a semantic value to the uttered sentence, and their doing so is identical to their interpreting the utterance a particular way. In that case, assuming that the contents that utterances tend to express are propositions, then semantic values will tend to be propositions as well.

The picture is too simple, we know, because of the need to account for context-sensitive expressions like indexicals. Semantic values, it is assumed, are compositional and remain fixed across contexts of utterance. But of course, what is expressed by an utterance with an indexical depends not just on the compositionally determined semantic value of the uttered sentence, but also on features of context. On this view, semantic value is not identified with utterance content, but it may very well be understood in terms of it. On Schiffer's view, considered in Chapter 2, for instance, the semantic value of an sentence is a function from contexts to the proposition that is expressed by literal, unembedded utterances of the sentence. Assignment of semantic value and interpretation of an utterance are not identical on this view, but at the very least (DET_{sys}) , (DET_{ling}) , and (DET_{exp}) would all obtain. Upon hearing an utterance of a sentence, an interpreter assigns a semantic value to a sentence, and then their

doing so is a partial cause of their interpreting the way they do. Furthermore, the interpretation they end up with will depend systematically on the semantic value of the uttered sentence. Semantic values, here, will very often not be propositional, but assignment of semantic value will be a necessary step towards attributing propositional content to an utterance.⁶¹

On my view, not only is semantic value not to be understood in terms of utterance content, but none of (DET_{sys}), (DET_{ling}), or (DET_{exp}) obtain. In that case, we can be neutral on whether semantic values are propositional or sub-propositional.⁶² Indeed, it could be that there really are two semantic value theoretical roles that need to be filled—one which is propositional, and one which is sub-propositional. Assignment of semantic value, whether semantic value is propositional or not, however, would not suffice for an interpretation of an utterance.

What, then, does assignment of semantic value *do* for us as communicators? Here, I think we need to consider the kinds of explanations that facts about semantic value contribute to in semantic theorizing. For instance, at the very least, facts about semantic value are taken to help explain certain entailments and semantic anomaly.⁶³ Facts about entailment and anomaly, in turn, help to explain speakers' judgments of unacceptability and entailment that are used as a primary source of evidence for thinking that a certain construction is anomalous, or that an entailment relation obtains.

But what are speakers doing when they judge 'Mary is completely tall' to be unacceptable, or more generally, defective? Typically, there is no pretense by the linguist that they are making some communicative speech act when they utter a sentence they are trying to collect data on, so it is implausible that subjects take a linguist's *utter-*

⁶¹Here I am thinking of semantic as being like Kaplan's (1989) "character".

⁶²It might be thought that semantic value must be propositional if facts about semantic value are to explain semantic entailments. I think that facts about semantic values do explain entailment facts, but as I will (in effect) argue in Chapter 5, this does not require that semantic values be propositional.

⁶³How facts about semantic value figure in these explanation will be the topic of discussion in Chapter 5.

ance (qua speech act) to be unacceptable when they ask about anomalous sentences. Perhaps what they are doing when they judge an anomalous sentence to be unacceptable is judging that utterances of such a sentence would be unacceptable to make. But why should we think that? Whether or not an utterance is acceptable or not depends heavily on context, and it wouldn't be difficult to come up with a context where an anomalous sentence is interpreted in some way or another without any sort of unacceptability (again, qua speech act).

My suggestion, then, is that speakers are (or at least often are) offering judgments about the expression types themselves. Or, rather, their judgments are best explained by appealing to properties of expression types at least, and so their judgments bear fairly directly on the semantic values of expression types. Consider, again, the ungrammatical 'Gilgamesh failed often Herbology'. When we judge the sentence to be somehow defective, we certainly need not be thinking about potential utterances of such a sentence.⁶⁴ Rather, there is a particular phenomenology associated with hearing or reading it. Even if we can make sense of an utterance of that sentence (and again, this wouldn't be difficult), there is a residue of defectiveness that accompanies encountering it. Even in a case where we can easily interpret an utterance of the ungrammatical sentence, hearing an utterance of it might at least result in momentary perplexity. Speakers are sensitive to this feeling of defectiveness or perplexity, and can report it to linguists upon being confronted with ungrammaticality. Linguists, in turn, can take the report as being evidence for the linguistic competence that is causally responsible for the report and accompanying phenomenology.

The situation is similar for anomalous sentences like 'Mary is completely tall'. There is a similar feeling of perplexity and defectiveness, and we judge the sentence to be unacceptable in some sense. In this case, the linguist posits that the source of this

⁶⁴For ease of discussion here, I refer to the ungrammatical string as a sentence. Whether it is properly considered a sentence or not is not of importance here.

reaction is semantic, rather than syntactic. The speaker, herself, likely is not and need not be sensitive to the cause of the feeling of perplexity. Nevertheless, if the semanticist's explanation of anomaly is good, then the assignment of semantic value is at least partly responsible for the feeling of perplexity. In that case, assignment of semantic value is presumably also partly responsible for the acceptability (i.e. non-defectiveness) of sentences that are not anomalous.

But assignment of semantic value, presumably, does not merely determine whether a sentence is “semantically acceptable” or not. If that's all semantic value did for us, then there would be little sense in thinking of it as being proposition-like, or as having much to do with meaning at all. Indeed, we might ask *why* it is that anomalous sentence perplex us. Informally, the answer might just be that—aside from how we might interpret utterances of anomalous sentences—we can't make sense of them. Informally, again, we don't know what they mean.

A possibility, then, is that assignment of semantic value constitutes the mental state which we might informally call “hearing a sentence as meaning such-and-so”.⁶⁵ Consider, for instance, when your friend utters a sentence which has some ambiguity, so that there are two available readings of the sentence, but one such reading is the obviously intended one. Suppose, for instance, you live in New York and your friend says to you, “Let's call the senator from Texas and give him a piece of our minds!” Out of context, it's unclear whether you are calling the senator who represents Texas, or whether the call is to be made from Texas. Since you live in New York, you have no trouble interpreting the utterance—i.e. grasping which proposition your friend intends for you to believe. Nevertheless, you might initially ‘hear’ the utterance as having the unintended meaning, either coincidentally or perhaps because the unintended reading

⁶⁵This sketch is inspired by McDowell's (1981, 239-240) suggestion that when one has learned a language, one acquires a kind of perceptual ability to perceive what one says in an utterance. See also Heck (2007, §5).

is humorous, etc. The trouble with anomalous sentences is that we don't hear them as meaning anything at all.

I am not in a position to adequately defend or fully characterize the 'hearing an expression as meaning such-and-so' relation. I want to leave open, for instance, whether it is a relation that speakers bear to propositions or something sub-propositional.⁶⁶ I also want to be clear that I use 'hear' and 'meaning' in labeling the relation for illustrative convenience only. I do not have any view about the relationship between this relation and ordinary uses of the English 'means'—the relationship is presumably very complex and uninteresting from the standpoint of generative linguistics.

We can say a bit more about the relation, however, and the role it plays in communication. First, the relation does not suffice for belief, nor for the interpretation of an utterance. Consequently, it does not follow from the fact that one assigns a semantic value to an uttered sentence that the utterance expresses any particular content, even if semantic values are propositions. It is thus more like entertaining a thought in this respect.

However, hearing a sentence as meaningful, as the choice of terminology suggests, is more like a perceptual ability than it is like entertaining a thought. I want to quickly elaborate on this idea, though I want to be clear that whether or not our semantic competence constitutes (or partly constitutes) a perceptual ability is beyond the scope of this work. Further, whether it does or not will likely be of little interest to linguistics, since what is of interest is what the nature of our linguistic abilities are, not the extent to which these abilities are to be considered as being of the same kind as other abilities.

First, like many of our perceptual abilities, hearing a sentence as meaning such-and-so occurs automatically and more or less instantaneously. Second, we have limited

⁶⁶If it is a relation between speakers and propositions, then it might be that semantic value only partly constitutes one's bearing the relation to a proposition, since semantic values might be sub-propositional. This would be the view, for instance, of Glanzberg (2014) and Pietroski (2010), who argue that full propositional meaning results only with input from extra-semantic faculties.

control over how we hear a sentence. In some cases, it takes some effort to hear an ambiguous sentence in a particular way. In many cases, we simply can not hear or “read” a sentence in any number of ways. In other cases, we might be able to go back and forth between readings, as we might (more or less) voluntarily switch between perspectives on the Necker Cube.⁶⁷ Third, the development of the ability to hear sentences as meaningful is something that develops more or less automatically in childhood.⁶⁸ The linguistic ability, of course, takes much longer to reach maturity than the paradigm perceptual abilities, but it is noteworthy that it does not require special training.

Fourth, the relation is perception-like in that, while it does not suffice for belief, it typically instantaneously and effortlessly causes particular sorts of beliefs. For instance, our visual experience typically causes us to have beliefs about not only what we see, but about what is around us. In the case where we hear an uttered sentence as meaning ϕ , we effortlessly and automatically move to the belief that we hear the sentence as meaning ϕ , and also the belief that the speaker means ϕ in making their utterance. As with vision, the belief-formation process can be overridden. Many of the examples I’ve used to argue against the determination principle are cases where, because of other beliefs that we have, we override the tendency to interpret an utterance as we hear it.

Finally, hearing a sentence as meaning ϕ is a strong, but defeasible source of justification for the beliefs that it tends to effortlessly cause. In the case of vision, visual experience is typically an excellent source of justification for thinking that the world is as it appears. Very often, however, we know that certain illusions obtain, or that we are in conditions where illusions are likely to occur. In those cases, the evidence

⁶⁷For example, in ‘I wonder who the men expected to see them’ we can hear ‘them’ as referring to the men or to some other specified group. In ‘The men expected to see them’, we cannot hear ‘them’ as referring to the men. The examples are from Chomsky (1986, 8).

⁶⁸See, e.g., Guasti (2002) for a textbook approach to language acquisition in generative linguistics. I should be clear here that the relevant linguistic abilities pertain to spoken language, and not written language.

provided by our visual experience is defeated, we have good reason to believe that the world is not as it appears. Similarly, the fact that you hear an uttered sentence as meaning ϕ is typically good reason to interpret the speaker as meaning ϕ . Arguably, however, more care is needed in the linguistic case than in the visual case. Linguistic “illusions” are rampant in natural language, given not only its inherent vagueness and ambiguity, but also the flexibility of communication. Further, hearing an uttered sentence as meaning ϕ will only be good evidence that the speaker means ϕ to the extent that their linguistic background and conventions of communication are similar to the interpreter. When we speak to someone who has an accent that we are not used to, or speak an unfamiliar dialect, we are naturally (and reasonably) more hesitant to interpret utterances in accordance with how we hear them. Sometimes, we can ignore our hearing sentences in a particular way altogether, as in the micro-convention cases raised in earlier sections.

Our semantic competence, then, performs at least two main functions in communication on the kind of view I am sketching. First, it very often is a good source of justification for our interpretations of others—i.e. for our beliefs about what others mean, what they think, given what they’ve said, and so on.⁶⁹ It is thus very often crucial for our coming to know about what speakers mean and what they think, even if being semantically competent is neither necessary nor sufficient for obtaining this knowledge.⁷⁰

Second, our semantic competence greases the wheels of communication by making us much more efficient as communicators. Much of the time, we can quickly and effortlessly move from hearing a sentence as meaning ϕ to interpreting an utterance as

⁶⁹Just how this justificatory link works, though, will depend on the nature of the relationship between utterance-content in the theoretical sense, and folk-content-expression (see §3.4).

⁷⁰That it is not necessary can be demonstrated by considering micro-convention cases where you friend speaks in a language you are unfamiliar with. Nevertheless, you can recognize when their utterance has only three syllables.

expressing ϕ . The processing that results in hearing the sentence a particular way is typically extremely rapid and automatic, and the move to interpretation is similarly quick and requiring little effort on the part of the speaker. Consequently, interpretation can occur rapidly and without effort much of the time. Sometimes interpretation does take some time and effort, however. In those cases, the fact that a speaker utters a sentence which you hear as meaning ϕ might be an important starting point in arriving at an interpretation of the utterance.⁷¹ In some cases, as I've stressed, assignment of semantic value will not play either of these roles, and hence, no non-trivial versions of the determination principle obtain. However, given that it does so much of the time, this vindicates the underlying motivation for the determination principle—namely, that assignment of semantic value (and our semantic competence, generally) is extremely important for our ability to communicate as we do.

I want to stress, once again, that this sort of view is merely a sketch of the role that semantic value might play in communication that is consistent with the rejection of the determination principle, and the plausible assumption that our semantic competence plays an important role in communication. It also does justice to the role that semantic value plays in semantic explanations of anomaly and certain entailments, and to the flexibility and complexity of communication.

8.2 Mental Content and Utterance Content

The sketch is consistent with semantic projects that are more self-conscious about their place in generative linguistics. I want to briefly describe why this is so, because it might not be obvious that such projects are consistent with my arguments, and it also raises

⁷¹Semantic value, in that case, would bear some resemblance to Cappelen & Lepore's (2005) minimal content with respect to the role it would play in communication. Assignment of semantic value, on the view I am sketching however, would not constitute anything like an expectation that one means ϕ in uttering some sentence.

an important distinction between utterance content and mental content.

I have in mind here the sort of semantic research program typified by Jackendoff's "Conceptual Semantics".⁷² On Jackendoff's view, the semantic value and content roles have the same realizers, which he refers to as "concepts". The aim of his research program is to describe the nature of concepts and conceptual structure. Thus, he should not be understood as endorsing some version of the determination principle when he writes, "The fundamental goal of Conceptual Semantics is to describe how humans express their understanding of the world by means of linguistic utterances", and "the theory aspires to describe the messages that speakers intend for their utterances to convey" (2011, 81). The aim here is not to describe which messages attach to particular expressions, but rather to describe the nature of the messages—i.e. concepts—themselves. Since he also thinks of concepts as being constituents of thoughts, the project also seeks to describe the (narrow) contents of thoughts.⁷³

As I have stressed throughout, nothing I've said denies that the realizers of the expressed-content and semantic value roles be the same kinds of objects. On Jackendoff's view, when a speaker hears an utterance of some sentence they understand, the semantic component of the language faculty constructs a semantic structure.⁷⁴ This structure, which, in our terms, is the semantic value of the sentence, can also play role of the content of a thought. When we communicate with one another, we try to express these same sorts of objects to our interlocutors.

On the sketch of the relationship between semantic competence and communication I gave above, the construction of one of the semantic structures in the semantic faculty that Jackendoff's theory posits would, in effect, constitute (or cause) our hearing a

⁷²Jackendoff (2011). See also Jackendoff (1990, 2010).

⁷³See, e.g., Fodor (1987) for a discussion of narrow content.

⁷⁴I am glossing over details here. For instance, Jackendoff does not recognize a separate, autonomous "semantic faculty". Rather, he posits a conceptual faculty, which interfaces with memory, perceptual systems, etc., and is not dedicated solely to the interpretation of language. See, e.g., Jackendoff (1990, 2010, Ch. 3).

sentence as meaning ϕ . From here, again, it is a complicated process which results in our interpreting an utterance in a particular way, which will be constituted by the instantiation of a possibly distinct semantic structure which will constitute the content of the utterance. Semantic processing and the process of interpretation terminate in the instantiation of the same *kinds* of objects—conceptual structures in Jackendoff’s sense—but there’s no reason to think that the particular conceptual-structure-types that are tokened will be the same.

The upshot here is just that if we think of mental contents as being thoughts or the components of thoughts, there is plenty of room to hold that semantic value determines mental content, or perhaps just is identical to it.⁷⁵ If the semantic value and mental content roles are both realized by Jackendoff’s conceptual structures, then mental contents will just be attached to expressions via their lexical entries. In that case, it would seemingly be trivial that the semantic value of an expression will determine, in some sense, the mental content that is associated with it.

But this association is not to be understood in terms of utterance-content. Not only can we not assume that utterances will be interpreted as having a sentence’s associated mental content, we can’t even assume that all utterances of particular sentences will trigger the construction of a particular mental content (i.e. semantic value). Cases of micro-conventions in which a speaker recognizes the uttered sentence, but only pays any attention to prosodic features of the uttered sentence tell against this view. In that case, the speaker recognizes the uttered sentence as being, say, an English sentence, but if they are keenly concentrating only on prosodic features, it might be that they fail to even hear the sentence as having meaning.

Whether or not semantic value is identical to or determines mental content, in some sense, is a substantial and difficult empirical question, which requires better

⁷⁵It’s not immediately obvious that we could identify the two, since semantic value might just contribute part of the mental content that is associated with expressions.

understanding the nature of semantic competence, and the role it plays in our broader cognitive architecture. Thus, it can only be answered by doing semantics. We can, in our current state of understanding, however, reject all forms of the determination principle as guiding methodological principles in linguistics. These principles aim to tie semantic value tightly to our communicative abilities—to our ability to interpret utterances and convey thoughts in our utterances. If we identify semantic value with mental content, then this will only mean that the relationship between mental content and utterance content is as messy as is the relationship between semantic value and utterance content.

Chapter 4

Compositionality as Weak

Supervenience: Why

Compositionality Doesn't Require

Truth-Conditional Semantics

Chapter Overview: There is a common, a priori argument to the effect that semantic theories must be truth-conditional to be adequate. The argument goes as follows. First, a semantic theory, to be adequate, must be compositional. Second, the only semantic theories which are compositional are truth-conditional theories. Therefore, the only semantic theories that can be adequate are truth-conditional theories. The trouble with this argument is that it is very rarely made precise just what the principle of compositionality comes to, and it is even more uncommon to see arguments for one version or the other.

The aim of this chapter, then, is to argue for a particular reading of the principle of compositionality in the context of generative linguistics, and then to assess whether

it can only be satisfied by truth-conditional semantic theories. As a starting point, I consider a proposal by Szabó (2000), who argues that the principle of compositionality should be understood as the claim that, as a matter of psychological necessity, the meaning properties (i.e. semantic values) of complex expression strongly supervene on the syntactic structure of the complex and the meaning properties of its constituents.¹ I argue that if we understand compositionality Szabó's way, then compositionality can play no role in explanations of the acquirability of natural languages, because it makes these explanations circular. This, in turn, would undermine the primary motivation for thinking that natural language is compositional, and would thus undermine the importance of the principle in natural language semantics. Thus, even if Szabó's reading of the principle best accords with theorists' intuitions about what sorts of languages are compositional—as he claims it does—there is good reason to reject that reading.

I defend the position that compositionality should be understood, instead, as the claim that, as a matter of psychological necessity, the meanings (semantic values) of complex expressions *weakly* supervene on the syntactic structure of the complex and the meaning properties of its constituents. This understanding of the principle of compositionality, I argue, better accords with the explanatory role that compositionality is expected to play in linguistics. Understood this way, however, the demand that semantic theories be compositional fails to rule out paradigmatically non-truth-conditional approaches to semantics. Consequently, the demand for compositionality does not force us into thinking that semantic theories must make use of a notion of truth or truth-conditions.²

¹I use 'meaning', throughout, to refer to semantic value.

²§1-§5 of this chapter appear in Napoletano (2015).

1 Introduction

There is no consensus, or even standard understanding of the principle of compositionality in philosophy and linguistics. Often, while discussing how the principle bears on various philosophical or empirical projects, authors will simply acknowledge this and leave implicit exactly what the principle comes to. Even rarer than explicit formulations of what one means by ‘compositionality’ are arguments to the effect that we ought to understand the principle of compositionality in one way or another. One such argument has been given by Zoltán Gendler Szabó (2000), however. According to Szabó, we ought to understand the principle of compositionality as the claim that in natural language, the meanings of complex expressions *strongly* supervene on the meanings of their constituents and how the constituents are combined.³ The reason, he argues, is that this reading of the principle best accords with theorists’ intuitions about what it is for a language to be compositional.

While I agree that Szabó’s reading of the principle is largely in accord with these intuitions, I will argue that we have overriding reasons to reject both Szabó’s reading of the principle, and also his proposed methodology for fixing the meaning of ‘compositionality’. In particular, if we adopt Szabó’s reading, then the fact that natural language is compositional will play no role in explanations of its acquirability. But the primary motivation for claiming that natural language is compositional is that this is necessary to explain why it is that we are able to acquire natural languages. Whatever compositionality is, it should play some role in these explanations. In other words, ‘compositionality’ is, first and foremost, a theoretical term which is recruited in the context of linguistic theories to denote a property which plays a particular explanatory role in these theories.⁴ Thus, while appealing to theorists’ intuitions about

³Szabó provides some support for this view in his (2012a) as well.

⁴Historically, of course, the principle of compositionality has also played an important role in philosophical debates in the philosophy of language. As with Szabó (2000), my concern here is with

compositionality might be useful in attempts to fix the meaning of the principle of compositionality, we should at least take such attempts to *also* be constrained by the explanatory demands that are placed on the principle in linguistic theorizing.

Insofar as Szabó's reading of the principle fails to meet this constraint, I conclude that we ought to reject it, and that we should instead think of the principle as the claim that in natural language, the meanings of complexes *weakly* supervene on the meanings of their constituents and how they are combined. On this reading, the fact that a language is compositional does partly explain the acquirability of natural language, though it explains less than is often assumed.

In §3 I give a general sketch of the role that compositionality is usually taken to play in explanations of the acquirability of natural language. I will take it as a desideratum on an account of compositionality that it be able to play some role in these explanations. In §3, I present Szabó's strong supervenience reading of the principle of compositionality, and consider some objections to the weak supervenience reading which show that on that reading, compositionality would explain less about acquirability than is often assumed. In §4, I argue that these objections do not motivate Szabó's strong supervenience reading, and that in fact, his reading precludes compositionality from playing *any* role in explanations of the acquirability of natural language because it makes these explanations circular. In §5, I argue that the weak supervenience reading—despite explaining less about the acquirability of natural language than is often thought—is actually preferable to the strong supervenience reading, since it still has a place in explanations of acquirability. I then consider what additional assumptions about natural language and the psychology of speakers are required in order to fill in the explanatory gap that the weak supervenience reading leaves us with. Finally, in §6, I argue that the arguments from compositionality to the inadequacy of versions of the principle that appear mostly in the context of natural language semantics. See Szabó (2012a) for discussion of related principles and their roles in various other theoretical pursuits.

non-truth-conditional approaches to semantics fail.

2 Compositionality and Linguistic Explanations of Acquirability

It is widely held by both philosophers and linguists that natural language is compositional, and that this fact partly explains what it is about natural languages that makes them acquirable by humans as first languages—i.e. what makes them *natural*.⁵ In particular, the fact that natural language is compositional is usually taken to explain the fact that competent speakers of natural language possess certain linguistic abilities—the ability to produce and understand a large number of expressions which they’ve never encountered, for instance. Since possession of these abilities seems to be universal amongst competent speakers of natural language, it calls out for linguistic explanation, and this explanation is typically taken to consist at least partly in the fact that natural languages are compositional.

