Preface

I want this book, the one you're holding now,[1] to introduce you to a way of thinking about language that I've found very interesting and helpful. The idea is that we use grammar strategically to signal our intended meanings. By *strategically*, I mean that my choices as a speaker are conditioned by the choices you as a hearer will make in interpreting what I say. In short, I've found game theory   the theory of decision making when the outcome of the decision depends on the choices of others   to be enormously helpful in thinking about a wide variety of linguistic puzzles. Let me try to give you an idea of what I mean by this.

If you take a simple word like *and*, you'll find that it's capable of some quite complex behavior. Certainly, there is the familiar *and* of temporal sequencing: the sentence

(1) They got married and had a baby.

is decidedly different from

(2) They had a baby and got married.

The logician might shudder and point to the following pair of sentences, which are surely equivalent:

(3) a. The House of Representatives has 435 members and the Senate has 100 members.
     b. The Senate has 100 members and the House of Representatives has 435 members.

Both sentences in (3) amount to pretty much the same thing. While order matters in (1) and (2), it seems not to matter in (3), where all that is required is that the sentences on either side of the *and* be true. What about

(4) I added nitrate to the damned thing and it blew up!

Surely, more than temporal sequencing is going on in (4); we might infer that the reason it blew up was that I added nitrate to the damned thing.

Well, fine, we might say, we just need to define three kinds of *and*: one kind for temporal sequencing, another for causation, and a third as the logician's conjunction. Perhaps a little thought will reveal still more kinds of *and*.

Surely, though, we're missing something important by supposing that there are three different kinds of *and*. The treatment is compatible with the idea that there are three different words   one for temporal relations, one for logical relations, and one for causation   which just happen to sound alike, a peculiar accident of English.

We ought to entertain the idea that the three different "meanings" of *and* flow from different uses of the same semantic thing. Would a language have a different word for each of the three uses? Perhaps, but certainly most languages use a single word to serve each of the three different purposes. Something more than mere accident is going on.

The great philosopher of language H. Paul Grice thought that there was more here than mere coincidence. He argued that the regimented semantics of the logician didn't quite capture things. Rather, a different kind of logic was needed, a "natural" logic, that could never be replaced by the logician's regimentation:

Moreover, while it is no doubt true that the formal devices are especially amenable to systematic treatment by the logician, it remains the case that there are very many inferences and arguments, expressed in natural language and not in terms of these devices, which are nevertheless recognizably valid. So there must be a place for an unsimplified, and so more or less unsystematic, logic of the natural counterparts of these devices; this logic may be aided and guided by the simplified logic of the formal devices but cannot be supplanted by it. Indeed, not only do the two logics differ, but sometimes they come into conflict; rules that hold for a formal device may not hold for its natural counterpart. (Grice 1975, 43)

The idea is a compelling one   a more abstract logic guides our use of language in such a way that meanings emerge. But what kind of logic could it be? As Grice observes, it certainly isn't the formal logic we might learn in a philosophy or math class. It would have to be something prior, something we all share.

We might suppose that Grice's natural logic is really just the rational use of grammar to signal meaning. On this view, given a context, we use the grammar as a tool to signal meaning; the choices arise from our knowledge of the context, our knowledge of grammar, and our communi-

cative intention. On this view, grammar is a tool that we deploy to get things done.[2] Underlying our use of grammar is a logic of rational decision making. In order to get things done, we must make communicative decisions based on the (potential) decisions of our interlocutors.

Game theory explicitly concerns itself with rational decision making when the outcome of the decision depends on choices made by other (rational) agents. It therefore provides a mathematics that allows us to develop theories of the kind of decision making crucial to understanding linguistic behavior.

Linguistics, particularly in North America, has been dominated by two trends that, while initially helpful, have hardened into dogmas. The first trend is solipsism. The proper subject for linguistic theory, according to this line of thought, is grammatical representation, largely divorced from the content of these representations. The focus on grammatical representations led to an explosive growth of linguistic theory during the second half of the twentieth century. Certainly, the data uncovered and classified by this revolution have been crucial to our understanding of linguistic forms and language diversity.

The second trend dominating linguistics has been the conflation of linguistic explanation and grammatical explanation. As we have learned more about language meaning, for example, the tendency has been to make the semantic component   and, more often than not, the syntactic component   of the grammar more complex. Thus, nodes corresponding to pragmatic functions have been added to syntactic trees, and the semantics itself has been rendered more complex in the service of the goal of explaining aspects of language that might better be accounted for in terms of the use of grammar rather than the grammar itself.

Game-theoretic pragmatics runs counter to both trends. Game theory is, by its very nature, antisolipsistic. The solipsism in current linguistics, and in cognitive science more generally, has outlived its usefulness. The only way to properly understand meaning is to grapple with its social nature; language, after all, is the bridge between our private mental worlds and the public world of social interaction. I argue that, in fact, it is the social that gives content to our mental lives.

The idea that use determines meaning is hardly new; its roots lie in the work of Wittgenstein, Austin, Grice, and many others. Happily, game theory gives us a formal language for working out these ideas. The resulting theory of use will allow us to account for many aspects of linguistic meaning, and the grammar itself can be simplified. The resulting theories are nevertheless precise and subject to empirical testing.

One of the pleasant aspects of game theory is that it allows us to unify many aspects of linguistics that seem, at first glance, to be disparate. For example, we can begin to see profound connections between sociolinguistics and the study of meaning. These connections can be followed into neurobiology, as I suggest at various points in the book. The game-theoretic approach to language promises to open connections between behavioral economics, social evolution, and neuroeconomics.

