Games, Meaning, and Linguistic Signaling

Robin Clark
Department of Linguistics
University of Pennsylvania

May 24, 2016

The relationship between the meaning that arises from the use of language in context and the literal content utterances (meaning divorced from context) has long been a matter of dispute, generating a vast literature. In this paper, I will provide a brief overview of an approach to meaning based on game theory, a branch of mathematics intended to explicate the rational basis of strategic decision-making. I will, where possible, avoid mathematical formalism in favor of the intuitions that inform a game theoretic approach.[1] My goal is to give the reader a taste of the reasoning behind game theory and its application to the study of natural language. The coverage will be relatively modest, but the reader is encouraged to consult some of the growing body of literature on game theory and language (see Parikh (2001, 2010); Franke (2009) for a variety of approaches). We turn to an informal overview of game theory, glancing casually at some standard games in section 1. In section 2, we will look briefly at two applications of game theory to natural language. The first, in section 2.1, takes a historical example of adversarial speech and analyzes how the meaning of a particular statement can be treated as a special kind of signaling game. The second example, in section 2.2 concerns

[1]Binmore (2007) provides an excellent comprehensive introduction to game theory.

social coordination using socially shared salience.

1 Some Elementary Games

A game is taken to be an instance of decision-making where the outcome of a player's decision is contingent on the choices made by another player or players. It is assumed that players have preferences over the outcomes and that each player seeks to optimize his or her outcomes given these preferences and the choices made by the other players. Thus, a game is *strategic* in that each player selects a move based on his or her reasoning about the moves selected by the other players; a game is *rational* since the players seek to maximize the degree to which the outcome of the game satisfies their preferences. We should note here that the notion of rationality here is an idealization, involving unhindered access to computational resources and information. Behavioral game theory (Camerer, 2003; Thaler, 2015) studies strategy behavior in the laboratory and has done the invaluable service of showing how strategic reasoning is bounded.

In support of the analysis of rational behavior, information is the coin of the realm in game theory. Rational players have knowledge certain kinds of information and can assume that other players have equivalent knowledge. It is unsurprising, then, that we can develop a taxonomy of games based on the information available to the players. First, we can divide games into those of *complete information* and those of *incomplete information*. In a game of complete information, the player's know that actions available to each player, the outcomes of the choices and the preferences of the players. Not so a game of incomplete information, where the players might be uncertain about the moves available to the other players, the preferences of the other players, or the game that the other players are playing; we shall return to games of incomplete information below, since they provide a compelling model of signaling behavior.

We can further divide games into those with *perfect* and *imperfect* information. In a complete game of perfect information, games like tic-tac-toe; checkers; chess; and so on, the players know the structure of the game, as above, as well as their opponents choice. Thus, in chess, a player's move is announced on the game board for all to see. It might seem odd to class chess with tic-tac-toe but mathematically the two differ only in that the space of possible tic-tac-toe games in exceedingly small and easy to search, while the space containing only the possible chess moves is vast and intractable to search.

In a complete game of imperfect information, the players have asymmetrical access to certain types of information. In poker, for example, the players all know what a possible hand is, how the cards are ranked, that their opponents are endeavoring to win and so forth; they lack precise information about the contents of their opponents' hands, a situation that makes interesting behaviors like bluffing possible.

To make things more concrete, let's look at a concrete example of a game of imperfect information, the *Stag Hunt*. Imagine two players who plan on going hunting for food. Each player can choose either to *cooperate* with the other player to catch a stag or to *defect* and catch a hare. The outcomes of the actions are as follows:

(i.) If both players choose to cooperate, they get a nice result in the form of a stag, which provides a lot of food.

(ii.) If one player chooses to cooperate and the other chooses to defect, then the former fails to catch anything and goes hungry while the latter gets a small meal of a hare.

(iii.) Naturally, both players can choose to defect, hunt their own hares and have a guaranteed meal.

The next step is to spell out the preferences of the players. We can assume that, all else being equal, both players would like to catch a Stag, which

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$\boxed{2},\boxed{2}$	$0, 1$
	Defect	$1, 0$	$1, 1$

Figure 1: A Stag Hunt

provides the most food. Barring that, each one would gladly take a hare, since it yields a meal. The worst case is not catching anything:

$$\text{Stag} > \text{Hare} > \text{Nothing}$$

By associating the integer 2 to Stag, 1 to Hare, and 0 to nothing, we translate the preferences of the players to numbers while preserving the order of their preferences.

Next, we establish an order of play. In this case, we will assume that the players make their choice of action simultaneously. That is, they choose *Cooperate* or *Defect* without knowing their opponent's choice. This means that, when each player chooses their action, they do not have full knowledge of the state of the world; they are in an *information set* consisting of two states. In one state, their opponent has chosen *Cooperate* while in the other state their opponent has chosen *Defect*. It is in this sense that the game is a complete game of imperfect information. Both players know the structure of the game, the outcomes of the actions, the preferences of the players, but they are, at the crucial moment, ignorant of the other player's choice.