Thus, the fact that a language is compositional figures in what I will call a ‘linguistic explanation of acquirability’ (LEA) for that language. I will use ‘acquirable’, throughout, as shorthand for ‘acquirable (largely in early childhood) as a first language, given typical human psychology and typical exposure to linguistic stimuli’. The longhand is, in turn, a standard way of understanding what it is to be a natural language within the generative linguistics framework, which I will be assuming in the background of this paper.⁶ The LEA for a language is thus what constitutes the fact that the language is

⁵Of course, neither claim is completely uncontroversial. See, e.g. Travis (1985), Pelletier (1994), and Fodor (2001) for arguments against the claim that natural language is compositional. Horwich (1998) claims that natural language is compositional, but that this fact does not explain the features of speakers’ competence that most take it to explain.

⁶For detailed presentations of this framework, see Chomsky (1986) and Ludlow (2011).

a natural language.⁷ It should also be noted that ‘first language’, here, is not simply intended to imply that the language is temporally the first to be acquired, but that the language in question serves as the speaker’s primary language.⁸ More specifically, this is the language that a speaker “knows” or “has” in virtue of their language faculty having settled into a more or less stable state at some point during their linguistic development. Thus, while a child is still in the process of acquiring their first language, and their language faculty is undergoing considerable change, they may be said to “know” or “have” any number of languages—corresponding to the varying states of their language faculty—which are not natural or acquirable in the sense of those terms that I am employing here.

The possible stable states of the human language faculty, given typical human psychology and linguistic conditions, are what correspond to the natural languages. As it turns out, there are considerable limits on which such states of the language faculty are psychologically possible for humans, and thus there are considerable constraints on what it takes to be a natural language. One of the primary aims of linguistic theorizing, then, is to determine which linguistic (phonological, syntactic, semantic, etc.) features are constitutive of natural languages, and thus to be able to determine what the space of psychologically possible first languages is like.⁹ The LEA for a language will simply consist in the fact that it possesses all of these features. We can think of artificial

⁷We could just as well talk about ‘linguistic explanations of naturalness’ and leave ‘acquirability’ out of it. I incorporate ‘acquirability’-talk both because ‘natural language’ is a technical notion which does not wear its meaning on its sleeve, and because of the awkwardness of ‘naturalness of natural language’.

⁸Szabó (2000, 479) describes natural languages as being, in addition, languages that are “suitable for the expression and communication of a wide range of thoughts”. I intend for this to be included in what it is to be a speaker’s “first” or “primary” language.

⁹Languages in this context are abstract objects which have their linguistic properties essentially (see, e.g., Ludlow (2011, Ch. 2), and George (1989)). They can be thought of as consisting of a lexicon (a pairing of clusters of phonemes—the simple expressions—with their syntactic and semantic properties), syntactic rules for combining items of the lexicon into complexes, and semantic rules for deriving the meanings of complexes. Included amongst the natural languages, then, are languages which no human has ever spoken, and which might never be spoken.

languages, by contrast, as just being those languages which lack one or more of these features, and which thus, as a matter of psychological necessity, are not acquirable as first languages under normal human conditions (even if they are “learnable” or “acquirable” in the ordinary senses of those terms).¹⁰

What is the principle of compositionality and how does it figure in explanations of acquirability? For now, we can think of the principle of compositionality in terms of (C):

(C) The meaning of a complex expression is determined (solely) by the meanings of its constituents and by its structure.

This particular formulation is rough and ambiguous several different ways,¹¹ but it will serve as a placeholder until more precise formulations are introduced in the next section. Typically, the principle is taken to explain, and to be motivated by one or more of the following claims about the competence of speakers:

(PROD) Competent speakers of natural languages can understand expressions which they have not yet encountered.¹²

(INF) Competent speakers of natural languages understand an infinite number of expressions.¹³

(SYS) If a competent speaker of a natural language understands complex expressions e_1, \dots, e_n , then they can understand any complex expression which are among the

¹⁰It follows that a child that is exposed only to some artificial language will not acquire that language, though they will likely acquire a language which is perhaps superficially similar, but which meets the various constraints that all natural languages must meet. This claim is borne out by the existence of creoles, which children acquire in just these kinds of circumstances.

¹¹See Szabó (2000; 2012a; 2012b), and Pagin and Westerståhl (2010a) for discussion of the various disambiguations.

¹²(PROD) figures as a premise in arguments that are variously called arguments from ‘productivity’, ‘novelty’, or ‘creativity’.

¹³Arguments for compositionality utilizing (INF) are usually called arguments from ‘unboundedness’ or ‘acquirability’. See, e.g., Davidson (1965). Note that on my use of ‘acquirable’ in this paper, to say that a language is acquirable is not to imply anything about the number of expressions it contains.

constituents of e_1, \dots, e_n , and which is the result of applying only syntactic rules which are used to construct e_1, \dots, e_n .¹⁴

None of those claims, if true, would *necessitate* that natural language is compositional, since what they require is just that speakers have some way of arriving at the meanings of complexes. This is compatible with the claim that the meanings of complexes is determined by their structure, the meanings of their constituents, and something else—the intentions of speakers or features of the context, for instance. However, part of the best explanation for (PROD), (INF), and (SYS) is (SUFF):

(SUFF) To be able to understand a complex expression in a natural language, it is sufficient to understand its constituents and how they are combined in the complex.

(SUFF), in turn, is at least partly explained by the fact that the meanings of natural language complexes are determined by the meanings of their constituents and how they are combined—i.e. the compositionality of natural language. Thus, the fact that natural language is compositional will be part of the best explanation for the above facts about the competence of natural language speakers. Absent an account of linguistic understanding, and without saying more about what sense of ‘determine’ is at play in (C), this does not make precise *how* the compositionality of natural language figures in an explanation of (SUFF). The rest of the paper will be partly devoted to filling in these details. The aim here, though, is just to sketch the general location of the principle of compositionality in explanations of acquirability—the fact that a language is compositional should figure *somewhere* in these explanations.

What’s called the ‘systematicity of natural language’ is often thought to consist in the fact that one can understand a natural language complex if and *only if* one

¹⁴See, e.g., Szabó (2012a, 76; 2012b) and Fodor & Lepore (2001a).

understands its constituents and how they are combined. I will call a language in which to understand a complex, it *suffices* to understand its constituents and how they are combined, ‘weakly systematic_{PSY}’.¹⁵ More precisely,

a language L is weakly systematic_{PSY} if and only if, if a speaker of L understands expressions e_1, \dots, e_n in L (simple or complex) and syntactic rules $SynR_1, \dots, SynR_n$ of L , then they can understand any complex expression in L which has only members of $\{e_1, \dots, e_n\}$ as constituents, and which is composed only via syntactic rules which are members of $\{SynR_1, \dots, SynR_n\}$.¹⁶

(SUFF) should be understood as the claim that all natural languages are weakly systematic_{PSY}. The weak systematicity_{PSY} of natural language, then, is part of the best explanation for (PROD), (INF), and (SYS). Weak systematicity_{PSY} is not a linguistic property of a language, but a relational psychological property, and so the fact that a language is weakly systematic_{PSY} will not itself figure in LEA’s. However, the weak systematicity_{PSY} of natural language is a fact that calls out for linguistic explanation, and this is typically taken to at least partly consist in the fact that natural language is compositional. This, again, is the motivation for the principle of compositionality and the general place of the principle in linguistic explanations of acquirability.

My aim here is not to assess these or related abductive arguments for the claim that natural language is compositional.¹⁷ I will assume that they do support the claim that

¹⁵The ‘PSY’ subscript is needed to differentiate this relational psychological property from its linguistic counterpart, which I introduce in §5.

¹⁶The appeal to syntactic rules here instead of mere syntactic structure captures the fact that the syntactic structure of a complex is derived from syntactic rules and syntactic features of the constituent expressions. This explains why someone who understands ‘and’ will be able to understand a sentence with ten conjuncts, even though they may have never encountered a sentence with exactly that structure before. I take it that talk of syntactic rules just makes explicit by what most have in mind by ‘structure’ or ‘the way the constituents are combined’. See, e.g., Robbins (2005), Pagin and Westerståhl (2001a), where the appeal to derivation and syntactic rules in the principle of compositionality is explicit.

¹⁷For critical discussion of these arguments, see Szabó (2012a; 2012b), Pagin and Westerståhl

natural language is at least largely compositional, and thus that the compositionality of natural language explains features of natural language competence. In particular, I will assume that (PROD) is sufficient to motivate the claim that natural language is compositional. Both (INF) and (SYS) are stronger claims which entail (PROD), assuming the finitude of speakers, that they understand some complexes, and that natural language syntax is recursive. Neither is required to motivate the compositionality of natural language, and so I ignore them the rest of the way.¹⁸

Going forward, I will simply assume that natural languages are weakly systematic_{PSY} as well—i.e. that (SUFF) holds—and that this is part of the best explanation of (PROD), the productivity of natural language. Since the weak systematicity_{PSY} of natural language is at least partly explained by its compositionality, I take it that the compositionality of natural language partly explains (PROD), and is partly constitutive of the acquirability of natural language—i.e. the property *being compositional* is partly constitutive of *being a natural language*. My primary concern will be with whether (and to what extent) the compositionality of natural language can explain its weak systematicity_{PSY} if we understand the principle, with Szabó, as the claim that in natural languages, the meanings of complexes strongly supervene on the meanings of their constituents and how they are combined.

(2010b), and Cohnitz (2005).

¹⁸(SYS)—the claim that natural language competence is systematic—is far more controversial than (PROD), for instance, because it is taken to motivate not just the compositionality of natural language, but also its “reverse compositionality”—i.e. that the meanings of the constituents in a complex are determined by the meaning of the complex and its syntax. For discussion, see Fodor and Lepore (2001a; 2001b), Pagin (2003), Patterson (2005), and Robbins (2005).

3 Motivation For the Strong Supervenience Reading

As Szabó (2000) notes, most agree that (C) expresses the principle of compositionality. It has also been pointed out by a number of authors that (C) is itself imprecise and ambiguous, and different authors have different views about how best to understand it. A common reading is in terms of (F):

(F) The meaning of a complex expression in a natural language is a function of the meanings of its constituents and of its structure.

What (F) guarantees is that, in every natural language, there are no two complex expressions such that they have synonymous constituents, identical syntactic structure, and yet differ in meaning. We can put this by saying that in every natural language, there is a function from the constitution properties of complexes to their meaning properties. For an expression to have a certain meaning property in a language is just for it to have a certain meaning in that language. For it to have a certain constitution property in that language is for it to have “some constituents with such-and-such meanings combined in such-and-such way” in that language (2000, 495). Following Szabó, we can recast (F) in terms of (WS), which is the claim that in natural languages, the meaning of a complex “weakly supervenes” on its constitution property:

(WS) For all natural languages L , for any meaning property M and any complex expression e in L , if e has M in L , then there is a constitution property C such that e has C in L , and if any complex expression e' in L has C in L then e' has M in L (2000, 497).¹⁹

¹⁹As Szabó (2000, 496-497) notes, (WS) has the general form of a weak supervenience claim, understood along the lines of Kim (1984, 64). One crucial difference is that in (WS), the necessity operators are treated as quantifying over *natural languages* instead of possible worlds.

I will use ‘the principle of compositionality_{WS}’ to refer to (WS). A language L is compositional_{WS}, then, if and only if the meaning properties of L ’s complex expressions weakly supervene on the constitution properties of the complexes—if and only if there is a function from the constitution properties of L ’s complexes to the meaning properties of L ’s complexes.

Szabó offers two objections to this reading of the principle of compositionality. First, it leaves open the possibility that a complex expression in some natural language L_1 has the same constitution property as some complex in some distinct natural language L_2 , and yet the two differ in meaning. But, he says, this is counterintuitive; we think compositionality is supposed to apply cross-linguistically, such that it rules out this possibility.

Second, suppose that the meaning of a sentence in some language changes, but the syntactic rules remain fixed. Szabó says that in that case, we would suppose that the meaning of the sentence changed because one of the meanings of the constituents had changed. Further, he says that we would make this assumption because we think that natural language is compositional. But (WS) is too weak to support that explanation; so long as there is still a function from constitution properties of complexes to their meaning properties, the language will be compositional on the (WS) reading. It would not have to be the case that the meanings of one of the constituents changes in order for the language to remain compositional. Thus, since we would think that the meaning of at least one constituent had changed, and since we would do this because we think the compositionality of natural language requires it, this shows that we must have something stronger than (WS) (or (F)) in mind when we think of the principle of compositionality (Szabó 2012a).

Framed as they are, I don’t think these objections are very powerful. Even if Szabó is right about our intuitions regarding the principle of compositionality, it’s not obvious

that these should be taken to bear so heavily on an account of what compositionality consists in. Again, ‘compositionality’ is a theoretical notion whose primary purpose has been to serve in explanations of acquirability. If accommodating these intuitions forces us into a notion of compositionality which could not play the usual role in these kinds of explanations, then I think we’d have ample reason to alter or ignore our intuitions.

The objections do nicely highlight the relative weakness of (WS) as a constraint on natural languages, however. In fact, what they illustrate is that the compositionality_{WS} of natural language does not explain its weak systematicity_{PSY}. Suppose that in natural languages, the meanings of complex expressions were derivable via syntactic rules which specify how the meanings of constituents are to be “put together”—as weak systematicity_{PSY} tries to capture. It should then be impossible for any two complex, natural language expressions to have identical constitution properties and distinct meanings. In at least one of these languages, you need to know more than that a given complex has a particular constitution property in order to arrive at its meaning. But the compositionality_{WS} of natural language does not rule this possibility out, and so does not suffice to explain its weak systematicity_{PSY}.²⁰

Put this way, I think Szabó’s objections have more force—they bring out the fact that if we understand compositionality in terms of (WS), then the compositionality of natural language explains less than is typically supposed. Indeed, since it is often assumed that compositionality suffices to explain the weak systematicity_{PSY} of natural language, it is no surprise that we have the intuitions about compositionality that Szabó points to in his objections. This might lead us to conclude, as Szabó does, that (WS) is “unreasonably weak”, and that we should look for ways to strengthen the principle (Szabó (2012a, 68).

In particular, he takes the objections to motivate the “strong supervenience” read-

²⁰Given the sorts of things languages are, they cannot, strictly speaking, change their linguistic properties. Thus, Szabó’s second objection collapses into the first.

ing of the principle (SS):

(SS) For all natural languages L , for any meaning property M and any complex expression e in L , if e has M in L , then there is a constitution property C such that e has C in L , and for any natural language L' if any complex expression e' in L' has C in L' then e' has M in L' (Szabó 2000, 499).

I will use ‘the principle of compositionality_{SS}’ to refer to (SS). In addition, we can say that a language L is compositionality_{SS} if and only if the meaning properties of L ’s complex expressions strongly supervene on the constitution properties of the complexes, in the manner spelled out in (SS).

Szabó’s idea is that (SS) captures what it is for the meanings of complexes to be *determined* by their constitution properties better than (WS) does, and so is a better reading of (C). The idea is that there is a single function across all natural languages from constitution properties of complexes to their meaning properties. Understanding the principle of compositionality in terms of (SS) also avoids the counterintuitive results that Szabó raises against the (WS) reading, and so Szabó takes this as further evidence that (SS) captures what authors mean by ‘the principle of compositionality’.

While I agree that the (SS) reading avoids the counterintuitive results that the (WS) reading faces, this does not show that it does any better at explaining the weak systematicity_{PSY} of natural language. Thus, it does not show that it does any better in addressing the underlying source of the intuitions either. I argue in the next section that in fact, not only would the compositionality_{SS} of natural language not suffice for its weak systematicity_{PSY}, it doesn’t play *any* role in explanations of the acquirability of natural language. If that’s right, then not only do Szabó’s objections to the (WS) reading fail to give us any reason to abandon it in favor of the (SS) reading, but we ought to reject the latter reading as a viable account of the principle of compositionality.

4 (SS) and Linguistic Explanations of Acquirability

Ultimately, my objection to Szabó's interpretation of the principle of compositionality is that the fact that a language is $\text{compositional}_{SS}$ can play *no* role in linguistic explanations of acquirability (§4.3), because it makes these explanations circular. Before making this objection, however, I want to argue (§4.1-3.2) that it does not suffice to explain the weak systematicity $_{PSY}$ of natural language. Thus, supposing the compositionality $_{SS}$ of natural language could play some role in explanations of acquirability, it—like the compositionality $_{WS}$ of natural language—would explain less than is ordinarily assumed. The discussion in §4.1-3.2 will help to lay the groundwork for the stronger objection to the (SS) reading in §4.3.

4.1 Compositionality $_{SS}$ and Weak Systematicity $_{PSY}$

Let's suppose that natural language is $\text{compositional}_{SS}$. Would the fact that natural language is $\text{compositional}_{SS}$ suffice to explain (PROD), then? No. The first reason is that any language which has no structure—which has no complex expressions—will be counted vacuously as $\text{compositional}_{SS}$. A language which has no complex expressions will also be vacuously weakly systematic $_{PSY}$, but in that case, weak systematicity $_{PSY}$ does not explain productivity, since there will be no novel complexes for the speaker to understand.

Adding the premise that natural language is structured does not suffice to explain (PROD) either, however. In this case, the fact that natural language is structured and $\text{compositional}_{SS}$ does not suffice for the weak systematicity $_{PSY}$ of natural language. To see this, note that the compositionality $_{SS}$ and structure requirements do not rule out a language like L^* as a natural language. L^* consists of the words 'John', 'Bill', 'Mary', and 'hit'—where these words have the same meanings as in English—and the relevant

syntax of English, so that all of the complexes that can be formed from the analogous English expressions are contained in L^* as well. The sentence ‘John hit Bill’ has the meaning it does in English—viz. John hit Bill, and also the constitution property it has in English, which I will write as $\langle \underline{\text{John}}, \underline{\text{hit}}, \underline{\text{Bill}}, [[_{NP}N][_{VP}VN]] \rangle$.²¹ ‘Mary hit Bill’ in L^* has the constitution property $\langle \underline{\text{Mary}}, \underline{\text{hit}}, \underline{\text{Bill}}, [[_{NP}N][_{VP}VN]] \rangle$, but it means Cats sleep in L^* .²²

There is a function from the constitution properties of complexes to their meaning properties, but it is not a sufficiently systematic function; the meaning of ‘Mary hit Bill’ is not related in the appropriate way to the meaning of its parts, given the way they are combined. Thus, it’s not clear how knowing what ‘Mary’, ‘hit’, and ‘Bill’ mean, and grasping the syntax of ‘Mary hit Bill’ would help one know what that sentence means in L^* ; knowing the syntactic rules of L^* apparently does not help one determine the meanings of complexes the way they do in English, for instance. (WS) certainly does not rule out the possibility that languages like L^* could be natural languages, but neither does (SS). Such languages are not weakly systematic_{PSY}, however, and are not the sort of languages that we could acquire as natural languages. Thus, the fact that a language is compositional_{SS}—together with the fact that the language is structured—does not suffice to explain its weak systematicity_{PSY}, and thus explains less about the

²¹This is obviously an oversimplification of English syntax which I make for the sake of readability; it will not have any bearing on the thrust of the discussion.

²²One might worry here that L^* is not even a possible language because it is metaphysically impossible for an expression with that constitution property to mean anything other than Mary hit Bill. Put another way, one might think that in L^* , ‘Mary hit Bill’ must be an idiom, and so doesn’t have the suggested constitution property. This worry only gets off the ground if we assume a version of what Szabó (2012a, 78) calls the “philosopher’s principle” of compositionality: “Complex expressions have the meanings they have *in virtue of* having such-and-such constitution property.” Linguists typically do not, and need not, make claims about why it is that a complex expression means what it does, and so this principle should not be assumed in the context of linguistic theorizing. Further, since the philosopher’s principle presumably makes a claim about meaning properties of complexes *generally*, it would follow that all languages—natural and artificial—are compositional in pretty much any sense of ‘compositional’, including the senses given by (WS) and (SS). Thus, the resulting notion of ‘language’ would be one that is entirely non-standard in empirical work on natural language. Szabó’s criticisms of (WS) also rely on there being languages like L^* . See Szabó (2000, 487-488) for his arguments to the effect that there are such languages.

acquirability of natural language than is often expected of compositionality.

Note, though, that if L^* were a natural language, then it would follow from (SS) that every natural language expression with the constitution property $\langle \underline{\text{Mary}}, \underline{\text{hit}}, \underline{\text{Bill}}, [[_{NP}N][_{VP}VN]] \rangle$ would mean Cats sleep. But this is absurd. We know that there are lots of natural languages—including, e.g., English—with expressions which have that constitution property, but which mean Mary hit Bill, and not Cats sleep. Perhaps, then, one could argue that while compositionality_{SS} does not explain weak systematicity_{PSY} by itself, it does once we add some very uncontroversial premises about the constitution properties and meaning properties of expressions in natural languages. It might even be thought that this gives an advantage to the (SS) reading of compositionality over the (WS) reading, since adding these premises to (WS) would not exclude a language like L^* from being a natural language. Perhaps (SS) is more explanatory than (WS), then, and so accords better with standard views about the explanatory role of the principle of compositionality.

There are problems with this response, though. First, it will require our explanations of acquirability to be infinitely long. L^* is just one among infinitely many languages that fail to be weakly systematic_{PSY}. Worse, there are an infinite number of ways to fail to be weakly systematic_{PSY}. For compositionality_{SS} to entail weak systematicity_{PSY}—or at least exclude all of the non-weakly systematic_{PSY} languages from the natural languages—we'd need to add, for each constitution property had by some natural language complex, a premise which pairs that constitution property with its corresponding meaning property. Call these premises—the true premises like the one that says that there is some natural language which has a complex expression with constitution property $\langle \underline{\text{Mary}}, \underline{\text{hit}}, \underline{\text{Bill}}, [[_{NP}N][_{VP}VN]] \rangle$ and meaning property Mary hit Bill—‘C-M premises’. Let's also say that for a language to have an expression with such-and-such constitution property and such-and-such meaning property is for it

to have such-and-such C-M property. The C-M premises are all of the true statements which say that there is some language which is a natural language, and which has such-and-such C-M property.

From the C-M premises and (SS) it will follow that every natural language which has an expression with a constitution property that is mentioned in the C-M premises must have the meaning property that is paired with that constitution property in the relevant C-M premise. If a language has a C-M property which “clashes” with the C-M premises—i.e. pairs a C mentioned in the C-M premises with the wrong M—then the language is rightly counted as artificial. The thought, then, is that since the function from the constitution properties to meaning properties of natural language complexes is systematic enough to ensure the weak systematicity_{PSY} of natural languages, we can say that a language which clashes with the C-M premises is unacquirable because it is not weakly systematic_{PSY}. Thus, we get weak systematicity_{PSY} from compositionality_{SS} and the C-M premises.

The problem with this approach is that natural languages have infinitely many meaningful complexes, and thus infinitely many constitution properties, so we’d need an infinite number of C-M premises to rule out all of the languages that fail to be weakly systematic_{PSY}. Now, we are assuming that all natural languages are weakly systematic_{PSY}, and that this calls out for linguistic explanation. Whatever linguistic properties are needed to explain the weak systematicity_{PSY} of natural language will be part of our account of what it is to be an acquirable language. On the C-M approach, we’d need infinitely many premises to do this. Assuming that we don’t want our explanations of acquirability to be infinitely long, this seems like a reason not to try to derive weak systematicity_{PSY} via compositionality and C-M premises, but instead to grant that the compositionality_{SS} of natural language is at best only a necessary condition on the weak systematicity_{PSY} of natural language.