I would go one step further and argue that game theory returns linguistics to the heart of the social sciences. In recent years, game theory has helped pave the way toward a systematic study of the development of conventions, the evolution of altruism, and reciprocal behavior. Language provides a platform to study all these things; the evolution of Gricean implicature is but one instance of the broader evolution of cooperative behavior.

This book is intended as an accessible introduction to game theory and the study of linguistic meaning. I have tried throughout to keep the tone light and to presuppose little specific knowledge; my intention is to make the ideas available to a wide audience. Many of the ideas I touch on are, by their nature, obscure; nevertheless, I believe that discussions about the nature of meaning, meaning as the outcome of strategic reasoning, are vital to a wide audience. I hope that undergraduates, graduate students, and general readers with an interest in language will find something useful here. The time has come for linguists and other cognitive scientists to make these ideas available to a mass audience, lest we become another obscure guild, open only to a few specialists.

In order to make the book as accessible as possible, I have been sparing in footnotes, have left most bibliographic matters for the end of chapters, and have tried to keep the mathematics down to some simple algebra. Formal definitions have been placed in boxes outside the main text so that readers who are not interested in that level of detail can simply pass them by while still reading the main text.

The first part of the book, chapters 1–3, is an extended argument in favor of the social basis of meaning. While a definitive argument is, of course, impossible, I think the weight of evidence strongly supports the social nature of linguistic meaning. I have occasionally wandered into the realm of the memoir, my hope being that I can show why meaning matters so much to me. The issues here, grounded as they are in analytic philosophy, can seem arcane to a nonspecialist; nevertheless, the questions I raise are of general importance. The main arguments in favor of the economic and ecological nature of meaning are in chapter 3. Chapter 9

attempts to work out the nature of the system more precisely. Part I is a prelude to the study of games and meaning; it motivates the use of game theory in the study of language without being about game theory proper.

Part II turns to game theory. Chapter 4 is a brief, informal introduction to game theory with a particular eye toward coordination games and cooperation. I have tried to keep the mathematics as accessible as possible, but a little math is inevitable. Such a brief chapter cannot do the work of a full introduction to game theory, but I hope enough of the ideas are introduced that a general reader, unfamiliar with the theory of games, can benefit from the chapter and comfortably read the rest of the book.

Chapter 5 turns to a particular application of game theory: games as a model of formal logic. The correspondence between first-order logic and the theory of zero-sum games of perfect information is delightful. I give a logic whose "formulas" are English sentences and show how to evaluate them using games. This is only a small corner of the broader relation between games and logic, but I think that general readers will enjoy seeing a small part of this bigger subject. The use of zero-sum games in logic, as opposed to the coordination games in pragmatics, is also useful in understanding how semantics and pragmatics differ from each other.

Part III turns to the development of bounded rationality. Decisions are constantly made under computational bounds; we do not have perfect knowledge, and we must often make leaps of faith. Chapter 6 explores the problem of common knowledge in some detail. Game theory offers a model of common knowledge, since the players are assumed to know the game they are playing. The mutual knowledge that the speaker and hearer must have in order to communicate can be incorporated as part of the game they are playing. We can avoid the puzzle of infinite regress: my model of your knowledge includes your model of my knowledge, which includes my knowledge of your knowledge, and so on. We can assume that the required knowledge is included in the game and is therefore public. Because of inherent bounds on knowledge, the model of the game is always imperfect. We can use these bounds to think about a variety of interesting phenomena, including linguistic accommodation at one end of the spectrum and misunderstandings at the other. In fact, we can use bounds on common knowledge to model some of the pragmatics of definite descriptions.

Chapter 7 turns to games of partial information, a type of game developed by Prashant Parikh. These games are particularly useful in modeling communication, linguistic and otherwise. I use games of partial information to develop a neo-Gricean model of word finding. This model is

sensitive to both the absolute frequencies of lexical items and to the contribution of context. The games are used to model the lexical side of garden path sentences. As a further illustration of games of partial information, I develop a small model of irony that involves selecting a high payoff state or avoiding risk. In the former case, the speaker and hearer know enough about each other to get the implication of irony; in the latter case, the irony is missed and only the literal content of the sentence is taken.

Chapter 8 illustrates the use of games with two examples. First, a model of discourse pronouns is developed. This model forms only a part of a broader theory of discourse anaphora; the latter theory is beyond the scope of the present work. However, a game analysis can be developed for some simple texts; doing so allows us to identify principles that can be generalized to the study of anaphora in general. The chapter continues with a discussion of the analysis of politeness and how we can use politeness to elucidate conversational implicature. Once again, I can only allude to a larger theory that goes well beyond the present scope; nevertheless, the example illustrates the game methodology and suggests some avenues for future research.

The last chapter turns to the problem of lexical content given a context. We will use the important notion of focal point, due ultimately to Thomas Schelling, to develop a system of social coordination of reference. I argue that these focal points are conventionalized, via social practice, into the concepts associated with lexical items. Furthermore, the process of conventionalizing these focal points has an economic and ecological character whose logic can be formalized, understood, and tested empirically. The resulting system gives insight into the difference between homophony, when two unrelated meanings are associated with the same phonological sign, and polysemy, where a single form extends its hegemony over a semantic space. The chapter ends with some thoughts on how to simulate such a system; the resulting approach takes meaning to be an emergent property of social signaling.

I have found game theory to be a useful way of thinking about linguistic meaning and have more than once been charmed by avenues it has opened up to me. I hope readers will be similarly charmed. The resulting social theory of meaning is a useful anodyne to the relentlessly solipsistic world we have come to inhabit. I hope readers will come and join the fun.