I've shown the *strategic normal form* of the game in Figure 1. The representation is a bimatrix where each cell contains a pair of numbers "a, b", where a is the payoff accorded to the row player (Player 1, in Figure 1) and b is the payoff accorded to the column player (Player 2, in Figure 1). I have, in addition, indicated equilibrium outcomes using squares and circles around the payoffs. An equilibrium is an outcome where no player has incentive to unilaterally change his or her choice. Thus, in the case where both players have chosen *Cooperate*, they've caught a Stag and gotten a payoff of 2 each. If row player, for example, had unilaterally selected *Defect*, she would get a hare with payoff 1. Since Stag is preferred to Hare ($2 > 1$), she has done worse. She has no motive to change her decision, so the pair of actions (*Cooperate*,*Cooperate*) is an equilibrium.

Equally, suppose that both players have chosen *Defect*; then both players get a payoff of 1. Suppose that column player were to unilaterally change his decision, by doing so he no longer gets a Hare and instead gets nothing since, by hypothesis, his opponent has not changed her choice. Since each player prefers Hare to Nothing ($1 > 0$), column player does worse and clearly has no reason to change his choice from *Defect* to *Cooperate* and the pair of actions (*Defect*,*Defect*).

Compare the equilibria (*Cooperate*,*Cooperate*) and (*Defect*,*Defect*), with the non-equilibrium (*Defect*,*Cooperate*). In this case, if the row player changes her action from *Defect* to *Cooperate*; in this case, her payoff improves from 1 to 2, so she has a reason to change her decision unilaterally. Equally, in the same case, if column player unilaterally switches from *Cooperate* to *Defect*, he will improve his payoff from 0 to 1 (nothing to eat to a hare), so he, too, has a reason to unilaterally defect.

The game in Figure 1, then, has two equilibria. One, (*Cooperate*,*Cooperate*) has the payoff (2,2), while other, (*Defect*,*Defect*), has the payoff (1,1). It happens that the payoffs in the former case are greater than the payoffs in the latter case. We term the former equilibrium *Pareto-dominant* since at least one player gets a higher payoff than the other equilibrium and no player does worse (in fact, both players do better in this case). The other equilibrium, (*Defect*,*Defect*), has an interesting property, however. Although neither player does as well as they could by selecting the Pareto-dominant outcome, both players risk less by choosing *Defect* since each player is guaranteed a payoff of 1 no matter what move their opponent decided to make. Notice that choosing the move *Cooperate* requires a certain level of faith that the other player will also choose *Cooperate*; if not, then the poor fellow who chooses *Cooperate* gets nothing. In other words, the equilibrium (*Defect*,*Defect*) is *risk dominant*; it minimizes the risk that the player must tolerate. Which should we choose? The answer is clearly contingent on our assessment of what the other player will do.

The charm of game theory is that it can convert preferences to quantities and, by so doing, allows us a numeric method for grounding our decisions. We can use the notion of *expected utility* of the choices in the game. The expected utility of a choice is the average payoff I can expect if I play that choice. Suppose I want to know the expected utility of playing *Cooperate* in the Stag Hunt. Let's suppose that p is the probability that my opponent plays *Cooperate* so that $(1-p)$ is the probability that my opponent plays defect. Multiplying the payoffs by the probabilities and adding the results we get:

$$(p \times 2) + ((1-p) \times 0) = 2p$$

as the expected utility for playing *Cooperate*. Likewise, my expected utility for playing *Defect* is:

$$(p \times 1) + ((1-p) \times 1) = 1$$

A player who defects can expect a payoff of 1 no matter what, as we noted above.

Suppose that we find the value of p, the probability that my opponent plays *Cooperate*, by setting my expected utility for playing *Cooperate* to be the same as my expected utility for playing *Defect*. That is, the value for p where I am indifferent to the choice between *Cooperate* and *Defect*:

$$2p = 1$$

Solving for p gives us 0.5. This is the *Nash equilibrium* of the game and, in the present case, it corresponds to an indifference point. If I believe that the probability that my opponent chooses *Cooperate* is exactly 0.5, then I have no basis for preferring *Cooperate* over *Defect*, or vice versa; I can just select an action at random. The analysis is shown graphically in Figure 2, where the expected utilities of playing *Cooperate* and *Defect* are shown relative to the probability that my opponent plays *Cooperate*; the point where the lines intersect is the Nash equilibrium, the point where I am indifferent to playing either strategy since their expected payoffs are the same.

I might size up my opponent and decide that the probability of her choosing *Cooperate* is greater than 0.5, based on prior knowledge of her behavior, for example. If my estimate for p is greater than the indifference point, then this suggests that I would be well-advised to choose *Cooperate* since, on average, I would stand to get a better outcome. Choosing to cooperate is suggestive of trust in my opponent which should be based on my assessment of her and my belief in her assessment of me.