4.2 The Circularity of the C-M Approach

I think we already have sufficient reason to reject the C-M approach to deriving weak systematicity_{PSY} from the principle of compositionality_{SS}. However, there is a deeper reason why the approach fails, *viz.* that it makes linguistic explanations of acquirability circular. This point is worth bringing out here because it will bear on the argument against the (SS) reading of the principle of compositionality in §4.3.

Recall that C-M premises say that there is some language L which is a natural language and which has such-and-such C-M property. On the C-M approach, these premises are part of an LEA for a language. The idea, then, is that the property *being a natural language* partly consists in the property *being such that there is some language which is a natural language and which has such-and-such C-M property*. Call this latter property ‘*being C-M*’. But generally, *being such that something is F and G* is explanatorily posterior to both *being F* and *being G* , so *being C-M* must be posterior to *being a natural language*. Thus, trying to explain what it is to be a natural language in terms of *being C-M* makes the explanation circular.

We thus have yet another reason to reject the C-M approach. More importantly, we can draw the more general conclusion that the fact that there is some natural language should not be assumed in an LEA,²³ because that commits us to the idea that *being a natural language* partly consists in—and is thus explanatorily posterior to—*being such that there is some natural language*. But it is not, and so our LEA’s would become circular. These points will play an important role in the argument against the (SS) understanding of compositionality in the next section.

²³Indeed, this fact is itself one of the target explananda of an LEA.

4.3 Why The Principle of Compositionality is Not (SS)

I've argued that compositionality_{SS} does not suffice to explain weak systematicity_{PSY} by itself, but I have not shown that the former does not at least partly explain the latter.²⁴ I now want to argue, though, that the claim that natural language is compositional_{SS} will play *no* role in our LEA's, and so contributes nothing towards an explanation of weak systematicity_{PSY}. The reason is that if we understand compositionality this way, then the principle is either explanatorily otiose, or else it makes LEA's circular. This gives us a compelling reason not to think of the principle of compositionality in terms of (SS), given the explanatory demands placed on the principle.

To see this, let's consider what languages are counted as compositional_{SS}, and consequently, what languages are ruled out as artificial by (SS). A language L (natural or artificial) will be counted as compositional_{SS} if and only if any of the following obtain:

- (a) L has no complex expressions.
- (b) No complex expression in L has a meaning property.
- (c) There are no natural languages.
- (d) Not (a), not (b), not (c), and for every complex e in L with a meaning property M , e has a constitution property C , and every natural language expression with C has M .

The claim, then, is that if we understand L 's being compositional as the claim that the meaning properties of its complexes strongly supervene on its constitution

²⁴I argue in §5 that compositionality_{WS} does partly explain weak systematicity_{PSY}.

properties, then what we are saying (implying) is that at least one of (a)-(d) is true. Now we have to ask if any of these does any work towards explaining L 's acquirability.

(a)-(c) represent the ways in which a language can be vacuously counted as $\text{compositional}_{SS}$. It is pretty easy to see that the fact that a language is $\text{compositional}_{SS}$ will be explanatorily otiose if it is vacuously $\text{compositional}_{SS}$ in one of these ways. No acquirable language will satisfy (a), because having complex expressions—having structure—is an independent constraint on acquirability; it is not part of the explanation of a language's acquirability that it has no structure. (b) obviously is not satisfied by any natural language either, since all such languages—indeed all languages, by definition—have some meaningful expressions. That a language satisfies (b), then, is not part of an explanation of its acquirability. If (c), then every language is trivially $\text{compositional}_{SS}$, provided we make the plausible assumption that any complex expression has some constitution property, and so attributing $\text{compositional}_{SS}$ to a language would not explain anything.²⁵

Presumably, then, when one attributes $\text{compositional}_{SS}$ to a language, one does so because it satisfies (d). Furthermore, it should do so non-vacuously. A language will vacuously satisfy (d) if it has some complex expression with C , but there's no natural language expression with C . No natural language will vacuously satisfy (d), then, and so the fact that a language vacuously satisfies (d) will not play any role in explanations of its acquirability.²⁶

²⁵This is a very counterintuitive result. If it turned out that there were no natural languages, we would still think that there could still be plenty of artificial languages which were not compositional. Note too that any number of languages would still fail to be $\text{compositional}_{WS}$ if there were no natural languages. Thus, while (SS) entails (WS), the claim that a language (unqualified) is $\text{compositional}_{SS}$ does not entail that it is $\text{compositional}_{WS}$. Thus, the sense of 'strong supervenience' at play in (SS) isn't quite like the sense used in 'moral properties strongly supervene on physical properties', since the fact that properties strongly supervene on physical properties at w entails that moral properties weakly supervene on physical properties at w .

²⁶That a language which satisfies (d) vacuously is counted as $\text{compositional}_{SS}$ is another counterintuitive result. There are artificial languages which only have complexes which have constitution properties not had by any natural languages. These languages are $\text{compositional}_{SS}$ even though they might have two complexes with the same constitution property, but which differ in meaning.

The problem is that a language satisfies (d) (vacuously or not) only if (c) isn't true. That is, it must be that there is some language which is a natural language. This assumption is uncontroversially true, but as I argued in §4.2, it's not the sort of thing we can assume in our explanations of acquirability, since that assumption is explanatorily posterior to the facts which constitute an LEA. To include such an assumption in an LEA is to try to explain *being a natural language* in terms of *being such that there is some natural language*. But this clearly gets things explanatorily backwards. Thus, if a language is compositional in virtue of satisfying (d), then when we attribute compositionality_{SS} to a language in a LEA, we make the explanation circular. It seems, then, that the fact that a language is compositional_{SS} contributes nothing to an LEA. Since whatever compositionality is, the fact that a language is compositional is supposed to contribute to an explanation of the acquirability of that language as a first language, I think the (SS) reading of the principle of compositionality ought to be rejected.

This criticism is based on what seems to be a general fact about constitutive explanation. Namely, that a thing's being *F* does not partly consist in its bearing some relation to the *F*s. In other words, claims with the following general form are false:²⁷

(CONST) *Being F* partly consists in *being R-related to all of the things that are F*.

A thing can be *R*-related to all of the things that are *F* either vacuously or non-vacuously. In the former case, the purported explanans—*being such that there are no F*s—is not explanatory. If there are no things that are *F*, then everything will be *R*-related to the *F*s, and so standing in this relation will not explain what it is to be *F*. As

Such languages seem paradigmatically non-compositional—the meanings of these complexes are *not* determined by their constitution properties.

²⁷Note that this is a claim about constitution, and does not commit one to the idea that all *definitions* with a form analogous to that of (CONST) ought to be rejected. See, e.g., Gupta (2012, ch. 3) for arguments to the effect that circular definitions can be useful in specifying the sense of a term. Further, if there are exceptions to the general rule concerning constitution, it seems doubtful that *being a natural language* is one of them.

in the compositionality case, if the explanans is to have any hope of being explanatory, it needs to be satisfied non-vacuously. In the non-vacuous case, the explanation of x 's being F will partly consist in its being such that a_1, \dots, a_n exhaust the F 's, and its being R -related to each of a_1, \dots, a_n . But *being such that a_1, \dots, a_n exhaust the F 's* is explanatorily posterior to *being F* , and so can't partly explain what it is to be F . The purported explanans doesn't actually explain anything, then, since it is explanatorily posterior to the explanandum, and thus makes the explanation circular. Thus, since constitution is an explanatory relation, facts with a form like (CONST) will be false.

The claim that part of what it is for a language to be a natural language is for the meaning properties of the complexes in that language to strongly supervene on their constitution properties is a claim with the form of (CONST). In this case, *being a natural language* is taken to partly consist in having a certain property, which a language bears in virtue of standing in a certain relation to the natural languages—the property *having a mapping from constitution properties to meaning properties which does not “clash” with the mapping from constitution to meaning properties of any natural language*, in the sense of ‘clash’ that was introduced in §4.1. It is this feature of the strong supervenience reading of the principle of compositionality which makes it unusable for the explanatory purposes that the principle is typically put to. Since whatever the principle of compositionality comes to, it should play some role in explanations of the acquirability of natural language, this gives us ample reason to look for an alternate reading of the principle.

Now, it is worth emphasizing that none of this is intended to show that (SS) is false or that natural languages are not compositional_{SS}. The claim is just that, given that *being a natural language* partly consists in *being compositional*, (SS) is not the principle of compositionality. Nothing I've said rules out the possibility that there is a function (call it ‘ f ’) which, for any complex expression in any natural language, maps

that expression's constitution property to its meaning property. This, of course, is the guiding idea behind compositionality_{SS}. It would even be consistent with what I've said that it is partly constitutive of a language's being a natural language that it only have complexes whose constitution properties that are mapped by f to their meaning properties. Let me briefly explain why.

Let's say that a language which has this property is "mapped by f ". Even supposing, though, that *being mapped by f* is partly constitutive of *being a natural language*, this would not provide any support for the idea that *being compositional_{SS}* is partly explanatory of *being a natural language*. First, note that while it follows from the claim that *being a natural language* consists partly in *being mapped by f* that all natural languages are compositional_{SS}, the fact that a language is mapped by f does not itself entail that it is compositional_{SS}. The fact that a language is mapped by f entails nothing about bearing a certain relation to the natural languages—we only get this entailment when we add the premise that all natural languages are mapped by f . Generally, though, it's not the case that if the fact that *being F* partly consists in *being G* entails that all F 's are H , then a thing's being H partly explains why it is F . For example, *being a prime number* partly consists in *being a number*. It follows that all prime numbers have a square. But something's having a square does not at all explain why it is a prime number. Thus, while it could be that a language's being mapped by f partly explains why it is a natural language, this would not offer any support to the idea that its being compositional_{SS} would do so as well.²⁸

²⁸While it would not be circular to identify *being compositional* with *being mapped by f* , I think we have independent reasons to reject the identification. One reason would just be that while compositionality is usually taken to entail facts about the relation between a language's constitution properties and its meaning properties, it is not taken to entail facts about what the particular meaning properties of complex expressions are, given their constitution properties. In short, *being mapped by f* is too strong to play the theoretical role that compositionality is supposed to play.

5 Compositionality as Weak Supervenience

The objections to the weak supervenience reading of the principle of compositionality were that it doesn't match our intuitions about what kinds of constraints compositionality places on natural languages, and that the compositionality_{WS} of natural language doesn't suffice to explain its weak systematicity_{PSY}, and so explains less about the acquirability of natural language than is often thought. I think both are correct. As I've argued, however, this does not provide any support for the strong supervenience reading of the principle. In attempting to accommodate our intuitions about compositionality, the strong supervenience reading precludes the principle from figuring in explanations of the acquirability of natural language, and thus violates a crucial constraint on an adequate account of compositionality.

The weak supervenience reading of the principle clearly does not suffer from the circularity problem that arose for the strong supervenience reading. For a language to be compositionality_{SS} was for it to have a certain relational property, which it bore in virtue of standing in a certain relation to the natural languages. This, in effect, was the source of the problems I raised for that reading in §4.3. A language is compositionality_{WS}, however, in virtue of bearing a certain intrinsic property, which it has so long as there is a function from the constitution properties of its complexes to the meaning properties of the complexes. Attributing compositionality_{WS} to a language, then, does not require assuming that there are any natural languages, or any other facts which are explanatorily posterior to an LEA. Further, the weak systematicity_{PSY} of natural language requires, and is partly explained by, its compositionality_{WS}—if the meanings of complexes were not functions of their constitution properties, then understanding the meanings of their constituents and how they are combined would not suffice for understanding the meanings of complexes. Thus, the compositionality_{WS} of natural

language *does* feature in explanations of the acquirability of natural language. The weak supervenience reading of the principle is thus preferable to the strong supervenience reading.

However, it's not as though (SS) and (WS) exhaust the possible readings of the principle of compositionality. Far from it. So I can't conclude that (WS) gives us *the* principle of compositionality. Indeed, there are alternative readings of the principle which avoid the pitfalls of (SS), accommodate our intuitions concerning what sorts of constraints compositionality places on natural languages, and suffice to explain the weak systematicity_{PSY} of natural language. (WS) essentially provides us with the minimal satisfactory reading of the principle; there are other readings which do more, but which also bring with them further substantial empirical commitments. I want to suggest one such principle now, and to draw out the kinds of commitments it would bring with it.

I've said that one constraint on an account of the principle of compositionality is that it needs to accommodate the fact that the principle plays a certain explanatory role in the acquirability of natural language. More specifically, the fact that natural language is compositional should at least *partly* explain the weak systematicity_{PSY} of natural language. Alternatively, one might think that compositionality consists in whatever linguistic property *suffices* to explain weak systematicity_{PSY}. In a number of places in this paper, I've suggested that compositionality is often thought of this way, and that this is reflected in our intuitions about what it is to be compositional.

A language is weakly systematic_{PSY} just in case it stands in a certain psychological relation to speakers. The fact that a given language is weakly systematic_{PSY} is something which calls out for linguistic explanation, however. We want to know which linguistic features of a language explain why it stands in this psychological relation to speakers. I will call the property which explains the weak systematicity_{PSY}

of a language ‘weak systematicity_{LING}’—the underlying linguistic explanation of weak systematicity_{PSY}:

A language L is weakly systematic_{LING} if and only if, for all syntactic rules $SynR$ of L , all semantic rules $SemR$ of L , and all complex expressions e of L , if e has constituents c_1, \dots, c_n , and has structure which can be derived from $SynR_1, \dots, SynR_n$, then e ’s meaning can be derived from the meanings of c_1, \dots, c_n by rules $SemR_1, \dots, SemR_n$, which correspond to $SynR_1, \dots, SynR_n$.

The claim that natural languages are weakly systematic_{LING} is a version of what Szabó (2012b) calls the ‘rule-to-rule’ principle. The thought here is that for each syntactic rule in a natural language which tells you how to put together simple expressions into complexes, there is a semantic rule which tells you how to put together the *meanings* of simple expressions to arrive at the meanings of complexes. The important difference, then, between compositionality_{WS} and weak systematicity_{LING} is that the latter requires that there be a *derivation* from the meanings of constituents to the meanings of complexes, which proceeds via some suitably constrained linguistic rules. (WS) merely requires that there be a *function* from some rules and the meanings of constituents to the meanings of complexes. The relevant rules in this case are semantic rather than syntactic just because, typically, syntactic rules are posited to explain how speakers grasp the syntactic structure of an expression, while distinct semantic rules are posited to explain how they arrive at meanings of those structures.

The claim that natural languages are weakly systematic_{LING}—and other versions of the rule-to-rule principle—are themselves fairly natural readings of the principle of compositionality, since it is not clear in (C) whether what matters is how the constituents of a complex are combined, or how the *meanings* of those constituents are combined.²⁹ (WS) opts for the former reading, while versions of the rule-to-rule prin-

²⁹This ambiguity in (C) is pointed out by Szabó (2012a, 70).

principle opt for the latter.³⁰ Notice though, that the claim that natural languages are weakly systematic_{LING} brings with it substantial commitments about the relation between syntax and semantics. For instance, it commits us to the idea that the derivation of the meaning of a complex cannot involve more semantic rules than the number of syntactic rules needed to derive its structure. This isn't obviously false, but it's not obviously true either.³¹

In fact, if weak systematicity_{LING} is going to explain weak systematicity_{PSY}, we need to add some additional constraints on what semantic rules can be like. If the only constraint on being a semantic rule is that it specify a function from the meanings of constituents to the meanings of complexes, then weak systematicity_{LING} will not get us much more than compositionality_{WS}, and will not rule out, for instance, the possibility that L^* is weakly systematic_{LING}. But L^* is an example of a language that is paradigmatically not weakly systematic_{PSY}, and so it should not be weakly systematic_{LING} either, if the latter explains the former. Typically, however, semanticists do place additional constraints on what semantic rules can be like. For instance, we could require that semantic rules are sensitive only to the syntactic arrangement of the complex expression, and to perhaps some coarse-grained semantic properties of the constituents—e.g. their semantic type. Heim and Kratzer's (1998) "Functional Application" would be an instance of a semantic rule which meets these constraints:

(FA) If α is a branching node with β and γ as daughters, and $\llbracket\beta\rrbracket \in \llbracket\gamma\rrbracket$, then $\llbracket\alpha\rrbracket = \llbracket\gamma\rrbracket(\llbracket\beta\rrbracket)$

³⁰Pagin and Westerståhl (2010a), for instance, call a version of the rule-to-rule principle 'the function version' of 'basic compositionality'.

³¹Whether this particular formulation of the rule-to-rule principle is at all plausible will depend on what syntax framework one adopts. For instance, in Chomsky's minimalist approach (Chomsky, 1995), there is a single syntactic operation 'Merge' which is responsible for combining lexical items into complexes. Assuming the weak systematicity_{LING} of natural language, this would conflict with a standard thought that there are at least a handful of semantic rules for deriving meanings of complexes (e.g., Heim and Kratzer (1998)).

(FA) just says that if a complex expression α has β and γ as constituents, and the meaning of β is in the domain of γ , then the meaning of α is what you get when you take the meaning of β as an argument for the meaning of γ . Since the rule only applies if $\llbracket\beta\rrbracket \in \llbracket\gamma\rrbracket$, it will need to be sensitive to the semantic type of the constituents in addition to the syntactic makeup of the complex.³² With rules like this in place, there is no way that Cats sleep could be derivable from Mary, hit, and Bill, and English syntax, given that John hit Bill is derivable from John, hit, and Bill and English syntax—i.e. L^* is not weakly systematic_{LING}. If we suitably constrain what counts as a semantic rule, then, the claim that natural languages are weakly systematic_{LING} is substantially stronger than the claim that natural language is compositional_{WS}. We would lose this result if, for instance, we permitted semantic rules to be sensitive to the particular meanings of the constituents.

Adding the above constraints, however, would not be sufficient for the claim that if complex expressions in two distinct languages share a constitution property but differ in meaning, then one of the languages must fail to be weakly systematic_{LING}. But I said in section §3 that in that kind of case, one of the languages must not be weakly systematic_{PSY}, so one should fail to be weakly systematic_{LING} as well. The problem here is that weak systematicity_{LING} does not guarantee that the mapping from syntactic rules to semantic rules is the same in both languages, so we do not yet have any reason for thinking that they can't both be weakly systematic_{LING}. We need to also assume, then, that the required relation between syntax and semantics obtains across all natural languages, so that every natural language maps its syntactic rules to the same semantic rules.

Finally, something needs to be said about linguistic understanding and the psychol-

³²The rule also assumes that at least some meanings are functions which can take other meanings as arguments. This assumption is not crucial to the discussion though, and (while fairly common) is simply an artifact of the example.

ogy of speakers if we are to have an adequate explanation of weak systematicity_{PSY} from weak systematicity_{LING}. I want to briefly sketch how I think this explanation can be filled out, roughly in line with standard views in generative linguistics.

In the framework of generative linguistics, one’s understanding or “having” a particular language depends entirely on their internal, psychological state. Which language is understood is determined by the fine-grained details of the state one’s language faculty.³³ To understand a certain expression, then, is just for one’s language faculty to be such that it associates the expression (or more precisely, a speaker’s phonemic representation of the expression) with syntactic and semantic features. For one to understand a particular rule is just for their language faculty to be ‘following’ that rule, in whatever sense of ‘follow’ is most apt here.³⁴ To be able to understand a complex expression is just to have the resources to derive its meaning: one must understand the simple constituents, the syntactic rules to generate the syntactic structure of the complex, and the semantic rules for deriving the meanings of the complexes. If we also assume that competent speakers of natural language will understand the semantic rules which correspond to the syntactic rules, then the weak systematicity_{PSY} of a language will follow from, and be explained by, its weak systematicity_{LING}. Thus, even if the meanings of complexes are not strictly derivable from the meanings of the constituents by the syntactic rules used to form the complex—and so the syntactic rules don’t really tell you how to arrive at the meaning of the complex—as a matter of psychological necessity, understanding the constituents and syntactic rules will suffice for being able to understand the complex. We thereby secure the explanatory link from the weak systematicity_{LING} of a language to its weak systematicity_{PSY}. The latter,

³³This conception of linguistics and linguistic understanding was, of course, made prominent by the work of Noam Chomsky. See, e.g., Chomsky (1986) and Ludlow (2011) for discussion of the theoretical framework of generative linguistics.

³⁴See Chomsky (1986) and Ludlow (2011) for relevant discussion, and for defense against standard Kripkean/Wittgensteinian worries about the use of ‘rule-following’ in linguistics.

again, explains the productivity of a language, which is essential to natural language competence.

It should be clear, however, that the explanation from the weak systematicity_{LING} of natural language to its weak systematicity_{PSY} brings with it substantial empirical commitments concerning the nature of semantic rules, the relation between syntax and semantics, and the nature of linguistic understanding. This is not, by any stretch, a decisive cost for the weak systematicity_{LING} understanding of compositionality—the empirical assumptions it brings with it are substantial, but are neither implausible nor uncommon in linguistic theorizing. If one is uncomfortable with any of these assumptions, however, or at least thinks that they should not be entailed by compositionality, then they have reason to avoid this stronger conception of compositionality.

I have no substantive objections to understanding the principle of compositionality as the claim that natural languages are weakly systematic_{LING}. Indeed, that the compositionality of natural language is often taken to account for its weak systematicity_{PSY} suggests that many authors already have some comparable principle in mind. The choice between this principle and the more minimalistic (WS) seems to me to be a matter of preference. The latter seems preferable to me both because it is a fairly simple and common formulation of the principle, and because it is comparatively empirically modest. It carries fewer empirical commitments with what is already a substantial, and sometimes contested, empirical commitment—*viz.*, that natural language is compositional_{WS}. In addition, adopting the minimal conception forces one to be explicit about, and distinguish between distinct properties which go into the explanation of the productivity of natural language. Since everyone agrees that compositionality is at least compositionality_{WS}, the discussion can then move past the issue of how we ought to understand ‘compositional’, to more interesting questions concerning which

properties of language explain various features of our linguistic competence, or which properties of thought explain certain features of our conceptual competence.

Alternatively, one might think Szabó's methodology of consulting our intuitions about compositionality is the way forward. If it is, then I concede that (WS) probably does provide an unreasonably weak reading of the principle of compositionality, and that one should adopt a stronger principle which accords better with our intuitions. The thrust of this paper has been that some ways of strengthening (WS) are out of bounds, however, even if the resulting principle is plausibly true. (SS) is one such principle, since it has no place in explanations of the acquirability of natural language.³⁵

Szabó says that the principle of compositionality should be thought of as being “on a par with other bold hypotheses, such as that all syntactic operations are binary, that lexical categories are universal, that logical form and syntactic structure are intimately related, and so on,” (2012a, 65). But if we understand compositionality in the way he suggests, then while it may be on a par with these other hypotheses in terms of its boldness, it will not be on a par in terms of its place in linguistic theorizing. Szabó's principle may very well be true, but it is not one which should constrain semantic theories the way that semanticists typically take compositionality to constrain their theories. This, I've argued, gives us ample reason to reject the strong supervenience reading of the principle, and to adopt the weak supervenience reading in its place.