On the other hand, I might estimate that the probability that my opponent will cooperate is less than my indifference point of 0.5; I judge my opponent to be of a suspicious caste of mind and averse to risk. I would, in that case, have a reason to choose *Defect* since I would thereby minimize my risk and guarantee myself a payoff of 1. I have converted the payoffs specified in the game to probabilities and, then, to a plan for making my decision.

Let us now consider another game, the well-known *Prisoner's Dilemma* (PD). I've shown the strategic normal form of the game in Figure 3. The PD is an important game in the history of game theory and has played an interesting role in studying the evolution of cooperative behavior and animal

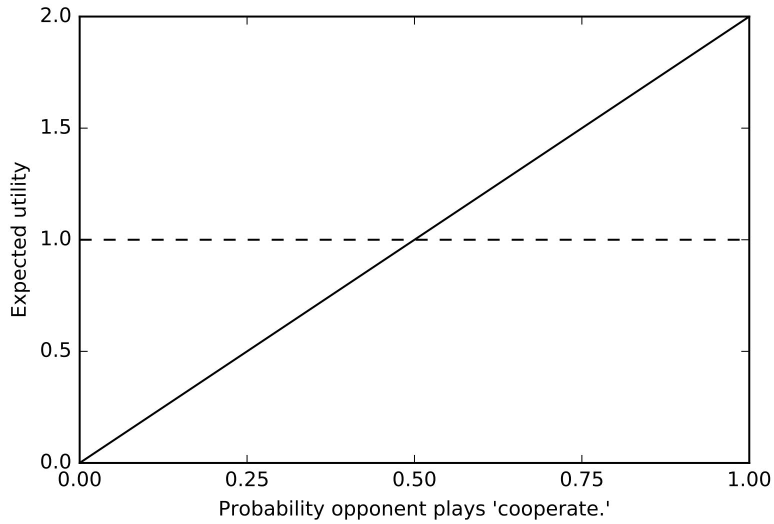

Figure 2: Expected utilities of *Cooperate* (solid line) and *Defect* (dotted line) with respect to the probability that the opponent plays *Cooperate*.

signaling systems (McElreath and Boyd, 2007; Maynard Smith and Harper, 2003; Searcy and Nowicki, 2005).

Suppose that the players of the game can, as in the Stag Hunt, either *Cooperate* or *Defect*, and, as above, the players make their choices simultaneously. In cooperating, the player confers a benefit (b) on the other player, worth 2 points, but does so at a cost (c): 1 point is deducted from the cooperator's payoff. If both player's cooperate, then they both get $b - c = 1$. Suppose one player cooperates while the other player defects. The defector gets a benefit, $b = 2$, at no cost while the cooperator suffers a cost, $c = -1$, with no benefit. Finally, if both players defect, then no benefit is transferred and no cost is suffered, so both players get a payoff of 0.

Unlike the Stag Hunt, the PD has a unique equilibrium. To see this,

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	1,1	$-1, 2$
	Defect	$2, -1$	0,0

Figure 3: A Prisoner's Dilemma

notice that the action *Defect* always has a better payoff than *Cooperate*. If both I and my opponent play *Cooperate*, for a payoff of (1,1), then I get a better payoff by choosing *Defect*, since will get a payoff of 2; my opponent, of course, gets the worst possible payoff of -1. If I play *Defect* and my opponent opponent also plays *Defect*, then our joint payoff is 0, which is still better than the payoff of -1 I would get if I unilaterally switched to *Cooperate*. Thus, *Defect* dominates *Cooperate*. We are left with the uncomfortable conclusion that rational play excludes joint cooperation, with its mildly profitable joint payoff of 1, in favor of joint defection, with its less satisfying joint payoff of 0. This, of course, raises the interesting question of how cooperation could evolve if it involves costs; this would take us well beyond the scope of this paper, though the interested reader should consult McElreath and Boyd (2007) and the references cited there.

Let's suppose, though, that the players in the PD are permitted to signal one another, so that the game is divided into stages: A signaling stage followed by play as defined above. Suppose further that my opponent sends me the message: **I plan to cooperate**. Should I listen? It may be that my opponent is sincere, but it is at least as likely that she is trying to persuade me to play *Cooperate* so she can play *Defect*, collect a pure benefit without any cost, and leave me with pure cost. Signaling in a pure PD, where the game is played as a one-off, is *cheap talk*: I should ignore the signal. But if I'm well-advised to ignore the signal, then my opponent shouldn't bother to send the signal; signaling itself could not evolve. This puzzle has stimulating a great deal of research on repeated games and costly signaling.

With the background on the Stag Hunt and the PD in mind, let's now suppose that we live in a world divided into players of the Stag Hunt—we'll refer to them as type 1 or t_1 players—and players of the PD, type 2 (t_2) players. As I move through the world, I'll encounter the two types—t_1 with probability p and t_2 with probability $1 - p$. In each possible game, I have imperfect information about my opponent's intended move, but we have also moved on to an apparently new type of uncertainty since I can no longer be certain about which game is being played. This is reflected in my uncertainty about the type of my opponent which, of course, translates into an uncertainty about which game is being played. This is now a game of incomplete information. Games of incomplete information were studied by John Harsanyi (1967-68) who gave them a solid mathematical characterization and showed that they could, in principle, be solved using the same techniques that we used above for games of imperfect information.