³⁵Note too that the claim that natural languages are weakly systematic_{LING}—together with the assumptions concerning semantic rules and the relation between natural language syntax and semantics—does entail (SS), but the claim that a language is weakly systematic_{LING} neither entails nor is entailed by the claim that it is compositional_{SS}.

6 Addendum: Compositionality and Non-Truth-Conditional Approaches to Semantics

Assuming that the principle of compositionality is to be understood in terms of (WS), then, how does this bear on compositionality-based arguments in favor of truth-conditional semantics? The short answer is that it shows that they are not sound, because it is simply not true that only truth-conditional semantics is compositional.³⁶

It will be worth exploring one of these arguments here. Take, for instance, an objection Fodor and Lepore (1993, 23-24) make against inferentialism. They argue as follows. Natural language is compositional, and so the meanings of complex expressions in natural languages are a function of the meanings of their constituents and their syntactic arrangement—i.e. natural language is compositional_{WS}. If this weren't the case, then we would have no explanation for the productivity of natural language. For instance, every English speaker who understands 'brown' and 'cow'—and how those are combined in 'brown cow'—understands 'brown cow'. But, they say, consider a speaker who thinks that brown cows are dangerous, and so will infer from 'x is a brown cow' that 'x is dangerous', but who does not infer from 'x is brown' or 'x is a cow' to 'x is dangerous'. According to inferentialism, the inference from 'x is a brown cow' to 'x is dangerous' is going to be partly constitutive of the meaning of 'brown cow' for the speaker in question. How, then, does the speaker arrive at the meaning of 'brown cow' on the basis of the meanings of 'brown' and 'cow' and the syntax of 'brown cow', given that none of these would seem to contain any hint of the dangers that brown cows present?

The crux of the argument seems to be as follows:

³⁶I should note here that the arguments are no better is Szabó is right that the principle of compositionality should be understood as (SS). Indeed, I'm not sure that there are any versions of the principle of compositionality which tell in favor of truth-conditional semantics and against all other theories of meaning.

- (i) (SUFF): To be able to grasp the meaning of a complex expression in natural language, it is sufficient to grasp the meanings of its constituents and how they are put together in the complex.
- (ii) Grasping the inferential role of ‘brown’ and ‘cow’ (for the speaker in question) and how they are combined in ‘brown cow’ is not sufficient for being able to grasp ‘brown cow’.
- (iii) Therefore, meaning is not inferential role.

Now, whether or not this argument works, it’s not immediately obvious how compositionality_{WS} fits into it. It seems as though Fodor and Lepore think—despite the fact that they explicitly identify compositionality with compositionality_{WS}—that it follows from (ii) that inferentialist semantics is not compositional. But what they’ve really argued for is just that natural language is weakly systematic_{PSY}, and that inferential semantics isn’t weakly systematic_{PSY}. But weak systematicity_{PSY} places a much stronger constraint on language than compositionality_{WS}, and so to show that a semantics or language is not weakly systematic _{PSY} is not to show that it isn’t compositional_{WS}.³⁷

Indeed, one quick glance at the example and at (WS) is enough to show that the argument doesn’t impugn the compositionality_{WS} of inferential role in English. (WS), again, just says that no natural language can have two complex expressions which share a constitution property, but which differ in meaning. The ‘brown cow’ example is one in which it might not be clear how the speaker arrives at the (inferential) meaning of the complex from the meanings of the constituents and syntax, but unless we have some distinct complex which has constituents with the same inferential roles as ‘brown’ and ‘cow’, respectively, the same syntax, yet has a distinct inferential role, we have

³⁷Strictly, compositionality is a property of languages. For the sake of convenience, however, we can say that a theory is compositional just in case it describes natural language as being compositional.

no counterexample to (WS). Thus, the argument does not provide any support for the claim that inferentialism isn't compositional_{WS}.³⁸

Now, perhaps one might take all of this to suggest that compositionality_{WS} is simply too weak, and so (WS) is too weak to serve as the principle of compositionality. As I mentioned in §5, I am amenable to the idea that the principle of compositionality just states that natural languages are weakly systematic_{LING}—i.e. the idea that some form of the rule-to-rule version of compositionality is correct. However, to expect compositionality to suffice to explain the weak systematicity_{PSY} of language, is, to my mind, to put far too great an explanatory demand on the principle, and is liable to invite confusion.³⁹

Further, it is by no means clear that inferentialist theories even fail to be weakly systematic_{LING}. In particular (and as Fodor & Lepore note), it matters whether the inferentialist thinks that all inferences are meaning-constituting. One might think, for instance, that one's inferring from 'x is a brown cow' to 'x is dangerous' is not constitutive of their competence with 'brown cow'.⁴⁰

The remarks here ought to generalize to other non-truth-conditional approaches to semantics. Indeed, the discussion vindicates some of Horwich's remarks concerning compositionality—for instance, that compositionality fails to place any constraint on

³⁸Similar remarks apply to Fodor and Lepore's (1996) argument against prototype theory: associating a prototypical pet with 'pet' and a prototypical fish with 'fish' is not enough for you to be able to associate a prototypical pet fish with 'pet fish'. Therefore (they argue), prototypes aren't compositional, and prototype theory is wrong.

³⁹Also, as I mentioned in §5, if one expects the fact that natural language is weakly systematic_{LING} to suffice to explain the fact that natural language is weakly systematic_{PSY}, then one will have to adopt additional, highly theoretical assumptions for the explanation to work.

⁴⁰Brandom's (2007) reply, in part, is that we should not overestimate the extent to which language is compositional. In other words, depending on how we understand what it takes to understand an expression, we might simply reject that natural language is weakly systematic_{PSY}. Understanding some complex expressions might require much more than knowing the meanings of its constituents and how they are put together. In the context of generative linguistics, I don't think this is a promising line, since the way Brandom understands what it is to understand an expression seems to go way beyond the explanatory scope of generative linguistics.

the meanings of simple expressions.⁴¹ On his view, the meaning of a complex is simply *constituted* by its having the constitution property it has (i.e. its having such-and-so constituents with such-and-such meanings, syntactically arranged in a particular way). Neither compositionality_{WS} nor weak systematicity_{LING} rule this sort of view out. Indeed, in order for either principle to have any hope of ruling out any theory about the meanings of simple expressions, one must also make what Horwich calls the “Uniformity Assumption”—roughly, the assumption that *being the meaning of a complex expression* and *being the meaning of a simple expression* must be grounded in the same sorts of properties.⁴² Nothing about compositionality_{WS} nor weak systematicity_{LING} forces us to adopt this assumption, however.

Similarly, Horwich’s view might seem problematic if one thinks that compositionality requires that the meanings of complex expressions literally be “built up” from the meanings of their constituents—i.e. that the meanings of constituents be *parts* of the meanings of complexes—or that compositionality requires that the meanings of complexes be *explained* by facts about the meanings of their constituents and their syntactic arrangement.⁴³ After all, on Horwich’s view, the meanings of complex expressions are not complex entities, and so they do not have parts. But, again, neither compositionality_{WS} nor weak systematicity_{LING} entail these stronger views. And more importantly, neither of these stronger views are forced on us in the context of semantics in generative linguistics.⁴⁴

I conclude that the requirement that semantic theories be compositional does not

⁴¹See, e.g., Horwich (1998, 2005 Ch. 8, 2003). Note that this does not make the principle of compositionality trivial. On either understanding of compositionality that I have been willing to accept, any number of possible languages will fail to be compositional.

⁴²In other words, the Uniformity Assumption entails that (for example) if ‘dogs’ and ‘bark’ have their meanings in virtue of some relation they bear to their uses, then ‘dogs bark’ must have its meaning in virtue of a similar relation to its uses.

⁴³See Fodor & Lepore (2001a) and Heck (2013), respectively, for endorsement of these stronger forms of compositionality.

⁴⁴And indeed, it’s not even clear how well truth-conditional semantics meets these stronger constraints.

force us to think that semantics in generative linguistics must be truth-conditional. A common a priori argument for thinking that truth and linguistic semantics are tightly intertwined, then, fails. At this point, however, the truth-conditional semanticist has an excellent reply to make—namely, that truth and linguistic semantics are tightly intertwined because the most successful research program in linguistic semantics, by far, is truth-conditional semantics. I turn to this argument in the next chapter.

Chapter 5

How Important Are Truth-Conditions For Truth-Conditional Semantics?

Chapter Overview: In Chapter 4, I argued that the requirement that semantic theories be compositional does not require that semantics be truth-conditional. Thus, the a priori argument from a compositionality requirement to truth-conditional semantics fails. However, there is a very good argument available to the truth-conditional semanticist. In particular, they can just argue from the (unmatched) success of truth-conditional semantics in generative linguistics to the claim that semantics is best understood the truth-conditional way.

Whether or not the argument is vindicated or not depends, of course, on how the empirical work in linguistic semantics develops. While I think that semantics is still in an exploratory phase, and thus that we should be careful about accepting such inferences, I think that the argument is fairly strong, and disputing it is not my primary aim here. My focus, rather, is on the role of truth in truth-conditional semantics. In

other words, supposing the correct approach to linguistic semantics is truth-conditional, what is the relation of a notion of truth to that research program?

The obvious (and standard) answer to this question is that a notion of truth-conditions plays a central and explanatory role in truth-conditional semantics. Truth, therefore, is tied up in explanations in linguistic semantics.

In this chapter, I argue against this answer. I argue that, in fact, the best explanations in truth-conditional semantics are non-truth-conditional. That is, facts about the truth-conditions of expressions do not play an explanatory role in the theories. Now, as I make clear, we cannot conclude that, therefore, notions of truth and truth-conditions are irrelevant to linguistic semantics. Specifying truth-conditions might still be necessary for reasons that go beyond the need to explain the central explananda of semantics. Nevertheless, I think there is plenty of room to wonder about the place of truth and truth-conditions in semantics. If notions of truth and truth-conditions do not play a part in the explanatory success of truth-conditional semantics, then the explanatory success of truth-conditional semantics does not tell so decisively against alternative semantic approaches.

1 Introduction

Here is a natural way of thinking about the explanatory role of truth-conditions in linguistic semantics. According to truth-conditional semantics—a very successful and ongoing research program in linguistics—the semantic value of an expression is its truth-conditions or contribution to truth-conditions.¹ The fact that an expression has the semantic value it has, generally, explains semantic phenomena relevant to the expression—e.g. why it is involved in some entailment pattern or why it can't be

¹Often, truth-conditional semantics is taken to identify the *meaning* of an expression with its truth-conditional content. I will opt for the less philosophically loaded “semantic value” throughout.

combined with another expression without anomaly. Therefore, facts about the truth-conditions or truth-conditional contributions of expressions explain, or at least play an important role in explaining, semantic phenomena.

I think that reflection on our best work in truth-conditional semantics suggests that this natural way of thinking is mistaken. While we can and should grant that truth-conditional semantics provides lots of good explanations of semantic phenomena, I do not think that we should conclude that facts about the truth-conditional contents of expressions—their extensions or intensions—play any role in the explanations of those phenomena. Indeed, this chapter defends the position that those facts do not play any role in explanations of semantic phenomena (properly understood). Rather, it will be argued, the good explanations of semantic phenomena in truth-conditional semantics appeal to non-truth-conditional structure that is associated with expressions via their lexical entries. The particular explanations will vary from case to case, but I will argue that we have good reason to think that, generally, explanations of semantic phenomena will appeal to facts other than facts about which truth-conditional contents expressions have.

Let me clarify my thesis by making explicit some background assumptions that will be in place throughout the paper. First, we will be concerned with truth-conditional semantics as a scientific research program in linguistics—more specifically, linguistics as it is conceived of in the generative linguistics framework. In that framework, linguistics is a branch of cognitive psychology, aiming to describe the workings of the human language faculty (typically at a very high level of abstraction).² To be clear, I intend to use ‘generative linguistics’ here quite broadly, so as to be neutral on debates concerning, for instance, the extent to which linguistic competence is innate, or the correct way to conceive of the relationship between syntax and semantics. For our purposes,

²See, e.g., Chomsky (1986, 2000) and Ludlow (2011) for detailed characterizations of the generative linguistics framework. See also Yalcin (2014) for discussion on how semantics fits into this framework.

what matters is just that generative linguistics is a broadly psychological research program which studies a particular faculty of the human mind. Truth-conditional semantics, then, is to be thought of as an empirical research program which claims that the semantic component of the language faculty—call this the “semantic faculty” for convenience—assigns truth-conditional content to lexical items and compositionally derives the truth-conditions of complexes.³

It is important to note, however, that pairing expressions with their truth-conditional contents is not the final theoretical aim of truth-conditional semantics in this context. Rather, the thought is that we can explain semantic phenomena—*anomaly* and *entailments*, for example—by appealing to their truth-conditional contents (and, e.g., an entailment rule which is sensitive to those contents). We take these linguistic phenomena to be among our best sorts of evidence for what the semantic faculty is like, so if a theory can explain why some sentence entails another, or why some complex expression is semantically anomalous, then that is evidence that the theory gives a good account of the nature of the semantic faculty.⁴

My claim, then, is just that while truth-conditional semantics often does a fine job explaining these phenomena, it is not facts about the truth-conditional contents of expressions which do any of the explaining. Rather, as I will try to show, it is the non-truth-conditional structure that is paired with expressions via their lexical

³See, e.g., Heim & Kratzer (1998), Larson & Segal (1995), and Ludlow (2011) for examples of theorists who explicitly locate truth-conditional semantics roughly within the generative linguistics framework. Of course, truth-conditional semantics is not always construed as fitting within generative linguistics. See, e.g., Partee (1979), Lewis (1975), and Soames (1992) for theorists who think of truth-conditional semantics as a project which is distinct from psychological approaches to linguistics. While I will not defend the claim here, it seems to me that it is best conceived of as a branch of generative linguistics if one’s aim is to stake the importance of truth and truth-conditions to an ongoing and successful research program in linguistics.

⁴I will not distinguish between Davidsonian, T-theoretic approaches and Montagovian, model-theoretic approaches to truth-conditional semantics in this chapter (see Glanzberg (2009, 2014) for justification of this simplification). Since the latter approach is the more dominant approach, however, I will adopt the model-theoretic notation and use ‘intension’ and ‘extension’ to refer to the truth-conditions of a sentence or the truth-conditional contribution of subsentential expressions.

entries that does the explanatory work. Since truth-conditional content does not play an important role in our best explanations of semantic phenomena, we have reason to think that truth-conditional content, generally, is extra-semantic. To be clear, when I argue that truth-conditional content is extra-semantic, the claim is not that it doesn't have anything to do with meaning, nor that it fails to constitute (or partly constitute) meaning, but rather that some other faculty of mind, outside of the language faculty proper, associates expressions with their truth-conditional content.⁵

The chapter contains two main arguments. Both arguments will require a great deal of clarification and support which the rest of the paper will aim to provide. Nevertheless, I think it will be useful to lay them out explicitly here.

The first argument, which I will call 'The Insufficiency Argument', is as follows:

Insufficiency Argument

- (i) If the fact that an expression has such-and-so truth-conditional content sufficed to explain facts about semantic entailment or semantic anomaly relevant to that expression, then disquotational lexical entries would support good explanations of these phenomena.
- (ii) However, disquotational lexical entries do not support good explanations of semantic entailment or anomaly.
- (iii) Therefore, the fact that an expression has the truth-conditional content it has does not suffice to explain relevant facts about semantic entailment or anomaly.

If correct, the Insufficiency Argument would suggest that truth-conditional semantics, in addition to assigning expressions truth-conditional contents, attributes other sorts of properties to expressions which figure in explanations of semantic entailment and

⁵The claim is thus not merely a verbal stipulation, but a substantive psychological hypothesis.

anomaly. The Insufficiency Argument, by itself, does not show that facts about truth-conditional content do not, nevertheless, play an important role in explanations of semantic phenomena. Such facts might still be a part of these explanations, even if they don't suffice to explain any phenomena by themselves.

I argue via the Superfluity Argument, however, that facts about the truth-conditional contents of expressions do not even partly explain facts about semantic entailment and anomaly:

Superfluity Argument

- (i) If facts about semantic entailment and anomaly are partly explained by facts about the truth-conditional contents of expressions, then the good explanations of entailment and anomaly will depend on facts about the truth-conditional contents of expressions.
- (ii) However, the good explanations of entailment and anomaly do not depend on facts about the truth-conditional contents of expressions.
- (iii) Therefore, facts about the truth-conditional contents of expressions do not even partly explain facts about semantic entailment and anomaly.

If this argument is right, then it would show that facts about truth-conditional contents do not play any role in explanations of a central class of semantic phenomena. This does leave open the possibility that such facts do play a role in explanations of other semantic phenomena—truth value judgments, for instance. I will argue, however, drawing on recent work from Michael Glanzberg (2014), that we have reason to think that these phenomena are not properly semantic, but are extra-semantic. Once again, this is not to deny that the phenomena are pre-theoretically meaning-relevant, but just to claim that they are best explained by appeal to faculties beyond the semantic faculty. My suggestion, then, is that the kinds of explanations that are distinctive of semantics are

non-truth-conditional explanations. *Having such-and-so truth-conditional content* is best understood as an extra-semantic property of an expression.

Throughout, I will use recent truth-conditional work on gradable adjectives and degree modifiers as a case study. I will argue that while truth-conditional semantics provides good explanations of the relevant semantic entailment and anomaly data, the fact that the expressions have the truth-conditional contents they have does not play any part in the explanations. While the paper will focus on details of explanations involving gradable adjectives and degree modifiers, I will give reasons for thinking that the points generalize to analyses of other sorts of expressions.

The structure of the paper is as follows. In §2, I will give a brief sketch of the semantics of gradable adjectives and degree modifiers defended by Kennedy & McNally (2005). In §3, I will present and defend a qualified version of the Insufficiency Argument which will be limited to semantic explanations involving gradable adjectives and degree modifiers. A similarly qualified version of the Superfluity Argument will be presented and defended in §4. In §5, I draw on recent work from Glanzberg (2014) to give reasons for thinking that there is nothing special about gradable adjectives and degree modifiers in the relevant regard, and that we should accept the unqualified versions of the Insufficiency and Superfluity Arguments. I will also argue that acceptance of these arguments does not force us to deny the successes of truth-conditional semantics. What is required, rather, is a reinterpretation of those successes.

2 The Semantics of Gradable Adjectives and Degree Modifiers

2.1 The Basics of Truth-Conditional Semantics

In this section, I will give an introductory sketch of the truth-conditional framework, and the semantics of gradable adjectives and degree modifiers given by Kennedy & McNally (2005).⁶

As I said in §1, the aim of truth-conditional semantics—as we are concerned with it—is to describe the workings of the semantic component of the human language faculty. It hypothesizes that this semantic faculty derives the truth-conditions of sentences on the basis of the truth-conditional contributions of their constituents, their syntactic arrangement, and a handful of rules of composition. A major part of truth-conditional theorizing, then, is to give an account of the truth-conditional contributions of simple expressions. These are represented in lexical entries, a handful of which are given in (1a)-(1d).⁷

- (5.1) a. $\llbracket \textit{Mary} \rrbracket = \textit{Mary}$
b. $\llbracket \textit{hikes} \rrbracket = \lambda x. x \textit{ hikes}$
c. $\llbracket \textit{dog} \rrbracket = \lambda x. x \textit{ is a dog}$
d. $\llbracket \textit{loves} \rrbracket = \lambda x. \lambda y. y \textit{ loves } x$

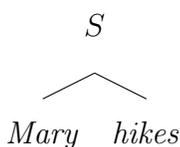
A lexical entry, in abstract, is meant to specify the linguistic properties of simple expressions. The property of an expression that is relevant to semantics, of course, is its semantic value. Truth-conditional semantics understands semantic values truth-conditionally, and its lexical entries pair expressions with their extensions (or intensions)—i.e. their truth-conditional contents. We can think of the job of lexical entries, for now,

⁶See also Kennedy & McNally (1999) and Kennedy (2007).

⁷I will be drawing mostly on the truth-conditional framework presented in Heim & Kratzer (1998).

as specifying the extensions or intensions of expressions.⁸ For instance, the lexical entry in (1a) should be read, “The semantic value of *Mary* is Mary.” (1b) should be read, “The semantic value of *hikes* is the smallest function which maps individuals to the value True if they hike, and False otherwise.” Equivalently, we can think of the extension of (1b) as the set of things that hike.

Consider a sentence like *Mary hikes*, then, the (oversimplified) syntactic representation of which is presented here:



According to the principle of compositionality (a suitably strong version of it), we should be able to compute its semantic value on the basis of its syntactic structure, and on the semantic values of *Mary* and *hikes*. The rule Functional Application tells us how to do this.

Functional Application (FA): If α is a branching node with β and γ as daughters, and $\llbracket \beta \rrbracket \in \llbracket \gamma \rrbracket$, then $\llbracket \alpha \rrbracket = \llbracket \gamma \rrbracket(\llbracket \beta \rrbracket)$.

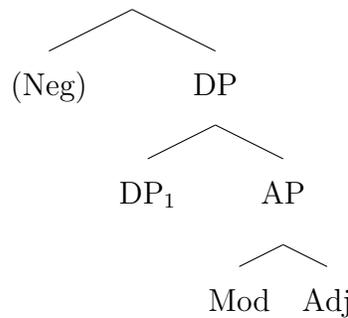
In the syntactic representation of *Mary hikes*, S is a branching node with *Mary* and *hikes* as daughters of S . Functional Application, in effect, tells you that if the semantic value of one daughter can take the semantic value of the other as an argument, let it do that, and the result will be the semantic value of the branching node. In this case, *hikes* denotes a function which takes individuals as arguments. Mary is one such individual, so we can use (FA) to derive the semantic value of *Mary hikes*. The end result will be what we want, a specification of the conditions under which the sentence is true: $\llbracket \text{Mary hikes} \rrbracket = \text{True}$ iff Mary hikes.

⁸I will largely ignore the distinction between extension and intension throughout, as it will not be relevant to my arguments. I will tend to make use of extensional lexical entries because of their comparative simplicity.

As the complexity of the sentences increase, so do the derivations. In some cases Functional Application, by itself, won't allow us to complete the derivation of the truth-conditions of a sentence. However, all of the simple constructions we will be concerned with here have this function-argument structure, and so we need not go beyond these basics.

2.2 Gradable Adjectives and Degree Modifiers

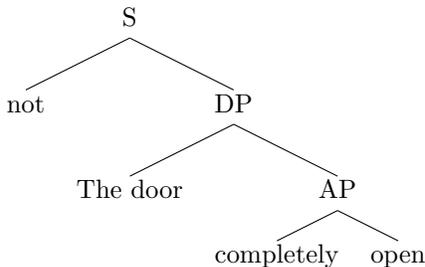
Let us now turn to the semantics of gradable adjectives and degree modifiers given in Kennedy & McNally (2005).⁹ Gradable adjectives are expressions like *tall*, *open*, *wet*, etc. They can be modified by degree modifiers like *very*, *completely*, or *half*. The simplified syntax that we will work with is as follows:¹⁰



Degree modifiers combine with the gradable adjectives they modify, and then the result will combine with a noun phrase or determiner phrase to result in a sentence. The extension of a gradable adjective like *tall* is, roughly, the set of things that have a certain, contextually determined amount of tallness (i.e. height). The extension of

⁹See also Kennedy & McNally (1999) and Kennedy (2007).