We can visualize a game of incomplete information as follows. There is a first move by Nature (a random move) that sets the type of your opponent; while your opponent knows her type, this remains unknown to you. With some probability, p, you will play a Stag Hunt; otherwise, you will play a PD with probability $1 - p$. In short, I am uncertain as to which game I am playing. My opponent's moves, in themselves, may not eliminate my uncertainty. The present example is rather too small to illustrate the problem in the most compelling way, but we can still get a flavor for it. Suppose that my opponent announces her move, thereby committing herself to the choice. Her announcement of "Cooperate!" does not definitively settle which game we are playing. If we are playing a PD, I could take advantage of her by defecting and get 2 points, but if we are playing a Stag Hunt, choosing defect will reduce my payoff. I won't know until I make my choice and the payoffs are distributed.

I can, though, try to reduce my uncertainty in a variety of ways: I might try to assess my opponent by observing her prior moves, perhaps in other plays of the game if I have access to such information. I might rely on personal information: I know her and have an opinion as to what kind of player she is. She might, in fact, be related to me so that I am willing to confer a benefit on her altruistically.[2] If the game is sufficiently complex with multiple moves (as in a card game like poler), I might observe her behavior and try to assess what kind of player she is. Collecting such information would reduce my uncertainty, allowing me to navigate the game and conditioning my choice of moves.

2 Two Examples

2.1 Signaling as a Game of Incomplete Information

We have, in the previous section, rehearsed some elements of game theory at an informal level. At this point, it is helpful to turn to a more concrete application to language. I have selected for this purpose an example drawn from American history, the Army-McCarthy hearings. A major issue at the hearings was whether Senator Joseph McCarthy or his aid, Roy Cohn, had used their influence to pressure the Army to give preferential treatment to an Army private, G. David Schine. A photograph purporting to show Schine with Secretary of the Army, Robert T. Stevens, was introduced as evidence, the implication being that Stevens could have used his influence

[2]This is Hamilton's rule, see the papers collected in Hamilton (1996) as well as the exposition in McElreath and Boyd (2007).

in favor of Schine. Special Counsel Joseph Welch suspected that Cohn, or someone connected to Cohn, had doctored the image by editing out other people in order to make it appear as though Stevens and Schine had a special relationship. In order to provide sufficient context, I will quote Thomas (1973, pp. 532-533) who describes the exchange as follows:[3]

> One such moment occurred during the argument over the cropped photograph. Jim Juliana was disclaiming any knowledge of the source of the original print. With a raffish leer Welch pried, "Did you think this came from a pixie?" The spectators laughed, priming Welch to continue. "Where did you think this picture I hold in my hand came from?" Quick to run interference for his staff, McCarthy moved in with, "Will that question be reread?" It was so ordered by the chair, and in an expressionless tone the reporter read, "Did you think this came from a pixie?"
>
> Facing Welch, McCarthy asked: "Would counsel, for my benefit, define—I think he might be an expert on this—the word 'pixie'?"
>
> Welch twinkled and replied roguishly: "I should say, Mr. Senator, that **a pixie is a close relative of a fairy**." The room rocked with mocking laughter, and when it died down Welch went on, with deadly innuendo: "Shall I proceed, sir? Have I enlightened you?"
>
> The intent was unmistakable, and this time McCarthy joined in the laugh, although his smile was strained. Attempting a light retort, he hit back: "As I said, as I said, I think you might be an authority on what a pixie is."

We will focus on the part of the exchange in boldface, with an eye to working out as best we can what, exactly, Mr. Welch could have been doing by saying

[3] The full exchange can be seen in the documentary *Point of Order* (de Antonio and Talbot, 1964)

such a thing. The passage of time and changes in norms have rendered some of the meaning of the exchange rather obscure, but the passage reflects the sense in which language is a cooperative activity (Grice, 1975), some of the ways that language can be used in an adversarial situation where the participants are not cooperating fully,and how speaker meaning and literal meaning can diverge (Kripke, 1977).

The nostrum that language is a cooperative activity is often offered up in discussions of adversarial language; if nothing else, this approach to conflict in language is rather incurious, to say the least. It is not unusual to hear analogies with sports: Opposing players in tennis find their interests opposed but, in determining the winner, they must cooperate at least as far as the rules of tennis go. Failing to observe the rules of tennis would be in neither player's interest since the resulting exchange would deviate from the standards a tennis match and the "victor" would hardly count as having won a tennis game. We can get further if we consider adversarial interactions as strategic engagements in which each participant intends to advance his interests as well as alter his opponent's behavior, often by changing the strategic landscape that the exchange takes place in. Linguistic interchanges are often far more than the simple exchange of information; they involve each speaker attempting to shape the other's behavior and attitudes. It may involve strategically revealing information from some participants in the exchange while concealing the same information from others. It is safe to assume, in the present example, that words like "pixie" and "fairy" are not being used with their conventional meaning.

McCarthy and Welch, the major players in the above exchange, were in a deeply adversarial relationship. McCarthy and Cohn had been investigating the Army, supposedly to expose the presence of communists and potential security risks in the Army, when the Army had pushed back with their own charges of corruption against McCarthy. Mr. Juliana was being questioned about the origin of a key piece of evidence, the origin of the cropped photograph. While Mr. Welch had every reason to expose the originator of the photograph, it is safe to suppose that Sen. McCarthy had a compelling interest in keeping the identity secret, and had so instructed Mr. Juliana. The interests of the key players are, then, diametrically opposed. Both sides realized that they were engaged in a struggle in a very public forum: The Army-McCarthy hearings were the first nationally televised congressional hearings and received a great deal of attention. As such, both sides had much to gain and much to lose.

The question at immediate issue was the source of the photo of Schine and Stevens that had been provided to the committee; we can frame this as a decision problem and take the solution of it to be the question that drives the utilities in the exchange (van Rooy, 2003); that is, a statement that answers the underlying question of the identity of the source of the photograph would derive a great deal of utility from this fact.

In order to answer the question of the source of the photograph, one other piece of information is necessary; Roy Cohn, McCarthy's aide, was a closeted gay man.[4] This piece of information clarifies Welch's quip. To see how, let's begin with the word *fairy*. The word could, of course, mean a magical creature; let's use *sprite* as a near-synonym which could be used as a paraphrase. In 1950s America, however, the word could also be used as a slur meaning *(male) homosexual*. The ambiguity in the word invites us to play a game of incomplete information. This game of incomplete information can be analyzed as consisting in two possible subgames; for the moment, let's treat these as separate games.

In one game, the speaker uses "fairy" with the intention of denoting MAGICAL BEING.[5] The speaker faces the strategic choice of saying either

[4]Note that McCarthy and Cohn were investigating the Army for infiltration by communists and for security risks, particularly homosexuals. Cohn never publicly acknowledged his sexuality, even up to his death from complications due to AIDS. See Miller (2006).

[5]I'll use bold small-caps to indicate the meaning of expressions; we needn't concern ourselves with details of the semantic notation at this level of discussion.

"fairy" or using a paraphrase, say, "sprite." If the speaker uses "fairy", then the hearer is faced with a strategic choice: What did the speaker mean? She might suppose that the speaker did indeed mean MAGICAL BEING by saying "fairy" but there is a chance that the speaker intended to utter a slur against homosexuals. If the speaker intends the conventional meaning of "fairy" but fears that the hearer will seize on the unintended slur, he might rephrase and avoid saying "fairy" although the paraphrase might not have quite the flavor he wishes to evoke. Thus, we can assume that if the speaker uses "fairy" to mean MAGICAL BEING and is so understood

$$(\text{fairy}, \textsc{magical being}) > (\text{sprite}, \textsc{magical being}) > (\text{fairy}, \textsc{homosexual})$$

The speaker prefers "fairy" since it evokes exactly what he intends, but will settle on "sprite" if need be; the worst outcome for both speaker and hearer is that the hearer misreads the speaker's intentions and concludes he meant HOMOSEXUAL.

The other game, where the speaker intends to use "fairy" as a slur, is topographically similar to the preceding game. The speaker must decide whether the hearer will grasp his intention to produce a slur and, if not, must select some other expression that will more nearly approximate his intention. As before, upon hearing "fairy" the hearer must decide what the speaker intended by using this word. In this game, we have the following preferences:

$$(\text{fairy}, \textsc{homosexual}) > (\text{homosexual}, \textsc{homosexual}) > (\text{fairy}, \textsc{magical being})$$

The speaker would prefer the hearer to arrive at the slur; if he is uncertain whether he can accomplish that with "fairy", he will move on to some other expression. The speaker does not want the hearer to conclude that the speaker intended MAGICAL BEING.

It remains to assemble the two games into a game of incomplete information. The two subgames are joined by a move by Nature (that is, a random move) that makes the speaker one of two types: Intending to denote a magical being or intended to make a slur; notice that the speaker's signal of "fairy" does not, in itself, determine for the hearer which game is being played. Only the speaker knows the type assigned to him by Nature; the hearer remains uncertain about the speaker's type and must use other cues to work out the intended meaning.