¹⁰So, for instance, *The door is not completely open* would be represented:



a degree modifier like *very* takes an extension of a gradable adjective like *tall* as an argument and outputs the set of things that greatly exceed the contextually determined standard of tallness, for instance. Kennedy & McNally assign to these expressions lexical entries like those in (1):

- (5.2) a. $\llbracket tall \rrbracket = \lambda x. \exists d[C(d) \wedge \mathbf{tall}(x) \geq d]$
 b. $\llbracket closed \rrbracket = \lambda x. \exists d[d = \text{MAX}(S_{closed}) \wedge \mathbf{closed}(x) = d]$
 c. $\llbracket half \rrbracket = \lambda G.\lambda x.\exists d[\text{DIFF}(\text{MAX}(S_G),d) = \text{DIFF}(d,\text{MIN}(S_G)) \wedge G(d) = x]$
 d. $\llbracket completely \rrbracket = \lambda G.\lambda x. \exists d[\text{MAX}(S_G) = d \wedge G(d) = x]$ ¹¹

The lexical entries in (1) specify the extensions I mentioned above, but in a somewhat complicated way. (1a), for example, says that *tall* denotes a function which maps individuals to the value True just in case there is some degree *d* such that *d* is the contextually determined standard, and the function **tall** maps the individual to a degree which is greater than or equal to *d*. **tall**, basically, maps things to their height values, and so the extension of *tall* is just the set of things that meet that standard of height. Importantly, though, these height values are values on a scale, which we will call ‘*S_{tall}*’. So we can represent **tall** and *S_{tall}* as follows:

$$\mathbf{tall}: f : H \subseteq U \longrightarrow \langle D_{(0,1)}, \leq, height \rangle$$

$$S_{tall} = \langle D_{(0,1)}, \leq, height \rangle$$

The domain $H \subseteq U$ consists of the individuals which have some height. ‘ $D_{(0,1)}$ ’ indicates a set of degrees, including all values between 0 and 1, excluding 0 and 1 themselves. **tall**, strictly, maps individuals to these values, which are ordered by the relation *less than or equal to* along a dimension which measures height. What is

¹¹I have altered the lexical entries for degree modifiers slightly for ease of exposition. Nothing will turn on these alterations.

important about **tall**, for our purposes, is the structure of the scale it is associated with, and in particular, that the scale is open on both its lower and upper ends, and thus has no maximum or minimum points.

Scales feature more explicitly in (1b-d). (1b), for instance, says that the semantic value of *closed* is the function which takes an individual *x* and returns the value True just in case there is a degree *d* which is the maximum value of the scale associated with *closed* and the degree to which *x* is closed is greater than or equal to *d*. The measure function **closed** is just like the measure function **tall**, except that it maps individuals to their degree of openness. Unlike S_{tall} , S_{closed} is closed on both ends, and thus, it has a maximum and minimum value:

$$S_{closed} = \langle D_{[0,1]}, \geq, openness \rangle$$

Whether or not a scale has maximum or minimum values will be important when the associated adjective takes certain modifiers. (1c), for instance, says (very roughly) that *half* takes a gradable property *G* and an individual *x*, and returns True just in case there is a degree *d* which is the midpoint of the scale associated with *G*, and *x* is *G* to degree *d*. But if the scale of the modified adjective is not closed on both ends, it will have no midpoint. Similarly for the semantic value of *completely*, which involves a function (MAX) which outputs the maximum value of the scale of the modified adjective. This feature of *completely* and *half* will be important when we consider explanations of semantic anomaly in §3.2.1.

So now consider sentences like *The jar is half empty* or *The door is completely closed*. Using lexical entries like those given above and Functional Application, we get semantic values like the following:

$$(5.3) \text{ a. } \llbracket \textit{The jar is half empty} \rrbracket = \text{True iff} \\ \exists d[\text{DIFF}(\text{MAX}(S_{empty}), d) = \text{DIFF}(d, \text{MIN}(S_{empty})) \wedge \mathbf{empty}(\text{the jar}) \geq d.]$$

- b. $\llbracket \textit{The door is completely closed} \rrbracket = \text{True}$ iff
 $\exists d[\text{MAX}(S_{\text{closed}}) = d \wedge \mathbf{closed}(\text{the door}) = d.]$

The truth-conditions in (2) capture the intuitive truth-conditions of *The jar is half empty* and *The door is completely closed*. They do so in a somewhat complicated way, but by exploiting our grasp of the formalism, we can see that, for instance, *The door is completely closed* is true just in case the relevant door has the maximum degree of closedness (or equivalently, the minimum degree of openness). These are the truth-conditions that we want.

In the next section, I will argue that the complexity of the specification of truth-conditions in (2) is of particular interest, and that it suggests that lexical entries need to do much more than merely provide a compositional specification of the truth-conditional contents of expressions in order to be explanatorily adequate.

3 The Insufficiency Argument

In this section, I present and defend the qualified version of the Insufficiency Argument—qualified because it will be restricted to explanations of anomaly and entailment relevant to gradable adjectives and degree modifiers only.

3.1 Premise (i): Disquotational and Non-disquotational Analyses of Gradable Adjectives and Degree Modifiers

First, I need to argue for premise (i) of the Insufficiency Argument, which ties the explanatory importance of facts about truth-conditional content to the explanatory power of disquotational lexical entries:

- (i) If the fact that an expression has such-and-so truth-conditional content sufficed

to explain facts about semantic entailment or semantic anomaly relevant to that expression, then disquotational lexical entries would support good explanations of these phenomena.

Consider, again, the complicated specifications of truth-conditions given in (2), above:

- (2) a. $\llbracket \textit{The jar is half empty} \rrbracket = \text{True}$ iff
 $\exists d[\text{DIFF}(\text{MAX}(S_{\textit{empty}}), d) = \text{DIFF}(d, \text{MIN}(S_{\textit{empty}})) \wedge \mathbf{empty}(\text{the jar}) \geq d.]$
- b. $\llbracket \textit{The door is completely closed} \rrbracket = \text{True}$ iff
 $\exists d[\text{MAX}(S_{\textit{closed}}) = d \wedge \mathbf{closed}(\text{the door}) = d.]$

These are the kinds of specifications of truth-conditions we get on Kennedy & McNally's account of the semantics of gradable adjectives and degree modifiers. And, as I mentioned, they succeed in compositionally specifying the right truth-conditions for the sentences in italics. Interestingly, though, they do so in a fairly complicated way, and with respect to specifying extensions, they do no better than simpler, more "disquotational" analyses:

- (5.4) a. $\llbracket \textit{The jar is half empty} \rrbracket = \text{True}$ iff the jar is half empty.
- b. $\llbracket \textit{The door is completely closed} \rrbracket = \text{True}$ iff the door is completely closed.

If this is not obvious, consider how it is that the truth-conditional theorist attributes truth-conditional content to expressions. She pairs a metalanguage expression with an object language expression, then exploits her ability to grasp the truth-conditions of the metalanguage sentence, positing that the object language has the same truth-conditional content as the metalanguage expression. So by exploiting our grasp of the formalism employed in (2), we can judge that, in fact, (2a) and (2b) specify the same truth-conditions as do (3a) and (3b), respectively. And in fact, this should be no

surprise. Where we have a disquotational specification of truth-conditions, as in (3), we can't fail to get the truth-conditions right (provided there aren't contextually sensitive expressions in the sentence). Thus, it would be at least *prima facie* problematic for the non-disquotational analyses if they didn't specify the same truth-conditions as the disquotational analyses.

This raises an interesting question: If the same truth-conditions are specified in (2) and (3), why is it that the best work on gradable adjectives and degree modifiers opts for the complicated specifications of truth-conditions like those in (2)? One thought might be that the complications are necessary to account for compositionality. To some degree, this is correct, since as far as I know, there is no way to compute completely disquotational semantic values for, e.g., *The jar is half empty*, given standard syntax and rules of composition. Maybe, then, we need the complicated lexical entries in (1) to ensure there our derivation of truth-conditions (given in (2)) is compositional.

But while we cannot compositionally compute completely disquotational specifications of truth-conditions like those in (3), compositionality does not force us to complicate our semantics that much. Consider the lexical entries in (4), which are significantly simpler and more “disquotational”—more on what it is to be more or less disquotational shortly—than the entries given by Kennedy & McNally in (1). The entries in (4) will suffice to compositionally specify truth-conditions given standard syntax and Functional Application:

- (5.5) a. $\llbracket \textit{empty} \rrbracket = \lambda x. x \text{ is empty.}$
 b. $\llbracket \textit{closed} \rrbracket = \lambda x. x \text{ is closed.}$
 c. $\llbracket \textit{half} \rrbracket = \lambda G. \lambda x. G(x)=1, \text{ halfway.}$
 d. $\llbracket \textit{completely} \rrbracket = \lambda G. \lambda x. G(x)=1, \text{ completely.}$

Using these lexical entries, we can compute the truth-conditions given in (5):

- (5.6) a. $\llbracket \textit{The jar is half empty} \rrbracket = \text{True}$ iff the jar is empty, halfway.
 b. $\llbracket \textit{The door is completely closed} \rrbracket = \text{True}$ iff the door is closed, completely.

These might sound slightly odd, but they are truth-conditionally equivalent to the truth-conditions given in (2) and (3). The specifications of truth-conditions are not purely disquotational as those in (3) are, but they do not incorporate the structure that the specifications in (2) do—they make no mention of scales, degrees, etc., which figure prominently in (2) and the lexical entries in (1). Instead, they specify truth-conditional content disquotationally, so that ‘completely’ and ‘closed’ appear on both the left and right hand side of the ‘=’ in (5b), for instance. Thus, we can think of the lexical entries in (4) and the specifications of truth-conditions in (5) as being disquotational in the sense that they forego the structure and complexity in (1) and (2) by making use of disquotation. (For our purposes, it will not be necessary to have a precise notion of what it is to be disquotational, so long as we can recognize that the lexical entries in (1) are less disquotational than those in (4).)

The constraint that the truth-conditions of sentences be derived compositionally, then, doesn’t explain why semanticists have moved towards the complex lexical entries like those in (1), and away from lexical entries like the disquotational ones in (4).¹² Why the complexity, then? The answer is simply that, despite the fact that the lexical entries in both (1) and (4) specify the same truth-conditional content, and suffice for the compositional computation of the truth-conditions of the relevant sentences, the lexical entries in (1) are explanatorily superior. They are thus truth-conditionally equivalent, yet empirically distinct. I will argue for this claim in the next section. For now, though, we can simply note that we have good reason to accept premise (i) of the

¹²It may turn out that if we consider a wider range of expressions, the lexical entries in (4) won’t be able to cope with the compositionality constraint. Even if that is the case, it should not be difficult to come up with entries that do satisfy the constraint, but which do not incorporate all of the complexity of the entries in (1).

Insufficiency Argument.

Since the lexical entries in (1) and (4) (and the resulting specifications of truth-conditions in (2) and (5)) specify the same truth-conditional content, if the fact that expressions have such-and-so truth-conditional content is what explained their entailment and anomaly behavior, then both the disquotational and non-disquotational entries would do equally well. This is true even if we add the constraint that the truth-conditions of sentences must be compositional—computable via standard syntax and compositional rules.¹³

3.2 Premise (ii): Explanations of Semantic Anomaly and Entailment For Gradable Adjectives and Degree Modifiers

Now that we have established premise (i) of the Insufficiency Argument, our focus can shift to premise (ii):

- (ii) However, disquotational lexical entries do not support good explanations of semantic entailment or anomaly.

In this section, I will argue for the qualified version of premise (ii) which restricts itself to gradable adjectives and degree modifiers (I argue in §5 that there is nothing special about gradable adjectives and that we have reason to accept the unqualified version). I will begin by arguing that explanations of anomaly and entailment for gradable adjectives and degree modifiers which appeal to disquotational lexical entries are substantially worse than the explanations given by Kennedy & McNally (2005), which appeal to complex, non-disquotational lexical entries. Premise (ii) is, thus, not

¹³There are many versions of the principle of compositionality, but those that are most relevant to semantics in linguistics will require that the semantic values of complex expressions be computable on the basis of the semantic values of simple expressions, syntax, and a handful of semantic composition rules (e.g., Functional Application). For further discussion of different versions of the principle of compositionality, see Szabó (2000, 2012), Pagin & Westerstähl (2010), and XXXXX.

meant to suggest that *no* explanation can be given in terms of disquotational lexical entries, but rather that these explanations are not very good, and so do not explain—in the metaphysical sense of ‘explain’—anomaly or entailment facts.

3.2.1 Semantic Anomaly

I will argue first that disquotational lexical entries do not support good explanations of semantic anomaly for gradable adjectives and degree modifiers.

Non-disquotational Explanations of Anomaly

Consider the cases of semantic anomaly given in (6). The double question marks indicate that they are quite badly unacceptable. They are not, however, ungrammatical, and the explanation is not syntactic.

- (5.7) a. ??Mary is completely tall.
b. ??Mary is half tall.
c. ??The beer is completely expensive.
d. ??Donald is half famous.

What explains the badness of these expressions? Kennedy & McNally point to the lexical entries of the degree modifiers, and the scales of the gradable adjectives that are being modified.¹⁴ Notably, *tall*, *expensive*, and *famous* all have scales which are open on the upper end, and thus lack a maximum value. But one of the things the modifiers *completely* and *half* need is to modify an adjective which has a scale with a maximum value:

¹⁴See Kennedy & McNally (2005, 351-355) for a more thorough presentation than will be given here.

- (1) c. $\llbracket half \rrbracket = \lambda G. \lambda x. \exists d [\text{DIFF}(\text{MAX}(S_G), d) = \text{DIFF}(d, \text{MIN}(S_G)) \wedge G(d) = x]$
 d. $\llbracket completely \rrbracket = \lambda G. \lambda x. \exists d [d = \text{MAX}(S_G) \wedge G(d) = x]$

If it does not, then the function MAX will just be undefined for the relevant scale. So, we can posit that when this happens, the result is anomaly and judgments of unacceptability. And we get a nice generalization as well, viz., that modifiers in which the function MAX takes a scale as an argument cannot modify adjectives with scales that are open on their upper ends.

I will argue in §4 that this kind of explanation is purely non-truth-conditional—i.e. that facts about the truth-conditional contents of expressions do not even partly explain the anomaly data. For now, however, I just want to argue that disquotational lexical entries lack the complexity and structure to give good explanations of the data in (6).

Disquotational Explanations of Anomaly

How might we give an explanation of the data in (6) if we were to make use of the disquotational entries in (4), then? The explanation cannot be the same as in the non-disquotational case since there will be no scales associated with gradable adjectives to appeal to. Thus, there will be no application of MAX to a scale which is not in its domain. Instead, using the lexical entries in (4) and Functional Application, we will derive truth-conditions for (6a) which look like the following:

- (5.8) $\llbracket Mary \text{ is } half \text{ tall} \rrbracket = \text{True at } c \text{ iff Mary's height exceeds the standard for height at } c, \text{ halfway.}$

Now, (7) does get something right. In particular, the metalanguage sentence paired with *Mary is half tall* is totally puzzling, and it's entirely unclear to me what its truth-conditions are. Thus, it attributes puzzling truth-conditions to *Mary is half tall*. The

explanation for anomaly, then, could go as follows. Sentences with entirely unclear truth-conditions are anomalous. *Mary is half tall*, as indicated by the semantic value in (7), has entirely unclear truth-conditions. Explanation complete.

This sort of approach would give us an explanation, for each such sentence like this, of its anomalousness. But the explanation is lacking in depth and generality. We lack the complexity to state any interesting generalizations of the sort that we were able to state with the non-disquotational entries. There, we were able to say that no modifier which makes use of the function MAX can combine with an adjective that has a scale that is open on the upper end. The proposed disquotational explanation does a poor job of saying *why* it is that *half* and *tall* can't combine without unacceptability, it only reflects that they can't. The disquotational explanation is, generally, inferior.¹⁵

Now, since the disquotational entries *do* succeed in specifying the truth-conditional contents of expressions (in a manner that will be compatible with a suitably strong compositionality constraint), we can conclude that the good explanations of anomaly for gradable adjectives and degree modifiers are not purely truth-conditional. That is, we can conclude that the fact that expressions have such-and-so truth-conditional content will not suffice to explain semantic anomaly facts. Good explanations of anomaly, for these expressions, will appeal to non-truth-conditional facts (facts other than facts about truth-conditional contents) about expressions—the structure of the scales they are associated with and the functions of scales employed by degree modifiers, for instance.

3.2.2 Entailment

In order to argue for the qualified form of premise (ii) of the Insufficiency Argument, I also need to argue that facts about the truth-conditional contents of gradable adjectives

¹⁵Similar criticisms can be made of an approach which says that (7) fails to assign any truth-conditions to *Mary is half tall*, and that a lack of truth-conditional content explains anomaly.

tives and degree modifiers (and the sentences containing them) do not suffice to explain the relevant entailment facts.¹⁶ The strategy will be parallel to the one employed in the previous section: I will argue that explanations of entailment which utilize the lexical entries provided by Kennedy & McNally are better than the ones that can be given in terms of disquotational lexical entries. Once again, we will see that it is the non-truth-conditional facts about expressions—facts other than those concerning the truth-conditional contents of expressions—that seem to be doing the work in the good explanations.

Non-disquotational Explanations of Entailment

Consider the following entailment data involving the interaction of gradable adjectives with negation:

(5.9) a. The door is not open. $\dashv\vdash$ The door is closed.

b. The table is not wet. $\dashv\vdash$ The table is dry.

(5.10) a. The table is not large. $\not\vdash$ The table is small.

b. The beer is not expensive. $\not\vdash$ The beer is inexpensive.

What explains the differences in entailments between (8) and (9)? To provide an explanation along the lines of Kennedy & McNally, we need to appeal to antonymy and the standards associated with particular adjectives. I clarify antonymy first.

Each pair of sentences involves an antonym pair (e.g. *open/closed*). In Kennedy & McNally’s terminology, *open*, e.g., is a “positive” gradable adjective, while *closed* is a “negative” gradable adjective.¹⁷ This just means that the measure functions **open** and **closed**, which appear in the semantic values for *open* and *closed*, respectively, map to

¹⁶This is especially noteworthy because if there’s anything truth-conditional semantics is supposed to do well with, it’s entailments. See, e.g., Schiffer (2015), who (following Partee (2011)) takes the explanation of entailments to be the primary aim of truth-conditional semantics.

¹⁷See Kennedy & McNally (2005, 354).

degrees on scales which differ only in that their ordering relations are inverses of one another:

$$\mathbf{open}: f : u \subseteq U \longrightarrow \langle D_{[0,1]}, \leq, \mathit{openness} \rangle$$

$$\mathbf{closed}: f : u \subseteq U \longrightarrow \langle D_{[0,1]}, \geq, \mathit{openness} \rangle$$

And in fact, we can think of what it is for gradable adjectives to stand in an antonymy relation with one another in just this way: two gradable adjectives are antonyms just in case their measure functions map individuals to scales which differ only in that their ordering relations are inverses of one another.¹⁸ We do not need to appeal to the truth-conditional contents of the relevant expressions to understand antonymy.

Now, the difference between the adjective pairs in (8) and those in (9) has to do with the “standards” associated with the adjectives. *Open/closed* and *wet/dry* are what Kennedy & McNally call ‘absolute adjectives’, while *large/small* and *expensive/inexpensive* are ‘relative adjectives’. Relative adjectives (like *tall*) have their standards fixed by context, so that a thing can rightly be called ‘large’ in one context, but not in another. Absolute adjectives, on the other hand, have their standards fixed lexically. And unlike relative adjectives, which may be associated with open scales, they must have scales that are closed on at least one end. For example, *closed* imposes a maximum standard—heuristically, for a door to be closed, it needs to be completely closed (or completely not-open). *Open*, on the other hand, imposes a minimum standard, so that any degree of openness greater than the minimum value on S_{open} suffices for being open. These standards are encoded into the lexical entries for *closed* and *open*:

$$\llbracket \mathit{closed} \rrbracket = \lambda x. \exists d[d = \text{MAX}(S_{\mathit{closed}}) \wedge \mathbf{closed}(x) \geq d]$$

$$\llbracket \mathit{open} \rrbracket = \lambda x. \exists d[d > \text{MIN}(S_{\mathit{open}}) \wedge \mathbf{open}(x) \geq d]$$

¹⁸Kennedy & McNally (2005, 351).

Thus, *The door is not open* entails *The door is closed* and vice versa. Likewise for *wet* and *dry*, since *wet* imposes a minimum standard, while *dry* imposes a maximum standard on a shared scale dimension measuring wetness.

The lexical entries for *closed* and *open* are instances of the templates for maximum and minimum adjectives, respectively.¹⁹ Thus, the lexical entries of *dry*, *empty*, *asleep*, *unknown*, etc., will have semantic values which differ from that of *closed* only with respect to their measure functions. These measure functions, in turn, will all map to a scale which has a closed upper end. Likewise, adjectives like *wet*, *awake*, and *flawed* will have semantic values like that of *open*. They will differ only with respect to their measure functions, all of which map to degrees on a scale with a closed lower end.

At this point, I think there are two different ways of construing the non-disquotational explanations of entailment. One way is to think of the non-disquotational lexical entries as allowing for deeper truth-conditional explanations than would otherwise be possible. These explanations would appeal to what I will call the ‘truth-conditional entailment rule’, which is the standard truth-conditional means of explaining entailment:

$$S_1 \vdash S_2 \text{ just in case } \llbracket S_1 \rrbracket \subseteq \llbracket S_2 \rrbracket$$

The rule states that a sentence S_1 entails S_2 just in case the intension of S_1 (for our purposes, a set of possible worlds) is a subset of the intension of S_2 .²⁰ As I will argue later in this section, this kind of explanation can also be given with disquotational lexical entries. However, the non-disquotational entries provided by Kennedy & McNally allow for greater explanatory depth because we not only can say that the entailment rule applies (or not) to the sentences in (8) and (9), but they give us the tools to explain *why* the entailment rule applies (or not) to those sentences. We can do this by

¹⁹See Kennedy & McNally (2005, 358).

²⁰For the sake of convenience, I will switch to talk of intensions in the context of entailment. The lexical entries for gradable adjectives and degree modifiers that have been presented are extensional, but it would be a trivial exercise to make them intensional.