Returning to our example, "a pixie is a close relative of a fairy" taken out of context provides very little information about what interpretation of "fairy" the speaker intends; the presence of "pixie" in the utterance would, all else being equal, bias a hearer toward the MAGICAL BEING interpretation, since "pixie" puts us in the realm of the fantastic. But now consider that the question under discussion is the source of the cropped photograph of Schine and Stevens. Anyone with the background information that Roy Cohn was a closeted homosexual, then, by taking "fairy" to be a slur can reason from "a fairy is a close relative of a pixie" to someone associated with a homosexual to Roy Cohn, thus answering the question and providing extra utility to the interpretation of "fairy" as a slur. The anodyne quip about pixies transforms itself to a veiled accusation with an implicit threat that some information that was then extremely dangerous could be revealed. Welch, of course, is using exactly the kind of smear to take down McCarthy that McCarthy himself had often used to devastating effect. The fact that only some members of the audience had the background knowledge to appreciate the full import of Welch's quip makes it an example if irony in the sense of Fowler (1965, p. 305):

> Irony is a form of utterance that postulates a double audience, consisting of one party that hearing shall hear and shall not understand, and another party that, when more is meant than meets the ear, is aware both of that more and of the outsiders' incomprehension.

We could continue the analysis of the interchange in question, looking at

Figure 4: A table with something on it.

the question of speech act recognition, but the underlying point that signaling, even in adversarial circumstances, involves a recognition that the interlocutors are seeking to fulfill their preferences the best they can given the interests and preferences of the other players. Games of incomplete information provide an elegant and explicit framework for investigating how strategic reasoning contributes to this process.

2.2 Reference and Focal Points

Suppose that you are in a room and I ask you, from another room, to bring me the thing on the table. You look at the table and you see the situation shown in Figure 4 (simple version of this game is considered in Clark (2012b)). When asked what they would do in this situation, most people unhesitatingly say that they would bring the book shown that is on the table. Notice that something is, in fact, going on here; there are two things on the table—the book and the tablecloth—so my use of the definite description should be false since there is more than one thing on the table, at least according to the standard theory of definite descriptions (Russell, 1905).

Take the same request and substitute the situation shown in Figure 5 for the one we considered in Figure 4. In this case, almost everyone changes their decision about what they would bring, abandoning the book in favor of the object to the right of the book (which is, in fact, a child's toy). The request itself is not rejected as anomalous; rather, a new referent is found for the definite description "the thing on the table." This happens even though there are now three things on the table and despite the fact that the vast majority had, a moment before, taken the book to be the referent for the definite description. When pressed, most people will say that the toy is a better candidate for "thing" than the book, or that in the situation shown in Figure 5 a competent speaker would use "book" if they meant the book, neither explanation really accounting for the set of facts.

Both of the above cases involve coordination between the speaker and the hearer; the speaker seeks to get the hearer to coordinate with her on the object she intended to single out by using an expression like "the thing on the table." In the first case, coordination is relatively simple given the context and the assumption that the tablecloth is not *jointly salient*; it is sufficiently part of the background that neither participant is likely to select it as a referent for "the thing on the table."

In the second case, the situation is more complex; in this case there are

Figure 5: A table with some more on it.

two objects that are sufficiently salient to be the referent for "the thing on the table." Figure 6 shows a simplified version of the game, where the row player is taken as the speaker and the column player is the hearer. We suppose that the speaker has an intention to communicate either BOOK or TOY and utters "thing" with that intention. Both the speaker and the hearer prefer to coordinate on the speaker's intention. Thus, if the speaker intends BOOK, they both want the hearer to choose BOOK and similarly for TOY. Thus, there are two equilibria as indicated by the boxes around the payoffs. Since both equilibria have the same payoff, neither player should have a preference

		Player 2	
		BOOK	TOY
Player 1	⟨"thing", BOOK⟩	$\boxed{1}, \boxed{1}$	$0, 0$
	⟨"thing", TOY⟩	$0, 0$	$\boxed{1}, \boxed{1}$

Figure 6: A Referential Coordination Game

		Player 2	
		BOOK	TOY
Player 1	⟨"thing", BOOK⟩	$\boxed{1}, \boxed{1}$	$0, 0$
	⟨"thing", TOY⟩	$0, 0$	$2, 2$

Figure 7: A Referential Coordination Game with Focal Point

as to which equilibrium is selected.

This analysis fails to capture the preference for taking the exotic-looking toy as the referent for "the thing on the table." We can revise the game to reflect this by giving the toy interpretation greater utility based on the fact that the toy is jointly salient, as in Figure 7; the fact that both participants anticipate that the toy is most salient to the other—and, thus, an appropriate choice in a referential coordination game—adds to the utility associated with this option. The notion of joint salience has come to play an important role in game theoretic analyses of social coordination where jointly salient options are referred to as *focal points*. The first use of focal points in the game theoretic literature is due to Schelling (1960, p. 57):

> People *can* often concert their intentions or expectations with others if each knows that the other is trying to do the same. Most situations—perhaps every situation for people who are practiced at this kind of game—provide some clue for coordinating behavior, some focal point for each person's expectation of what the other expects him to expect to be expected to do. Finding the key. or rather finding *a* key—any key that is mutually recognized as the key becomes *the* key—may depend on imagination more than on logic; it may depend on analogy, precedent, accidental arrangement, symmetry, aesthetic or geometric configuration, casuistic reasoning, and who the parties are and what they know about each other.

Since Schelling first discussed the idea in 1960, focal points have been extensively studied in the behavioral game theory literature (see, in particular, Bacharach (1999, 2006); Bacharach and Bernasconi (1997); Sugden (1995); Sugden and Zamarrón (2006), among many others). The capacity to jointly select a focal choice is deeply social, as illustrated by the study by Mehta et al. (1994), who divided subjects into "picking" and "coordinating" conditions; in the picking condition subjects were asked simply to choose items in response to a questionnaire, while in the coordinating condition subjects filled out the same questionnaire but were told that they would receive a payoff every time their choice matched that of another randomly chosen participant. Randomly pairing the answers in the coordinating condition yielding significantly more matches than randomly pairing the answers in the picking condition, suggesting that (social) joint salience, as suggested by Schelling, above, is a significant factor in coordination tasks.

Applying the intuition behind focal points to the case at hand, the word "thing" in "the thing on the table" contains little specific information and, therefore, begs to be filled in by the most jointly salient object in the environment. In the first case, considered above, the book is the most jointly salient object, so both the speaker and the hearer can reason that "the thing on the table" must be the book. In the second instance, where the book and the toy are both available as referents, joint salience dictates that the exotic toy is the focal point in the environment so the speaker and the hearer are mutually aware of its focality, making it the best referent for "the thing on the table." In this case, the book is no longer focal and the speaker will have to supply more information—for example, "the book on the table"—in order to refer to it.

Similar reasoning can extend the analysis to the discourse interpretation of indefinite noun phrases, definite noun phrases and pronouns more generally. Clark and Parikh (2007); Mayol (2009); Mayol and Clark (2010) all try to bring game theoretic reasoning to bear on discourse pronouns. Although not couched in game-theoretic terms, H. Clark (1992, Part I) has emphasized the importance of mutual information and common ground in the analysis of definite descriptions and pronouns. In general, pronouns are available when the referent is both in the discourse model and highly salient—that is, when it is a focal point. A definite description is available to referents that are in the discourse model but are no longer focal (Clark, 2012a).

To illustrate how the interplay of focality and discourse status might work, consider the following short text:

> A man walked into the room. He was wearing an odd hat. He approached Bill and offered him the hat. Bill was taken aback but politely declined the offer. The man shrugged and left.

The text begins with "a man," which is new to the discourse, but having been introduced in an empty model, it is the focal point and can be the target of the pronoun, "he," in the second sentence, The second sentence also introduces a new discourse object "an odd hat." The man continues to be focal and is the target of “he” in the third sentence, where the hat, being in the discourse model, but non-focal, is the target of a definite description. Bill takes over as topic in the next sentence where “the offer” is discourse old, since it was evoked by the preceding sentence, but is insufficiently focal to support a pronoun like “it.” Finally, in the last sentence we return to the man who is now both discourse old and non-focal.

Linguistic behavior allows us not only to transmit information, but also coordinate our behavior and influence the actions of other agents. The reasoning associated with focal points allows us to explore the social aspects of decision making in a precise and explicit way. The reasoning involved influences many aspects of our linguistic behavior, from reference to the way that we structure narratives and discourse. Focal points, themselves, are profoundly game-theoretic in nature, allowing us to guide our strategic decision-making in a way that if sensitive to social norms and joint salience.

3 Prospects

Game theoretic pragmatics is relies on decision sciences to study linguistic meaning. Having brought game theory to bear on meaning, there is every reason to suppose that its methods can be used to study language use more generally. For example, Ahern (2015); Ahern and Clark (2016) use a model of signaling developed by Crawford and Sobel (1982), to analyze changes in the expression of negation in English and, then, develop a precise mechanistic model of the change based on corpus evidence. Evolutionary game theory (see Weibull (1995) to cite but one of many examples) can be extended to build precise models of language change, language evolution and linguistic typology (Benz et al., 2006, 2010; Jäger, 2007). Although as yet there are few game-theoretic models of sociolinguistic variation, game theory and agent-based models provide a natural setting for studying such variation (see Clark and Kimbrough (2015); Stanford and Kenny (2013) for examples). Our hope is that we will be able to observe linguistic phenomena in the field or in corpora; build mechanistic models based on our observations using the standard tools (evolutionary) game theory, Agent-Based Models and mathematical biology; then test those models in the laboratory or in the field. Game Theory is one tool that can reinvigorate our study of language in the world.

References

Ahern, C. (2015). *Cycles and stability in linguistic signaling.* Ph. D. thesis, University of Pennsylvania.

Ahern, C. and R. Clark (2016). Conflict, cheap talk, and Jespersen's cycle. in submission.

Bacharach, M. (1999). Interactive team reasoning: a contribution to the theory of co-operation. *Research in Economics 53*, 117–147.

Bacharach, M. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory.* Princeton, NJ: Princeton University Press. edited by Natalie Gold and Robert Sugden.