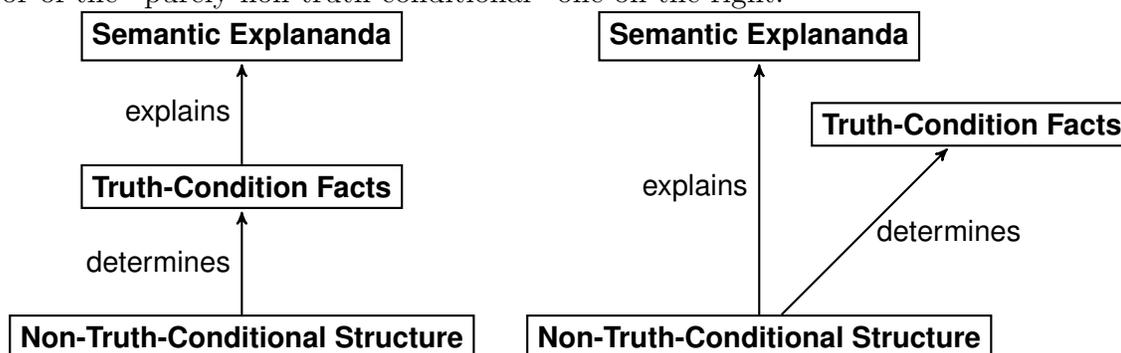
appealing to facts about the standards associated with the adjectives, the antonymy relations that obtain between the gradable adjectives in the sentences, and so on. By appealing to these kinds of facts, we can explain systematic patterns of subset relations obtaining between the intensions of sentences, and thus explain systematic entailment patterns.

This is arguably the kind of explanation of entailment that Kennedy & McNally have in mind.²¹ However, I will not appeal to it in support of premise (ii) of the Insufficiency Argument. The concern is that while the non-disquotational lexical entries might generally be explanatorily superior to disquotational entries, one might worry that the above explanation of *entailment* isn't any better than the one we get from disquotational entries. Both would explain the entailments in (8) by appealing to the fact that the sentences on the left have the same intension as those on the right. What the non-disquotational lexical entries allow for is a deeper semantic explanation for *that* fact—the explanans in the explanation of the entailments in (8). The reason is that, unlike with disquotational lexical entries, we can use non-disquotational structure to provide a semantic explanation of why simple expressions have the intensions they have. Facts about the intensions of simple expressions will, in turn, figure in explanations of why two sentences have the same intension (or why a subset relation obtains between the intensions). It thus might be that non-disquotational lexical entries are explanatorily superior to their disquotational counterparts, then, because they allow for the explanation of a wider range of facts than the disquotational lexical entries, but *not* because they provide better explanations of entailment—the explanans is the same in both cases.

I will, therefore, think of the non-disquotational explanations of entailment differently. In particular, I suggest that we can think of facts about antonymy, standards,

²¹See Kennedy & McNally (2005, 358-360) for their discussion of the relevant entailment patterns.

and so on, as playing a more direct role in explanations of entailments. On the above sketch, they figured in explanations of entailment by being a part of explanations of why sentences have the intensions they have. Entailment facts were more directly explained by these latter facts about the intensions of sentences and the subset relations that obtain between those intensions. I am suggesting, rather, that the explanation of entailment does *not* appeal to facts about intensions or the subset relations that obtain between them, but merely the facts about antonymy, standards, etc., that independently help to determine truth-conditional content. In other words, I argue against the truth-conditional (though not purely truth-conditional) picture on the left, and in favor of the “purely non-truth-conditional” one on the right:



My task right now, however, is to sketch these purely non-truth-conditional explanations of entailment more explicitly. I will argue in §4 that they are, in fact, purely non-truth-conditional. What matters for now is just that they are not disquotational and that they are superior to their disquotational counterparts.

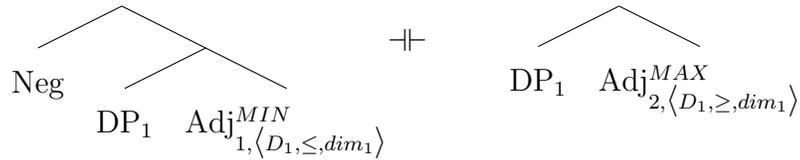
The thought behind the non-truth-conditional, non-disquotational explanation of entailment that I will sketch is just that entailment is to be explained by appeal to entailment rules which are directly sensitive to facts about syntax, antonymy, and the standards associated with the relevant gradable adjectives and degree modifiers. To be clear, these are not rules which Kennedy & McNally explicitly state. Nor do Kennedy & McNally endorse the idea that entailments are explained by appealing to

entailment rules other than the truth-conditional one. However, the entailment rules I will present borrow heavily from Kennedy & McNally’s discussion of the entailment patterns of gradable adjectives and degree modifiers, and the ways that facts about antonymy and standards help to determine these entailment patterns. Consequently, I think of these explanations as being very much friendly to—if not necessarily implicit in—contemporary work in truth-conditional semantics.

I will use ‘Adj_{1,⟨D_{1,≤,dim₁⟩}^{MIN}’ and ‘Adj_{2,⟨D_{1,≥,dim₁⟩}’ to indicate positive and negative antonym pairs, respectively. Superscripted ‘MAX’ and ‘MIN’ indicate maximum and minimum standards. If there is no superscript, then we assume the adjective is a relative adjective. We can capture the entailment data in (8), then, with the following rule:}}

$$(ENT-1): [\text{Neg} [\text{DP}_1 \text{Adj}_{1,\langle D_1, \leq, dim_1 \rangle}^{MIN}]] \dashv\vdash [\text{DP}_1 \text{Adj}_{2,\langle D_1, \geq, dim_1 \rangle}^{MAX}]^{22}$$

Or equivalently:



(ENT-1) provides us with an explanation for the inference patterns in (8). The data in (9) is simply explained by the fact that there is no entailment rule which licenses the entailments in (9) (since the adjectives in (9) have relative standards). The rule is general in that it applies to all antonym pairs of absolute adjectives.

(ENT-1) does not, however, account for the similar entailment patterns we find when we have negation with maximum adjectives, as in (10):

- (5.11) a. The door is not closed. $\dashv\vdash$ The door is open.
 b. The table is not dry. $\dashv\vdash$ The table is wet.

Thus, we can add another entailment rule, (ENT-2), which mostly mirrors (ENT-1):

²²‘DP’, here, stands for ‘determiner phrase’. These include phrases like *Mary, the door, three dogs*, etc. Also, I assume that negation takes scope over DP, though nothing hangs on this assumption.

(ENT-2): [Neg [DP₁ Adj_{2,⟨D_{1,≥,dim₁⟩}^{MAX}]] \dashv [DP₁ Adj_{1,⟨D_{1,≤,dim₁⟩}^{MIN}]}}

(ENT-1) and (ENT-2) do not suffice to explain the entailment patterns we find with degree modifiers, however. Consider, for instance, the following data involving *half* and *partially*:

(5.12) a. The plant is half dead. \vdash The plant is not dead.

b. The glass is partially full. \vdash The glass is not full.

(5.13) a. The door is half open. \vdash The door is open.

b. The table is partially wet. \vdash The table is wet.

The different entailment patterns in (11) and (12) are explained by the facts that *dead* and *full* are maximum adjectives, while *open* and *wet* are minimum adjectives, and that *half* and *partially* effectively return values on the relevant scales that are between the maximum and minimum values (i.e. they do not return the maximum or the minimum).

We can capture the data in (11) and (12) with the following additional rules, where ‘Mod_{<MAX}’ indicates a degree modifier which outputs a value on a scale which is lower than the maximum, and ‘Mod_{>MIN}’ indicates a degree modifier which outputs a value on a scale which is greater than the minimum:

(ENT-3): [DP₁ [Mod_{<MAX} Adj_{1,⟨D_{1,≥,dim₁⟩}^{MAX}]] \vdash [Neg [DP₁ Adj_{2,⟨D_{1,≥,dim₁⟩}^{MAX}]]}}

(ENT-4): [DP₁ [Mod_{>MIN} Adj_{1,⟨D_{1,≤,dim₁⟩}^{MIN}]] \vdash [DP₁ Adj_{2,⟨D_{1,≤,dim₁⟩}^{MIN}]}}

These rules account for the fact that we get the same entailment patterns as in (11) and (12) when we substitute modifiers like *somewhat*, *barely*, *partly*, etc. They also account for the more mundane facts that, e.g., [DP₁ *completely open/wet*] entails [DP₁ *open/wet*], and that [DP₁ *completely dead/full*] does not entail [Neg [DP₁ *dead/full*]].

(ENT-3) and (ENT-4) do not, by themselves, suffice to explain entailments like those in (12):

- (5.14) a. The plant is half dead. \vdash The plant is alive.
b. Mary is half asleep. \vdash Mary is awake.
c. The door is half open. \vdash The door is not closed.
d. The table is partially wet. \vdash The table is not dry.

To get the explanation for these entailments, however, we simply need to apply (ENT-1) and (ENT-2) in addition to (ENT-3) and (ENT-4). The former rules apply to the right hand sides of the latter.

To be clear, my aim here is not to provide an exhaustive account of how we can account for all of the relevant entailments via this approach. I suspect that there are ways to generalize the rules I have suggested here which would account for a wider range of data. The aim is just to sketch some entailment rules which would account for some of the entailment data that Kennedy & McNally canvass. The entailment rules I sketched, in effect, capture what I take to be the crux of their explanations. What is important about these rules, for our purposes, is that they appeal to features of gradable adjectives and degree modifiers which disquotational analyses simply cannot recognize. The explanation of entailment I've sketched, then, is not one which can be supported by disquotational lexical entries.

Disquotational (i.e. Purely Truth-Conditional) Explanations of Entailment

In order to complete the argument for the qualified version of premise (ii) of the Insufficiency Argument, I need to argue that explanations of entailment for gradable adjective and degree modifier sentences that appeal to disquotational lexical entries are inferior to the non-disquotational explanations given above.

The disquotational lexical entries for gradable adjectives and degree modifiers suffice to compositionally specify the truth-conditional contents of expressions, but cannot do much else. The complexity of the non-disquotational entries given by Kennedy & McNally allowed us to appeal to facts about standards, antonymy relations, and properties of degree modifiers in our explanations of entailment. These kinds of explanations will clearly not be available for disquotational entries, since they do not encode facts about standards or facts about scale structure (which are necessary to understand antonymy).

Since disquotational entries do suffice for the compositional specification of truth-conditions of sentences, however, the standard truth-conditional explanation of entailment will be available. In particular, entailments are explained by the truth-conditional entailment rule mentioned above:

$$S_1 \vdash S_2 \text{ just in case } \llbracket S_1 \rrbracket \subseteq \llbracket S_2 \rrbracket$$

Given its generality, this rule is quite powerful, and would seem to provide a parsimonious explanation of natural language entailments. It certainly can account for all of the entailment data we've considered so far, for instance.

But while the truth-conditional entailment rule is extremely general, it is too shallow to allow us to say anything interesting about gradable adjectives and degree modifiers in particular. In the non-disquotational explanation, we were able to state interesting generalizations in the form of entailment rules which were sensitive to the semantic structure of these expressions. This semantic structure (scale structure, standards, etc.) is what goes into the determination of the truth-conditional content of the expressions, and thus figure into an explanation of why the expressions have the truth-conditional contents they have. The non-disquotational explanation, then, is less general, but also less shallow, in that it appeals to facts about expressions which are more basic than facts about their truth-conditional contents.

It's also not clear that the extreme generality of the truth-conditional entailment rule is really a virtue. First, the rule generates *lots* of entailments, including entailments which are typically not recognized by speakers and entailments which we might think just aren't the sorts of entailments that a semantic theory is tasked with explaining. For instance, on the truth-conditional entailment rule, every sentence entails every necessary truth—*Mary hikes* entails *two plus two equals four*—and every necessary falsehood entails every other sentence. However, the vast majority of these entailments will go unrecognized by competent English speakers. This might be a completely acceptable position in logic, e.g., but it is *prima facie* problematic when speakers' judgments serve as the primary source of data for the theory.²³

In addition, the truth-conditional entailment rule would suffice to explain entailments which speakers do recognize, but which might not best be explained by a semantic theory. For instance, it would explain the entailment from *Fido is a dog* to *Fido is a living organism*. This is, of course, a perfectly good entailment, but it might be one which is best explained by appealing to a theory of the concepts DOG and LIVING ORGANISM, mastery of which will likely require general knowledge about dogs and living organisms, or at least a number of beliefs employing DOG and LIVING ORGANISM. But speakers' tacit semantic knowledge, generally, fails to constitute knowledge or beliefs about the things that their expressions are about. Thus, since semantic theories are theories of speakers' tacit semantic knowledge, it would be at least *prima facie* problematic for a semantic theory to posit that this tacit knowledge explained speakers' grasp of these extra-semantic, conceptual entailments.

If we adopt the non-disquotational approach to entailment that I sketched above,

²³One way to restrict the number of entailments generated by truth-conditional semantics is to think of semantics as assigning structured propositions or other hyperintensional structures to sentences. But note that this move more or less acknowledges the failure of the truth-conditional entailment rule in explaining entailment. Structured propositions are more fine-grained than intensions, but they are not more fine-grained representations *of truth-conditions*. See, e.g., Soames (1992, 2012). See also Larson & Ludlow (1993) for the “structured logical form” approach to hyperintensionality.

then the divide between the semantic and the conceptual will arise naturally. Where we can give these kinds of explanations of entailment, the entailment is a semantic one. Where we can do no better than disquotation, and thus a purely truth-conditional explanation, the entailment is a conceptual one. I will have more to say about this view of the semantic-conceptual distinction in §5, but for now, I will just register that the distinction presents a potential problem for truth-conditional explanations of entailment.

Finally, I want to briefly address the major advantage the truth-conditional explanation of entailment rule over the alternative I sketched, namely, simplicity. The concern is that even once we have simplified the kinds of entailment rules I sketched above as much as is possible, we will still require a very large number of these to explain all of the entailments which end up being “semantic” ones. The truth-conditional explanation, by contrast, requires only one entailment rule. Now, to be clear, the potential advantage here doesn’t amount merely to greater *theoretical* simplicity. An increase in that theoretical virtue would not give us sufficient reason to endorse the truth-conditional approach if it turned out, for instance, that it predicted or explained entailments which we have reason to think are simply not semantic explananda. The advantage, rather, is that the truth-conditional approach would appear to be *psychologically* simpler. The argument in favor of the truth-conditional explanation could go roughly as follows. One could claim that the non-truth-conditional explanation would require a number of rules which is simply psychologically implausible or impossible. Thus, whatever other virtues such an explanation might have, it’s simply not a viable contender. The best available explanation, then, is the truth-conditional one, which doesn’t suffer from this problem.

First, we should note that while this is a potential concern for my non-disquotational approach, it’s not clear it offers any support for the truth-conditional one, given the

reasons I've provided for thinking that the explanation is independently bad. Second, it's not clear how plausible the initial premise is. The typical child will amass a vast vocabulary and internalize very quickly and with little or no instruction the phonological, syntactic, and semantic subtleties of a vast number of lexical items.²⁴ It should not be at all obvious, then, that it would be implausible for them, as they master the entailments in their language, to also internalize a large number of entailment rules (which will be quite small in number in comparison to the lexical items). Third, there is a sense in which my non-disquotational explanation of entailment is, in fact, more psychologically simple than the truth-conditional one, at least with respect to the semantic faculty. On that approach, the semantic faculty needs to be sensitive to non-truth-conditional structure which may come in the form of properties of scale degrees, standards, and so on. It does *not* also need to be sensitive to the truth-conditional contents of expressions (as I will, in effect, argue for in §4). On the truth-conditional approach, the semantic faculty will need to be sensitive to truth-conditional content in addition to non-truth-conditional structure which is needed to explain semantic anomaly.

To sum up, the non-disquotational explanations of entailment for gradable adjective and degree modifier sentences that I sketched in §3.2.2 provide better explanations of the relevant entailments than the truth-conditional alternative. Indeed, there is good reason to think that the truth-conditional explanation of entailment is, in fact, problematic, at least for a certain range of entailments. Thus, since disquotational lexical entries can only support these truth-conditional explanations of entailment, we have good reason to think that disquotational lexical entries cannot support good explanations of entailment. We can therefore conclude the qualified form of premise (ii) of the Insufficiency Argument—that disquotational lexical entries do not support

²⁴See, e.g., Guasti (2002).

good explanations of semantic anomaly or entailment, at least with respect to gradable adjectives and degree modifiers. The good explanations are the ones I sketched above which require the non-disquotational lexical entries like those given by Kennedy & McNally.

From the qualified premise (ii), we can reach the conclusion of the qualified Insufficiency Argument: the fact that gradable adjectives and degree modifiers have the truth-conditional contents they have does not suffice to explain the relevant semantic anomaly and entailment facts. Again, if it did so suffice, then disquotational lexical entries—since they do succeed in compositionally specifying truth-conditional content—would be able to support good explanations of these phenomena. But they do not, and so at the very least, there are properties of gradable adjectives and degree modifiers other than their having the truth-conditional contents they have that figure in the relevant explanations.

4 The Superfluity Argument

Having argued that facts about the truth-conditional contents of gradable adjectives and degree modifiers do not suffice to explain the relevant anomaly and entailment data, I want to now argue that these truth-conditional content facts do not even partly explain anomaly or entailment facts. I argue for this claim via a qualified version of the Superfluity Argument. As with the Insufficiency Argument, I will provide reasons to generalize to the unqualified version of the Superfluity Argument in §5.

The (unqualified) Superfluity Argument, again, is as follows:

Superfluity Argument

- (i) If facts about semantic entailment and anomaly are partly explained by facts about the truth-conditional contents of expressions, then the good

explanations of entailment and anomaly will depend on facts about the truth-conditional contents of expressions.

(ii) However, the good explanations of entailment and anomaly do not depend on facts about the truth-conditional contents of expressions.

(iii) Therefore, facts about the truth-conditional contents of expressions do not even partly explain facts about semantic entailment and anomaly.

Premise (i) follows from a basic assumption about explanation which provides a link between two senses of “explain”: the activity sense of “explain”—explanation as something that theorists give—and the metaphysical sense of “explain”—explanation as a relation that obtains between events, properties, facts, etc. Generally, the good explanations for some phenomenon are just the ones that, minimally, get it right, metaphysically speaking.²⁵ So, our assumption is: if the fact that p is partly explained by the fact that q , then good explanations for the fact that p will be ones which depend on the fact that q . Thus, assuming the fact that p is partly explained (metaphysically) by the fact that q , if the proposed explanation does not depend on the fact that q , then it will fail to be a good one.

Now, what does it mean here for an explanation (in the activity sense) of the fact that p to depend on the fact that q ? For my purposes, I will define this dependence as follows: An explanation E for the fact that p depends on the fact that q if and only if the fact that q is part of E 's explanans (the collection of facts which it proposes explains the fact that p), or some facts in E 's explanans obtain in virtue of the fact that q . The thought behind premise (i), then, is that we can check to see if facts about

²⁵Of course, this is not always true of the best *available* explanation of some phenomenon, since we might have no good explanations available. However, it is an assumption that we typically make when, for instance, we report the findings of the sciences as telling us about how the world is. If one is less comfortable about making this assumption in the context of a young field like semantics, we could have a weaker version of the Superfluity Argument which just relied on assuming that we have *good reason to think* that our best explanations get it right.

the truth-conditional contents of expressions partly explain anomaly and entailment facts by seeing whether the good explanations of anomaly and entailment depend (in the relevant sense) on truth-conditional content facts.

How, then, can we test for this dependence? We can do this by supposing that those truth-conditional-content-facts—and facts which obtain in virtue of those facts—don't obtain, and then checking to see whether the non-truth-conditional explanations of anomaly and entailment would still be good. By supposing that the truth-conditional content facts and facts dependent on those facts don't obtain, we ensure that the truth-conditional content facts are not, in fact, a part of the explanans of our non-truth-conditional explanations of anomaly and entailment. If these explanations—which I argued in §3 provide the best explanations of those phenomena for gradable adjectives and degree modifiers—remain good in spite of our suppositions, then we show that facts about the truth-conditional contents of gradable adjectives and degree modifiers do not play any role in our explanations of anomaly and entailment.

Let me briefly explain why we need to be concerned not only with facts about the truth-conditional contents of expressions, but also facts which obtain in virtue of those facts. Consider the truth-conditional explanation of entailment: one sentence entails another because a subset relation obtains between the intensions of the two sentences. The explanans in this case is not a fact about what the truth-conditional contents of the sentences are, but a fact about a relation that obtains between them. But the relation obtains in virtue of the fact that the entailing sentence has the intension it has, and the entailed sentence has the intension it has. If we are to argue, then, that facts about the truth-conditional contents of expressions play no role in explanations of entailment, it is not enough to simply suppose that the sentences have truth-conditions other than they, in fact, have, and show that the explanations of entailment are still good. For the requisite subset relation could obtain between two sentences even if their

truth-conditions were different from what they actually are. In that case, it could still be that the obtaining of that subset relation is what explains entailment, and since in the actual world, that relation obtains in virtue of the fact that the relevant sentences (and their components) have the truth-conditional contents they have, this latter fact would be an important part of our explanation of entailment.

We have, then, a way to test for whether or not the non-disquotational explanations of anomaly and entailment for gradable adjectives and degree modifiers depend on facts about the truth-conditional contents of those expressions. The next step is to give an argument that these explanations do not, in fact, depend on those facts—i.e. we need to argue for the qualified premise (ii).

Semantic Anomaly

Consider, again, the data given in (6):

- (6) a. ??Mary is completely tall.
- b. ??Mary is half tall.
- c. ??The beer is completely expensive.
- d. ??Donald is half famous.

The explanation of these data given in §3.2.1 was that degree modifiers whose lexical entries include ‘MAX(S_G)’ must combine with gradable adjectives that are associated with scales with maximum values. Otherwise, we get semantic anomaly and judgments of unacceptability. *Completely* and *half* both employ MAX, but *tall*, *expensive*, and *famous* are all associated with scales which are open on the upper end, and thus lack a maximum value. Thus, we correctly predict the unacceptability of the sentences in (6).

The explanation does not explicitly appeal to facts about the truth-conditional contents of the expressions involved, but now I will argue that it doesn't implicitly depend on them either.²⁶ What I need to do, then, is show that the non-disquotational explanations are still good, even when we suppose that facts about the truth-conditional contents of gradable adjectives and degree modifiers, and facts which obtain in virtue of those facts, do not obtain.

But what are the truth-conditional facts that might implicitly explain semantic anomaly? While there isn't any standard truth-conditional explanation of semantic anomaly, there are some plausible candidates. For instance, it might be that if the truth-conditions of a sentence are unclear or incalculable in a theory, then the sentence is anomalous. Or it might be that if a sentence has no truth-conditions or is a necessary falsehood, then this could explain its anomalousness.²⁷ I will consider each of these options in turn, arguing either that they are not facts about truth-conditional contents in the relevant sense, or that the non-disquotational explanations of anomaly do not depend on these kinds of facts.

In §3.2.1, I suggested a possible truth-conditional explanation of anomaly, which appealed to the idea that if it is unclear what the truth-conditions are that are attributed to a sentence, it is anomalous. Perhaps, then, facts about the unclarity or incalculability of truth-conditions figure in the explanantia of non-truth-conditional explanations of anomaly. However, neither of these kinds of facts obtain in virtue of facts about the truth-conditional contents of expressions. To say that a sentence has

²⁶Note that the explanations *do* depend, in the relevant sense, on facts like those about the scale structure associated with the adjectives (for instance). The fact that *Mary is completely tall* is anomalous is explained by the fact that it involves the combining of a degree modifier which employs MAX, and a gradable adjective with an open scale. But this fact obtains in virtue of the fact that *tall* is associated with S_{tall} , for example.