Bacharach, M. and M. Bernasconi (1997). The variable frame theory of focal points: An experimental study. *Games and Economic Behavior 19*, 1–45.

Benz, A., C. Ebert, G. Jaeger, and R. van Rooij (Eds.) (2010). *Language, Games and Evolution.* Springer.

Benz, A., G. Jäger, and R. van Rooy (2006). An introduction to game theory for linguists. In A. Benz, G. Jäger, and R. van Rooy (Eds.), *Game Theory and Pragmatics*, Palgrave studies in pragmatics, language and cognition, pp. 1–82. Basingstoke, UK: Palgrave Macmillan.

Binmore, K. (2007). *Playing for Real: A Text on Game Theory.* Oxford, UK: Oxford University Press.

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Clark, H. H. (1992). *Arenas of Language Use*. Chicago: The University of Chicago Press & Center for the Study of Language and Information.

Clark, R. (2012a). *Meaningful Games: Exploring Language with Game Theory*. Cambridge, MA: The MIT Press.

Clark, R. (2012b). Social and physical coordination. *Interaction Studies 13*(1), 66–79.

Clark, R. and S. O. Kimbrough (2015). The spontaneous emergence of language variation from a homogeneous population the spontaneous emergence of language variation from a homogeneous population. In *https://computationalsocialscience.org/csssa2015-conference-schedule/csssa-2015-papers/*. Computational Social Science Society of the Americas.

Clark, R. and P. Parikh (2007). Game theory and discourse anaphora. *Journal of Language, Logic and Information 16*, 265–282.

Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society 50*(6), 1431–1451.

de Antonio, E. and D. Talbot (1964). *Point of Order*. directed by Emile de Antonio. Point Films.

Fowler, H. W. (1965). *Fowler's Modern English Usage*. Oxford: Oxford University Press.

Franke, M. (2009). *Signal to Act*. Ph. D. thesis, University of Amsterdam.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. P. Morgan (Eds.), *Syntax and Semantics*, Volume 3, pp. 41–58. New York: Academic Press.

Hamilton, W. D. (1996). *Narrow Roads of Gene Land: The Collected Papers of W. D. Hamilton: Evolution of Social Behavior.* Oxford: Oxford University Press.

Harsanyi, J. C. (1967-68). Games with incomplete information played by Bayesian players. *Management Science 14*, 159–182, 320–334, 486–502.

Jäger, G. (2007). Evolutionary game theory and typology: A case study. *Language 83*, 74–109.

Kripke, S. A. (1977). Speaker reference and semantic reference. In P. A. French, T. E. Uehling, Jr., and H. K. Wettstein (Eds.), *Contemporary Perspectives in the Philosophy of Language*, pp. 6–27. Minneapolis: University of Minnesota Press.

Maynard Smith, J. and D. Harper (2003). *Animal Signals.* Oxford Series in Ecology and Evolution. Oxford University Press.

Mayol, L. (2009). *Pronouns in Catalan: Information, Discourse and Strategy.* Ph. D. thesis, University of Pennsylvania, Philadelphia, PA.

Mayol, L. and R. Clark (2010). Overt pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics 42*, 781–799.

McElreath, R. and R. Boyd (2007). *Mathematical Models of Social Evolution.* Chicago, IL: University of Chicago Press.

Mehta, J., C. Starmer, and R. Sugden (1994). The nature of salience: An experimental investigation of pure coordination games. *The American Economic Review 84*(3), 658–673.

Miller, N. (2006). *Out of the Past: Gay and Lesbian History from 1869 to the Present.* New York: Alyson Books.

Parikh, P. (2001). *The Use of Language.* Stanford, CA: CSLI Publications.

Parikh, P. (2010). *Equilibrium Semantics.* Cambridge, MA: The MIT Press.

Russell, B. (1905). On denoting. *Mind 14*, 479–493.

Schelling, T. C. (1960). *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Searcy, W. A. and S. Nowicki (2005). *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems.* Monographs in Behavior and Ecology. Princeton, NJ: Princeton University Press.

Stanford, J. N. and L. A. Kenny (2013). Revisiting transmission and diffusion: an agent-based model of vowel chain shifts across large communities. *Language Variation and Change 25*(02), 119–53.

Sugden, R. (1995). A theory of focal points. *The Economic Journal 105*(430), 533–550.

Sugden, R. and I. E. Zamarrón (2006). Finding the key: the riddle of focal points. *Journal of Economic Psychology 27*(5), 609–621.

Thaler, R. H. (2015). *Misbehaving: The Making of Behavioral Economics.* New York: W.W. Norton & Company.

Thomas, L. (1973). *When Even Angels Wept: The Senator Joseph McCarthy Affair—A Story without a Hero.* New York: William Morrow and Company.

van Rooy, R. (2003). Asserting to resolve decision problems. *Journal of Pragmatics 35*, 1161–1179.

Weibull, J. W. (1995). *Evolutionary Game Theory.* Cambridge, MA: The MIT Press.