²⁷Fodor & Lepore (1998), for instance, don't seem to think that anomaly should be given any uniform treatment. They write that in some cases, it will be explained by the fact that the sentence is a necessary falsehood, but in other cases it may be better explained by appeals to ungrammaticality or pragmatic unacceptability.

unclear truth-conditions in this context is to say that a theorist is unable to recognize what they are. Likewise for the fact that the truth-conditions of a sentence are not calculable in some semantic theory. These kinds of facts can obtain regardless of what the truth-conditions of sentences actually are. Thus, these explanations do not actually depend on facts about the truth-conditional contents of expressions, and so even if facts about the unclarity or incalculability of truth-conditions do play some role in the non-truth-conditional explanations of anomaly, this does not conflict with the conclusion of the Superfluity Argument. The facts are not truth-conditional content facts in the relevant sense, and so we can put them aside.

Another possibility is that the anomalous sentences have no truth-conditions, and that this explains why they are anomalous. Let's suppose, for the sake of argument, that the fact that a sentence has no truth-conditional content is a fact about the truth-conditional content of that sentence. Consider the truth-conditions we arrive at if we calculate the truth-conditions of *Mary is completely tall* using the lexical entries of Kennedy & McNally:

$$(5.15) \llbracket \textit{Mary is completely tall} \rrbracket = \text{True in } w \text{ iff in } w, \exists d[d = \text{MAX}(S_{\textit{tall}}) \wedge \mathbf{tall}(\textit{Mary}) = d]$$

Arguably, (15) fails to assign any truth-conditional content to *Mary is completely tall*. After all, $\text{MAX}(S_{\textit{tall}})$ is undefined, and so one might think that the value of ' $d = \text{MAX}(S_{\textit{tall}})$ ' is similarly undefined, and likewise for the entire statement of truth-conditions in (15). In that case, it would seem that according to our best semantics for gradable adjectives and degree modifiers, *Mary is completely tall* simply has no truth-conditions.

Further, since the non-disquotational explanations of the data in (6) requires the combining of a degree modifier like *completely* or *half* with a gradable adjective whose associated scale has no maximum value, one might think that the non-disquotational

explanans will obtain if and only if the relevant sentence has undefined truth-conditions. In that case, I would not be able to show that the non-disquotational explanations of anomaly do not depend on the fact that the anomalous sentences have no truth-conditions.

This is too quick, however. It's true that we need 'MAX(S_{tall})' to show up in our specifications of the truth-conditions of *Mary is completely tall* in order for the non-disquotational explanation to work, but that doesn't necessitate that the truth-conditions of the sentence will then be undefined. For instance, suppose we define a new identity function ' $=_{und}$ ' which is just like the ordinary '=', except that when we have ' $x =_{und} f(y)$ ', where $f(y)$ is undefined, the sentence is false. In that case, we can alter the lexical entry for *completely* as in (16a), and arrive at truth-conditions for *Mary is completely tall* as in (16b):

- (5.16) a. $\llbracket \textit{completely} \rrbracket = \lambda G. \lambda x. \exists d[d =_{und} \text{MAX}(S_G) \wedge G(d) = x]$
 b. $\llbracket \textit{Mary is completely tall} \rrbracket = \text{True in } w \text{ iff in } w, \exists d[d =_{und} \text{MAX}(S_{tall}) \wedge \mathbf{tall}(\textit{Mary}) = d]$

If (16a-b) were true, then since d will always be some value between 0 and 1, and $\text{MAX}(S_{tall})$ will be undefined, *Mary is completely tall* would be a necessary falsehood. Thus, it would have an intension (i.e. truth-conditions). The non-disquotational explanation would remain as before, however, since we still apply a degree modifier containing 'MAX(S_G)' in its lexical entry with a gradable adjective whose associated scale lacks a maximum value. In that case, 'MAX(S_G)' is undefined and we can appeal to our principle which says that whenever, during the course of semantic composition, 'MAX' is applied to a scale which is not in its domain, we get anomaly and judgments of unacceptability. But since the non-disquotational explanation goes through even when the purported truth-conditional explanans does not obtain, we have shown that

the former does not depend on the latter. Given that we can come up with similar examples for the rest of the anomaly data we've considered, we can conclude that the fact that anomalous sentences lack truth-conditions (if that is a fact) does not even partly explain facts about their anomalousness.

That the non-disquotational explanation remains good whether (15) or (16b) is true is enough to show that the explanation does not depend on the fact (again, if it is one) that *Mary is completely tall* is a necessary falsehood either. In that case, it would seem that the non-truth-conditional explanations of anomaly do not depend on any of the candidate truth-conditional facts. Without any other contenders, we can conclude that the non-disquotational explanations of anomaly for gradable adjectives and degree modifiers do not depend on facts about their truth-conditional contents. And so facts about their truth-conditional contents do not even partly explain the relevant anomaly data.

Semantic Entailment

In the case of semantic anomaly, the purported truth-conditional explanantia—that anomalous sentences lack truth-conditions or that they are necessary falsehoods—were facts about the actual truth-conditional contents of expressions. In the case of entailment, however, facts about the actual truth-conditional contents of expressions are not part of the explanans of the direct, truth-conditional explanation of entailment facts. Rather, what matters is that a particular relation—the subset relation—obtains between the truth-conditional contents of expressions (once again, I will be thinking of these as intensions—sets of possible worlds).²⁸ The fact that a subset relation obtains

²⁸We could suppose, for instance, that *open* has the intension of *wet* and *closed* has the intension of *dry* and vice versa, and the truth-conditional entailment rule would still correctly predict the data in (8):

between the intensions of two expressions, however, obtains in virtue of the fact that those expressions have the intensions that they actually have. Thus, in order to provide support for premise (ii) of the qualified Superfluity Argument, I need to show that the non-disquotational explanations of entailment remain good even when the intension of the entailing sentence is not a subset of the intension of the entailed sentence. In that case, we could be sure that the non-disquotational explanations do not depend on the facts that the relevant sentences have the intensions they have. I will argue next that there are many such examples, although, as I will explain, they are limited to a particular subclass of the entailment data.

The explanation for the entailment data in (8), as I argued in §3.2.2, was that speakers have tacit knowledge of the entailment rule (ENT-1) and the fact that the rule applies to the sentences in (8).

- (8) a. The door is not open. $\dashv\vdash$ The door is closed.
 b. The table is not wet. $\dashv\vdash$ The table is dry.

$$\text{(ENT-1): } [\text{Neg} [\text{DP}_1 \text{ Adj}_{1, \langle D_1, \leq, dim_1 \rangle}^{MIN}]] \dashv\vdash [\text{DP}_1 \text{ Adj}_{2, \langle D_1, \geq, dim_1 \rangle}^{MAX}]$$

(ENT-1) is sensitive to the syntactic structures of the sentences involved, and particular semantic features of the gradable adjectives.²⁹ In particular, it applies to the sentences in (8a) in virtue of the facts that *open* and *closed* are antonyms, that *open* has a minimum standard, and that *closed* has a maximum standard, and so on. Two gradable adjectives are antonyms, recall, if their associated scales order along the same dimension (in this case, openness) and their scales employ inverse ordering relations. Now, unlike in the anomaly case, scale dimensions are relevant to the explanation. However, what

-
- (8) a. The door is not open. $\dashv\vdash$ The door is closed.
 b. The table is not wet. $\dashv\vdash$ The table is dry.

²⁹Similar remarks hold for the other entailment rules, except that (ENT-3) and (ENT-4) are also sensitive to facts about degree modifiers.

matters is simply that the adjectives have the same scale dimension. It does not otherwise matter what that dimension is.

For instance, we could alter the lexical entry for *open* and *closed* so that they employ measure functions **open*** and **closed***, which are just like **open** and **closed**, but they map to the dimension dim_1 , not *openness*. I will stipulate, for the sake of the example, that these measure functions map everything to a value of 1 along the dimension dim_1 . In fact, we can stipulate that it's not possible for anything to fail to have a value of 1, so that sentences of the form *x is completely open* are necessary truths.³⁰

We can call the resulting scales associated with *open* and *closed* ' S_{open*} ' and ' $S_{closed*}$ ', as in (16):

$$(5.17) \quad \llbracket open \rrbracket = \lambda x. \exists d[d = \text{MIN}(S_{open*}) \wedge \mathbf{open}^*(x) > d]$$

$$\llbracket closed \rrbracket = \lambda x. \exists d[d = \text{MAX}(S_{closed*}) \wedge \mathbf{closed}^*(x) \geq d]$$

$$\mathbf{open}^*: f : u \subseteq U \longrightarrow \langle D_{[0,1]}, \leq, dim_1 \rangle$$

$$\mathbf{closed}^*: f : u \subseteq U \longrightarrow \langle D_{[0,1]}, \geq, dim_1 \rangle$$

Since *The door is closed* and *The door is not open* have the same intension (they are both necessary falsehoods), we do correctly predict the data in (8a). We have thus failed to break the subset relation between the sentences. In fact, while I do not offer a proof for this claim, it seems to me that there are no such examples to be had for any of the entailments that were considered in §3.2.2. In other words, the claim is that if $S_1 \vdash S_2$ according to the non-disquotational explanation of entailment, then $S_1 \vdash S_2$ according to the truth-conditional explanation (i.e. the intension of S_1 is a subset of that of S_2). The reason is just that the non-truth-conditional features of sentences that

³⁰One might wonder what the point of having such a dimension and corresponding lexical entry would be. The point of the example, however, is just to show that facts about the relations between truth-conditional contents don't figure into what I've called the 'purely non-truth-conditional' explanations of entailment. Thus, the implausibility of such a lexical entry is beside the point.

(ENT-1)-(ENT-4) are sensitive to—syntax, antonymy relations, standards, etc.—while they do not suffice to determine truth-conditional content, go a long way toward doing so. Truth-conditional content is only fully specified once we fill in what the dimension of what relevant scale is, but even without doing that, we determine a lot about the relations between the intensions of sentences involving the relevant gradable adjectives and degree modifiers.

But while it might be that there is no way to alter the scale dimensions associated with expressions so as to preserve non-disquotational explanations of entailments and sever the subset relation between the intensions of the relevant sentences, there is another class of entailment facts to be considered. In particular, we not only want to explain why sentences entail one another, but we also want an explanation of why sentences *fail* to entail other sentences. That is, we want an explanation for negative entailment facts like those given in (17):

- (5.18) a. The door is closed $\not\vdash$ The door is open.
b. The door is completely closed $\not\vdash$ The door is completely open.
c. The door is half open $\not\vdash$ The door is closed.
d. The door is half open $\not\vdash$ The door is completely open.
e. The door is not open $\not\vdash$ The door is partly open.

The non-disquotational explanation of negative entailment facts like these, as I suggested previously, is just that there are no entailment rules like (ENT-1)-(ENT-4) which license the entailments from the sentences on the left hand side in (17) to those on the right. The truth-conditional explanation for facts like these is just that the intensions of the sentences on the left are not subsets of those on the right. Once again, the fact that a subset relation does not obtain between two sentences obtains in virtue of the facts that the sentences have the intensions they have, and so truth-conditional

explanations of negative entailment facts depend on facts about the truth-conditional contents of sentences.

However, if we alter the truth-conditional contents of *open* and *closed* as in (16), then *The door is closed*, *The door is not open*, and the other sentences on the left hand side of (17) are all necessary falsehoods. In that case, the fact that (for instance) the intension of *The door is closed* is not a subset of the intension of *The door is open*, would not obtain. *Mutatis mutandis* for (17b-d). The truth-conditional explanantia for (17), then, would not obtain—the truth-conditional entailment rule would predict that the sentences on the left hand side entail the sentences on the right. The non-disquotational explanation remains as before, however, since the relevant entailment rules are simply not sensitive to the scale dimension of the relevant adjectives. If no entailment rule licensed an entailment between the sentences in (17), altering the relevant scale dimensions will not change this.

We therefore have an example where the non-disquotational explanation of an entailment fact holds even though the relevant truth-conditional explanans (the fact that the intension of some sentence is not a subset of the intension of the other) does not obtain. Similar sorts of examples would be easy to construct for other sentences involving gradable adjectives and the relevant negative entailment facts. Thus, we can conclude that the non-disquotational explanations of negative entailment facts relevant to gradable adjectives and degree modifiers do not depend on facts about their truth-conditional contents. And in that case, facts about the truth-conditional contents of gradable adjectives and degree modifiers do not even partly explain negative entailment facts.

As I claimed above, however, there are not similar examples to be had with the positive entailment data. Thus, one might still hold that even though facts about the truth-conditional contents of expressions do not factor into explanations of negative

entailment facts, they do so for positive entailment facts. After all, it might still be a necessary condition on a positive entailment fact obtaining that the intension of the entailing sentence be a subset of the intension of the entailed sentence.

This strikes me as being a very implausible position, however. One would want to know why it is that facts about the truth-conditional contents of expressions are relevant to positive entailment facts but not negative entailment facts. And indeed, it's hard to imagine what the explanation for this could possibly be. More likely, it seems to me, is that while the obtaining of the non-disquotational explanans of a positive entailment fact may obtain only when the purported truth-conditional explanans obtains, facts of the latter sort do not figure into explanations of the relevant entailment data. Merely pointing to the fact that it is a necessary condition on the non-truth-conditional explanans obtaining that the purported truth-conditional explanans obtains is extremely weak evidence that the latter sorts of facts figure into explanations of the relevant entailment data (since, e.g., it is also a necessary condition on the non-disquotational explanans obtaining that $2+2=4$).

Indeed, it is worth stressing that even if wherever we predicted entailment (and negative entailment) via the non-disquotational explanation, the truth-conditional explanans obtained as well, it would not follow that premise (ii) of the Superfluity Argument is false. In fact, given that the non-disquotational explanations I gave for entailment in §3.2.2 made no apparent appeal to facts about the truth-conditional contents of gradable adjectives and degree modifiers, the burden of proof was already on the objector to show that those explanations do depend on those facts. Thus, while my case against the dependence of the non-disquotational explanations of positive entailments on facts about the truth-conditional contents of expressions is not as decisive as in the case of negative entailments, this is no reason to think that it is at all plausible that that dependence holds.

I take it, then, that we have good reason to accept the qualified form of premise (ii) of the Superfluity Argument. Even if we remove facts about the actual truth-conditional contents of gradable adjectives and degree modifiers from our explanations (and facts which obtain in virtue of those facts), we still correctly predict the anomaly and entailment data for gradable adjectives and degree modifiers. Our explanations of anomaly and entailment for these expressions, then, do not depend on facts about their truth-conditional contents. We thus have the qualified version of (iii): facts about the truth-conditional contents of gradable adjectives and degree modifiers do not even partly explain the relevant facts about anomaly and entailment. In other words, truth-conditional-content-facts do no work in these explanations.

5 Generalizing the Arguments: Disquotation and Semantic Explanation

Thus far, I have argued that facts about truth-conditional contents play no role in explanations of anomaly and entailment for gradable adjectives and degree modifiers. I now want to generalize beyond gradable adjectives and degree modifiers and beyond anomaly and entailment—i.e. I want to offer support for the unqualified Insufficiency and Superfluity arguments and conclude that facts about truth-conditional contents play no role in explanations of semantic phenomena generally. I will also explain why accepting these arguments does not commit us to rejecting the progress made in truth-conditional semantics. What is required is merely a reinterpretation of that progress.

The best way to argue for the unqualified Insufficiency and Superfluity Arguments, I think, would be to proceed case by case, comparing explanations of semantic phenomena for different kinds of expressions, arguing in each case that non-disquotational, non-truth-conditional explanations—i.e. explanations that do not depend on facts

about truth-conditional content—are better than their truth-conditional counterparts in each case. Practical issues aside, this approach will not work because, in fact, much of the relevant semantic research is recent or ongoing. Thus, in many cases, the kinds of non-truth-conditional explanations like the ones I’ve sketched will simply not be developed yet. This, of course, does not mean that such explanations will not be forthcoming, and the trend in semantics towards complex, non-disquotational analyses makes me think that they will be.

However, we can offer additional reasons for thinking that, generally, good explanations in semantics will not appeal to disquotational lexical entries, and thus that the good explanations will at least not be purely truth-conditional. There are also reasons—I will argue—to think that, generally, facts about the truth-conditional contents of expressions will not figure in semantic explanations at all, just as in the case of gradable adjectives and degree modifiers.

Here, I want to draw on recent work by Glanzberg (2014), who argues that semantics offers good explanations where it appeals to mathematical, non-disquotational structure, and that where it resorts to disquotation, it loses its explanatory power altogether. If this is right, then the unqualified premise (ii) of the Insufficiency Argument follows. The unqualified premise (ii) of the Superfluity Argument also follows provided we assume that disquotational lexical entries suffice to specify the truth-conditional contents of expressions and that if facts about truth-conditional contents did any explanatory work, then so would disquotational lexical entries. From here, we can get the unqualified conclusions of both of the main arguments.

Let me first assess Glanzberg’s reasons for thinking that disquotation marks an absence of explanatory power in semantics, and offer some reasons of my own.

Glanzberg offers two main reasons for thinking that disquotational lexical entries do not support good semantic explanations. The first is that disquotational entries

are explanatorily weak because “they can be generated with only minimal knowledge of the grammatical category an expression falls under, so long as the metalanguage in which the theory is being given includes the object language” (268).³¹ So, for instance, if one knows that *happy* is an adjective, one knows enough to be able to produce a disquotational lexical entry for it:

$\llbracket happy \rrbracket = \lambda x. x \text{ is happy.}$

It’s unclear to me, however, why this is a reason to think that disquotational explanations of semantic phenomena are weak. If the claim were that a speaker’s tacit knowledge of the syntactic properties of an expression sufficed for tacit knowledge of what is stated by the disquotational lexical entry, then the objection would be a good one. In that case, the semantics would be attributing tacit knowledge to speakers that would seem to be entirely redundant, given their syntactic knowledge. The attributed tacit knowledge could only figure in syntactic explanations, and even then, would be unnecessary. However, it’s simply not true that tacit knowledge of the syntactic properties of an expression suffices for tacit knowledge of what is stated by disquotational lexical entries. A speaker could, for instance, have tacit knowledge of the syntactic properties of *happy*, but tacitly “believe” that $\llbracket happy \rrbracket = \lambda x. x \text{ is sad.}$

Now, it is true that the *theorist* can generate disquotational lexical entries rather easily, provided they know the grammatical category of the expression, but it doesn’t follow that disquotational lexical entries are thereby explanatorily weak. What the objection highlights, I think, is that one could know (explicitly) that a disquotational specification of truth-conditional content for an expression is correct without being competent with the word (beyond its grammatical category). But to have tacit knowledge of what is stated by a disquotational lexical entry is to have tacit knowledge of its semantic type and its truth-conditional content. Unless we already assume that facts

³¹Glanzberg cites Higginbotham (1989) as the source of this point.

about truth-conditional content don't do much explanatory work, we haven't been given reason to think that disquotational lexical entries are explanatorily weak.

The other reason Glanzberg offers for thinking that disquotational lexical entries are explanatorily weak is that they amount to merely listing facts about reference or truth-conditions (268-269).³² The complaint here is that the explanations one can give in terms of disquotational entries are shallow because, while they succeed in listing referents and truth-conditions, they don't offer insight into why expressions have the referents or truth-conditions they have.

Now, while I certainly agree that disquotational explanations are objectionably shallow, I don't think the problem is that disquotational lexical entries simply *list* the truth-conditional contents of expressions. Rather, the problem has to do with what is listed—namely, facts about the truth-conditional contents of expressions. If, on the other hand, we were to list facts about the scales associated with gradable adjectives and their structure, for instance, we could offer the good explanations of semantic phenomena given by Kennedy & McNally. Thus, it's not listing, per se, which makes disquotational lexical entries explanatorily weak.

Nor, however, is it that they fail to show why expressions have the truth-conditional content they have. While it's true that facts about the structure of the scale associated with a gradable adjective, for instance, help determine facts about its truth-conditional content, the Superfluity Argument shows that our semantic explanations are just as good, even if expressions are assigned the wrong truth-conditional content. Thus, to provide good semantic explanations, it's not necessary that we have lexical entries which assign properties to expressions which go into the determination of their referents or truth-conditions.³³ To put it another way, lexical entries which encode

³²This objection originally comes from Field (1972).

³³It is worth repeating here that to attribute truth-conditional content to expressions is to make a theoretical posit, and thus facts about truth-conditional contents are not explananda of the theory (except in the sense in which we expect a compositional explanation of the truth-conditional contents

scale-structure facts (for instance) are not better because they allow for deeper truth-conditional explanation. On that view, scale structure facts would help explain truth-condition facts which, in turn, would explain facts about anomaly and entailment, for instance. What I am arguing, however, is that these more complicated lexical entries are better because they encode facts about scale structure and facts about scale structure figure *directly* into explanations of anomaly and entailment. It's true that they also, independently, determine facts about truth-conditional contents of expressions, but those latter facts do not enter into explanations of semantic phenomena.

Glanzberg is mostly right, however, when he says that by resorting to disquotational lexical entries, "You thus fail to learn any non-trivial generalizations, make any non-trivial predictions, or do anything else that might figure into offering explanations in semantics" (269). True, the truth-conditional entailment rule is a non-trivial generalization that allows us to make non-trivial predictions about entailment, and we might be able to make some generalization about contradictions, for instance. But we can't do much else. The reason is just that disquotational lexical entries only really tell you two things about an expression: its semantic type and its truth-conditional content. Facts about an expression's semantic type are more or less determined by its syntactic properties and tend to appear, if anywhere, in explanations of ungrammaticality. We are thus left trying to explain all of the relevant semantic phenomena by appealing to the truth-conditional contents of expressions—sets of individuals, properties, possible worlds, etc.

As we have seen, there are ways to offer explanations of anomaly and entailment, for instance, which appeal to facts about the truth-conditional contents of expressions, and the set-theoretic relations that obtain between those contents. However, it is unsurprising that these explanations end up being shallow, or generally inadequate

of complexes).

because there are just not that many properties of the sets or functions that constitute truth-conditional contents that we can plausibly appeal to in semantic explanations. Aside from appeal to subset relations in explaining entailments, we might, for instance, think that certain kinds of unacceptability can be explained by the fact that a sentence is a necessary falsehood.³⁴ But beyond that, if we restrict ourselves to trying to explain semantic phenomena in terms of facts about truth-conditional contents, we leave ourselves with a very impoverished explanatory toolkit, insufficient to capture many of the interesting differences between, say, different gradable adjectives. It is not terribly surprising, then, that disquotational lexical entries—which can only offer purely truth-conditional explanations—tend to be explanatorily shallow and generally inadequate in comparison with the non-disquotational lexical entries offered by, for example, Kennedy & McNally.

I think, then, that there is good reason to agree with Glanzberg that, generally, disquotational lexical entries do not support good explanations of semantic phenomena. My reason for thinking so, however, is that they only support purely truth-conditional explanations, and those kinds of explanations are unlikely to be the best ones, given how few explanatory resources facts about truth-conditional contents supplies us with.³⁵

Now, this is only reason to think that, generally, disquotational—and thus *purely* truth-conditional—explanations are inadequate. It doesn't follow that we shouldn't have facts about truth-conditional contents in our explanatory toolkits. As I've argued, however, truth-conditional explanations of anomaly and what I called 'semantic entailment' are simply not very good. I think we have reason to doubt that the

³⁴This seems to be the way that Fodor & Lepore (1998) think that semantic anomaly is to be explained, when the explanation is semantic. They suggest that otherwise, the explanation has to do with ungrammaticality or pragmatic unacceptability.

³⁵Glanzberg would not seem to endorse this claim. He seems to think, rather, that the good explanations are truth-conditional, but that where we have disquotation, semantics fails to determine truth-conditional content. Thus, disquotational lexical entries do not support good explanations because they mark expressions where the semantics fails to fully determine their truth-conditional contents.

truth-conditional entailment rule, for instance, ever figures in good explanations of the kinds of entailments that semanticists are interested in explaining. Further, the Superfluity Argument shows that despite the fact that non-disquotational lexical entries suffice to determine truth-conditional content, the fact that an expression has the truth-conditional content it has does not even partly explain anomaly and entailment.

One might rightly object here that I have not given any reason to think that there are, or will be, other semantic phenomena apart from anomaly and entailment that will, in fact, be amenable to a truth-conditional explanation (or at least a partially truth-conditional explanation). In response, I want to make a suggestion similar to one made by Glanzberg in his discussion of disquotation in semantics. His claim, again, is that where semantics resorts to disquotation, it loses its explanatory power—there are no good disquotational explanations of semantic phenomena. However, he does not think that because truth-conditional semantics often resorts to disquotation it is thereby *defective*. Indeed, he thinks that disquotation is ultimately ineliminable from semantics. His suggestion, then, is that disquotation in semantics marks where semantic explanation runs out because the semantic *explananda* run out.

Consider, for instance, the disquotational lexical entry for *dog*: $\llbracket dog \rrbracket = \lambda x. x$ is a dog. It might simply be that with respect to *dog*, semantics simply isn't going to have much more to say. It isn't, for instance, going to be able to provide a good explanation for why *Fido is a dog* entails (in some sense) *Fido is a living organism*.³⁶ Glanzberg's thought here is that this kind of entailment fact is not a semantic explanandum, but a conceptual one. That is, it is facts about the conceptual faculty, distinct from the semantic faculty (and the language faculty altogether), which explains why we judge

³⁶There are non-truth-conditional semantic theories, e.g. Jackendoff (1990) and Pustejovsky (1995), which would provide explanations of semantic anomaly facts like the fact that *Mary read the dog* is semantically anomalous. But facts like these are typically not addressed by truth-conditional semanticists.

the entailment to hold.³⁷

My suggestion, then, is that the kind of non-disquotational, non-truth-conditional explanations that we saw in the case of gradable adjectives and degree modifiers are the kinds of explanations that are distinctive of semantics. The explananda that are paradigmatically semantic—the ones that are *directly* explained by semantic theories—then, are the ones that get these non-disquotational, non-truth-conditional explanations. Like Glanzberg, I think we ought to take semantic theorizing (to some degree) at face value, and that disquotation does not obviously mark any defect in the theory. But since disquotational lexical entries don't support semantic explanations, and since they *do* suffice to specify truth-conditions as well as non-disquotational entries, I think we ought to conclude that truth-conditional explanations are not semantic—i.e. do not figure in the direct explanation of semantic phenomena like anomaly and entailment.

Note that it doesn't follow that facts about the truth-conditional contents of expressions do not explain anything whatsoever. Indeed, it wouldn't be surprising if explanations of entailments or truth-value judgments involving expressions like *dog*, for instance, end up appealing, at least in part, to its truth-conditional content. Perhaps more interestingly, it also does not follow that lexical entries do not need to still specify the truth-conditional contents of expressions. It could be, along the lines of Glanzberg's view, that lexical entries need to encode properties of an expression which are both semantic and conceptual.³⁸ Whether they do or not depends on the nature of the conceptual faculty, and on the ways in which semantic and conceptual information are stored in the lexicon. The point here is just that the fact that an expression has the truth-conditional content it has is not a semantic fact, in the sense that the fact

³⁷Glanzberg (284-285) agrees that many entailment judgments, and perhaps all truth-value judgments will not have purely semantic explanations. See also Pietroski (2010) for a distinct, but nearby account of why our ability to make these judgments depends on more than just our semantic competence.

³⁸For more on how the non-disquotational, semantic structure, and the disquotational, conceptual components of lexical entries might interact, see Glanzberg (2011).

does not figure directly in explanations of semantic phenomena.

Now, my suggestion may simply seem like a way of stipulating away counterexamples, but I am not simply stipulating that truth-conditional explanations are non-semantic. Rather, the idea is that—in the case where we find that the best explanation for some fact appeals to facts about truth-conditional contents of expressions—our theorizing will have uncovered two very different kinds of explanations for a class of phenomena. This would give us good grounds for thinking that, in fact, there are two different sorts of phenomena in question. Consider, for instance, the semantics/pragmatics distinction. What makes facts about implicatures, say, different from facts about semantic anomaly? Both certainly have a claim to being about “meaning” in some sense, but the explanations for the two phenomena are quite different. In the case of implicature, for example, we might invoke a speaker’s grasp of conversational conventions and their knowledge of what another speaker knows or believes, none of which will we invoke in explaining semantic anomaly. Thus, we think of implicatures as being a different kind of phenomenon from semantic anomaly, and one that whose explanation is not completely semantic.

I suggest, then, that we have good reason to accept the unqualified versions of the Insufficiency and Superfluity Arguments. In fact, we have reason to generalize even beyond those arguments, which restrict themselves to facts about anomaly and semantic entailment. What I am suggesting is that there is good reason to think that the semantic phenomena, generally, will get the sorts of non-truth-conditional explanations that we saw for anomaly and entailment with gradable adjectives. Anomaly and entailment may make up a core class of the semantic data to be explained, but there is no reason to assume that there won’t be other sorts of semantic phenomena. Whatever other phenomena are explained by other kinds of facts—truth-conditional content facts, for instance—are non-semantic. Thus, if this way of carving up the semantic and the

extra-semantic is right, we have reason to think that facts about the truth-conditional contents of expression do not partly explain any semantic facts whatsoever.

6 Conclusions

Where does this leave us with respect to the place of truth in linguistic semantics? At the very least, the role is less central and less substantial than is assumed. If there is a constraint on semantic theories to specify truth-conditions, this constraint does not come from the need to explain semantic explananda. The constraint may come from other sources, however—perhaps the relation the semantic faculty bears to the conceptual faculty or to other faculties of mind.

I think the arguments in this chapter should make one wonder, however, about whether there is such a constraint on semantics. At the very least, one should not assume that the explanatory success of truth-conditional semantics is sufficient to warrant it. In that case, the argument for truth-conditional semantics from the success of the truth-conditional semantics research program is, perhaps, not as strong as it otherwise would be. If the explanatory success of truth-conditional semantics is explained by features of the theory that are, in the relevant sense, “non-truth-conditional”, then the door remains open to non-truth-conditional approaches to semantics in generative linguistics.

Bibliography

- [1] Adger, David (2003). *Core Syntax: A Minimalist Approach*. Oxford University Press.
- [2] Alston, William (1999). *Illocutionary Acts and Sentence Meaning*. Cornell University Press.
- [3] Brandom, Robert B. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press.
- [4] Brandom, Robert (2007). "Inferentialism and Some of its Challenges". *Philosophy and Phenomenological Research* 74 (3): 651-676.
- [5] Burge, Tyler (1979). "Individualism and the Mental". *Midwest Studies in Philosophy* 4 (1): 73-122.
- [6] Burge, Tyler (1988). "Individualism and Self-knowledge". *Journal of Philosophy* 85 (November):649-63.
- [7] Cappelen, Herman & Lepore, Ernest (2005). *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*. Wiley-Blackwell.
- [8] Cappelen, Herman (forthcoming). "Why Philosophers Shouldn't Do Semantics", *Review of Philosophy and Psychology*.

- [9] Carston, Robyn (1988). “Implicature, Explicature, and Truth-Theoretic Semantics”. In Ruth M. Kempson (ed.), *Mental Representations: The Interface Between Language and Reality*. Cambridge Univ. Press. 155–181.
- [10] Chomsky, Noam (1957). *Syntactic Structures*. Mouton.
- [11] Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. MIT Press.
- [12] Chomsky, Noam (1980). *Rules and Representations*. Oxford: Basil Blackwell.
- [13] Chomsky, Noam (1986). *Knowledge of Language*. Prager.
- [14] Chomsky, Noam (1995a). *The Minimalist Program*. The MIT Press.
- [15] Chomsky, Noam (1995b). “Language and Nature”. *Mind* 104 (413): 1-61.
- [16] Chomsky, Noam (2000). *New Horizons in the Study of Language and Mind*. Cambridge University Press.
- [17] Cohnitz, Daniel (2005). “Is Compositionality an A Priori Principle?” In M. Wenging, E. Machery & G. Schurz (eds.), *The Compositionality of Concepts and Meanings: Foundational Issues*. Ontos.
- [18] Collins, John M. (2005). “Faculty Disputes”. *Mind and Language* 19 (5): 503-33.
- [19] W. Culicover, Peter & Jackendoff, Ray (2005). *Simpler Syntax*. Oxford University Press UK.
- [20] Davidson, Donald (1965). “Theories of Meaning and Learnable Languages. In Yehoshua Bar-Hillel (ed.), *Proceedings of the International Congress for Logic, Methodology, and Philosophy of Science*. North-Holland. 3-17.
- [21] Davidson, Donald (1984). “First Person Authority”. *Dialectica* 38 (2-3): 101-112.

- [22] Davidson, Donald (1986). “A Nice Derangement of Epitaphs”. In Ernest Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Blackwell. 433-446.
- [23] Dretske, Fred (1981). *Knowledge and the Flow of Information*. MIT Press.
- [24] Dretske, Fred (1990). “Does Meaning Matter?” In Enrique Villanueva (ed.), *Information, Semantics, and Epistemology*. Blackwell.
- [25] Dummett, Michael A. E. (1993a). “Introduction”. In *The Seas of Language*. Oxford University Press.
- [26] Dummett, Michael A. E. (1993b). “What is a Theory of Meaning?” In *The Seas of Language*. Oxford University Press.
- [27] Dummett, Michael A. E. (1993c). “What Do I Know When I Know a Language?” In *The Seas of Language*. Oxford University Press.
- [28] Dummett, Michael A. E. (1993d). “Language and Communication”. In *The Seas of Language*. Oxford University Press.
- [29] Field, H. (1972). “Tarski’s Theory of Truth”, *Journal of Philosophy* 69: 347–375.
- [30] von Fintel, K., and I. Heim: 2011, “Lecture Notes on Intensional Semantics”, ms., Massachusetts Institute of Technology.
- [31] Fodor, Jerry A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press.
- [32] Fodor, Jerry A. (2001). “Language, Thought and Compositionality”. *Mind and Language* 16 (1): 1-15.

- [33] Fodor, J. A. & Lepore, E. (1993). “Why Meaning (Probably) Isn’t Conceptual Role”. *Philosophical Issues* 3 (4): 15-35.
- [34] Fodor, Jerry A. & Lepore, Ernest (1996). “The Red Herring and the Pet Fish: Why concepts Still Can’t Be Prototypes”. *Cognition* 58 (2): 253-70.
- [35] Fodor, Jerry & Lepore, Ernie (1998). “The Emptiness of the Lexicon: Critical Reflections on J. Pustejovsky’s the Generative Lexicon”. *Linguistic Inquiry* 29: 269-288.
- [36] Fodor, Jerry & and Lepore, Ernie (2001a). “Why Compositionality Won’t Go Away: Reflections on Horwich’s ‘Deflationary’ Theory”. *Ratio* 14 (4): 350-368.
- [37] Fodor, Jerry & and Lepore, Ernie (2001b). “Brandom’s Burdens: Compositionality and Inferentialism”. *Philosophy and Phenomenological Research* 63 (2): 465-481.
- [38] Fodor, Jerry A. & Lepore, Ernest (2002). *The Compositionality Papers*. Oxford University Press.
- [39] Foster, John (2010). “Meaning and Truth Theory”. In Darragh Byrne & Max Kölbel (eds.), *Arguing About Language*. Routledge.
- [40] George, Alexander (1989). “How Not to Become Confused About Linguistics”. In A. George (ed.), *Reflections on Chomsky*. Blackwell. 90-110.
- [41] Glanzberg, Michael (2009). “Semantics and Truth Relative to a World”, *Synthese* 166 (1): 281-307.
- [42] Glanzberg, Michael (2011). “Meaning, Concepts, and the Lexicon”, *Croatian Journal of Philosophy* 11 (1): 1-29.

- [43] Glanzberg, Michael (2014). “Explanation and Partiality in Semantic Theory”. In A. Burgess and B. Sherman (eds.), *Metasemantics*, pp. 259–292. Oxford: Oxford University Press.
- [44] Gross, Steven (2010). “Knowledge of Meaning, Conscious and Unconscious”. *The Baltic International Yearbook of Cognition, Logic and Communication*.
- [45] Guasti, Maria T. (2002). *Language Acquisition: The Growth of Grammar*. MIT Press.
- [46] Gupta, Anil (2012). *Truth, Meaning, Experience*. OUP USA.
- [47] Heck, Richard (2007). “Meaning and Truth-conditions”. In D. Greimann & G. Siegart (eds.), *Truth and Speech Acts: Studies in the Philosophy of Language*. Routledge. 349-76.
- [48] Heck, Richard (2013). “Is Compositionality a Trivial Principle?” *Frontiers of Philosophy in China* 8 (1): 140-55.
- [49] Heim, Irene & Kratzer, Angelika (1998). *Semantics in Generative Grammar*. Blackwell.
- [50] Higginbotham, James (1989). “Knowledge of Reference”. In A. George (ed.), *Reflections on Chomsky*, 153–174. Oxford: Basil Blackwell.
- [51] Horwich, Paul (1998). *Meaning*. Oxford University Press.
- [52] Horwich, Paul (2003). “Meaning and its Place in the Language Faculty”. In Louise M. Antony (ed.), *Chomsky and His Critics*. Malden Ma: Blackwell. 162-178.
- [53] Horwich, Paul (2005). *Reflections on Meaning*. Oxford University Press, Clarendon Press.

- [54] Jackendoff, Ray S. (1990). *Semantic Structures*. Cambridge: MIT Press.
- [55] Jackendoff, R. (2010). *Meaning and the Lexicon: The Parallel Architecture 1975-2010*. Oxford University Press.
- [56] Jackendoff, R. (2011). “Conceptual Semantics”, in Claudia Maienborn, Klaus von Heusinger, and Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, Vol. 1, 688-709. DeGruyter Mouton.
- [57] Johnson, Kent (2007). “The Legacy of Methodological Dualism”. *Mind and Language* 22 (4): 366–401.
- [58] Kaplan, David (1989). “Demonstratives”. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes From Kaplan*. Oxford University Press. 481-563.
- [59] Kennedy, Christopher (2007). “Vagueness and Grammar: The Semantics of Relative and Absolute Gradable Adjectives”, *Linguistics and Philosophy* 30 (1): 1-45.
- [60] Kennedy, C., & McNally, L. (1999). “From Event Structure To Scale Structure: Degree Modification In Deverbal Adjectives”, In T. Matthews, & D. Strolovitch (Eds.), *Semantics and Linguistic Theory* 9: 163–180. Ithaca, NY.
- [61] Kennedy, C., and McNally, L. (2005). “Scale Structure, Degree Modification, and the Semantics of Gradable Predicates”, *Language* 81: 345-381.
- [62] Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Blackwell.
- [63] Kim, Jaegwon (1984). “Concepts of Supervenience”. Reprinted in *Supervenience and Mind*, Cambridge University Press, Cambridge, 1993, 53-78.
- [64] Kólbel, Max (2001). “Two Dogmas of Davidsonian Semantics”. *Journal of Philosophy*, 98 (12): 613-635.

- [65] Kripke, Saul A. (1980). *Naming and Necessity*. Harvard University Press.
- [66] Larson, Richard & Segal, Gabriel (1995). *Knowledge of Meaning: An Introduction to Semantic Theory*. MIT Press.
- [67] Larson, Richard K. & Ludlow, Peter (1993). "Interpreted Logical Forms", *Synthese* 95 (2): 305-355.
- [68] Lewis, David (1975). "Languages and Language". In Keith Gunderson (ed.), *Minnesota Studies in the Philosophy of Science*. University of Minnesota Press. 3-35.
- [69] Lewis, David (1980). "Index, Context, and Content". In Stig Kanger & Sven Öhman (eds.), *Philosophy and Grammar*. Reidel. 79-100.
- [70] Ludlow, Peter (2011). *The Philosophy of Generative Linguistics*. Oxford University Press.
- [71] Marr, David (1982). *Vision*. New York, W.H. Freeman.
- [72] Matthewson, Lisa (2004). "On the Methodology of Semantic Fieldwork", *International journal of American linguistics* 70 (4): 369-415.
- [73] McDowell, John (1981). "Anti-realism and the Epistemology of Understanding". In Herman Parret & Jacques Bouveresse (eds.), *Meaning and Understanding*. W. De Gruyter. 225-248.
- [74] Millikan, Ruth G. (1984). *Language, Thought and Other Biological Categories*. MIT Press.
- [75] Millikan, Ruth G. (forthcoming). *Beyond Concepts: Unicepts, Language and Natural Information*. Oxford University Press.

- [76] Napoletano, Toby (2015). “Compositionality as weak supervenience”. *Synthese* 192 (1): 201-220.
- [77] Napoletano, Toby (2016). “Why Truth-Conditional Semantics in Generative Linguistics is Still the Better Bet”. *Erkenntnis*: 1-20.
- [78] Ninan, Dilip (2010). “Semantics and the Objects of Assertion”. *Linguistics and Philosophy* 33 (5): 355-380.
- [79] Pagin, Peter (2003). “Communication and Strong Compositionality”. *Journal of Philosophical Logic* 32 (3): 287-322.
- [80] Pagin, Pagin & Westerståhl, Dag (2010a). “Compositionality I: Definitions and Variants”. *Philosophy Compass* 5 (3): 250-264.
- [81] Pagin, Pagin & Westerståhl, Dag (2010b). “Compositionality II: Arguments and Problems”. *Philosophy Compass* 5 (3): 265-282.
- [82] Partee, Barbara H. (1979). “Semantics—Mathematics Or Psychology?” In Rainer Bäuerle, Urs Egli & Arnim von Stechow (eds.), *Semantics From Different Points of View*, 1-14. Springer-Verlag.
- [83] Partee, Barbara H. (2011). “Formal Semantics: Origins, Issues, Early Impact”, *The Baltic International Yearbook of Cognition, Logic and Communication* 6, 1-52.
- [84] Patterson, Doug (2005). “Acquirability and Compositionality”. *Mind and Language* 20 (3): 326-352.
- [85] Pelletier, Jeff (1994). “The Principle of Semantic Compositionality”. *Topoi* 13: 11-24.

- [86] Pietroski, Paul (2003). “The Character of Natural Language Semantics”. In Alex Barber (ed.), *Epistemology of Language*. Oxford University Press, 217-256.
- [87] Pietroski, Paul M. (2006). “Character Before Content”. In Judith Jarvis Thomson (ed.), *Content and Modality: Themes From the Philosophy of Robert Stalnaker*. Oxford: Clarendon Press. 34-60.
- [88] Pietroski, Paul M. (2010). “Concepts, Meanings and Truth: First Nature, Second Nature and Hard Work”, *Mind and Language* 25 (3): 247-278.
- [89] Pustejovsky, James (1995). *The Generative Lexicon*. MIT Press. Cambridge, Mass.
- [90] Pustejovsky, James (1998). “Generativity and Explanation in Semantics: A Reply to Fodor and Lepore”. *Linguistic Inquiry*, 29: 2. 289–311.
- [91] Putnam, Hilary (1975). “The Meaning of ‘Meaning’”, *Minnesota Studies in the Philosophy of Science* 7: 131-193.
- [92] Putnam, Hilary (1987). *Representation and Reality*. MIT Press.
- [93] Rabern, Brian (2012). “Against the Identification of Assertoric Content With Compositional Value”, *Synthese* 189 (1): 75-96.
- [94] Recanati, François (2002). *Literal Meaning*. Cambridge University Press.
- [95] Rizzi, L. (1997). “The Fine Structure of the Left-Periphery”. In L. Haegeman (Ed.), *Elements of Grammar: A Handbook of Generative Syntax*. Dordrecht: Kluwer. 281-337
- [96] Robbins, Philip (2005). “The Myth of Reverse Compositionality”, *Philosophical Studies*, 125: 251-275.

- [97] Scharp, Kevin (2013). *Replacing Truth*. Oxford University Press UK.
- [98] Schiffer, Stephen R. (2003). *The Things We Mean*. Oxford University Press.
- [99] Schiffer, Stephen (2015). “Meaning and Formal Semantics in Generative Grammar”. *Erkenntnis* 80 (1): 61-87.
- [100] Sellars, Wilfrid (1951). “Language, Rules and Behavior”, *Journal of Symbolic Logic* 16 (3): 209-210.
- [101] Sellars, Wilfrid (1969). Language as Thought and as Communication. *Philosophy and Phenomenological Research* 29 (4): 506-527.
- [102] Soames, Scott (1989). “Semantics and Semantic Competence”, *Philosophical Perspectives* 3: 575-596.
- [103] Soames, Scott (1992). “Truth, Meaning, and Understanding”, *Philosophical Studies*, 65 (1-2): 17-35.
- [104] Soames, Scott (2012). *What is Meaning?*, Princeton University Press.
- [105] Speas, Peggy, and Tenny, Carol (2003). “Configurational Properties of Point of View Roles”. In DiSciullo, A. (ed.), *Asymmetry in Grammar*. Amsterdam: John Benjamins. 315-344.
- [106] Sprouse, J., and Almeda, D. (2012). “Assessing the Reliability of Textbook Data in Syntax: Adger’s Core Syntax”. *Journal of Linguistics* 48: 609-652.
- [107] Stalnaker, Robert (2014). *Context*. Oxford University Press.
- [108] Stich, Stephen (1996). *Deconstructing the Mind*. Oxford University Press.
- [109] Szabó, Zoltán Gendler (2000). “Compositionality as Supervenience”. *Linguistics and Philosophy*, 23: 475-505.

- [110] Szabó, Zoltán Gendler (2012a). “The Case for Compositionality”. In *The Oxford Handbook on Compositionality*, W. Hinzen, et al. eds. Oxford: Oxford University Press. 64-80.
- [111] Szabó, Zoltán Gendler (2012b). “Compositionality”. *Stanford Encyclopedia of Philosophy*.
- [112] Travis, Charles (1985). “On What is Strictly Speaking True”, *Canadian Journal of Philosophy* 15 (2): 187-229.
- [113] Vendler, Zeno (1957). “Verbs and Times”, *Philosophical Review* 66 (2): 143-160.
- [114] Yalcin, Seth (2014). “Semantics and Metasemantics in the Context of Generative Grammar”. In A. Burgess and B. Sherman (eds.) *Metasemantics*. Oxford University Press, 17-54.