How to Play a Language Game

Robin Clark
Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104
rclark@sas.upenn.edu

1 Introduction

2 Aumann's Assurance Game

Aumann (1990) proposes a signaling game of particular interest to natural language pragmatics; it is a variant of a "stag hunt" game. In a stag hunt game, there is a conflict between cooperation—which yields a relatively high payoff for both participants—and safety, wherein in a player receives an guaranteed, low-risk payoff.[1] The strategic normal form of Aumann's version of the game is shown in Figure 1.

Before turning to Aumann's scenario, it is worth saying a few things about the structure of the game. Both players—row and column—have two actions available to them; I've labeled the actions A and B in the interest of presenting the game in as neutral a way as possible, but you can also think of them as "cooperate" and "defect" respectively.

If both row and column play A then both plays get a payoff of 9; if both play B then they both get a payoff of 7. If one player plays A and the other plays B, then the one who played B gets a payoff of 8 and the one who played A gets the worst payoff, 0.

[1]Assurance games have received a good deal of attention. See Dixit et al. (2009) for a broad discussion and Skyrms (2004) for a more detailed study. Clark (2012) discusses stag hunts with particular reference to language games.

		Column A	Column B
Row	A	9,9	0,8
	B	8,0	7,7

Figure 1: Aumann's Assurance Game

We can observe that in situations where one player plays A and the other player plays B, the B-player has every reason to defect and become an A-player: Her payoff had she played B in this circumstance would have been 7 instead of 0; since she prefers the higher payoff, she should have played B. Clearly, then, neither the play $\langle$A,B$\rangle$ nor the play $\langle$B,A$\rangle$ can be an equilibrium, where a play is a (Nash) equilibrium when no player can do better by unilaterally changing her choice.

Notice that there are two "pure strategy" Nash equilibria.[2] First, there is the play $\langle$A,A$\rangle$; by jointly playing A, each player gets a payoff of 9. Unilateral defection to B by one of the players would net him a payoff of 8, which is strictly worse than 9. Second, there is the play $\langle$B,B$\rangle$ which yields a payoff of 7 to each player. If a player unilaterally defects from playing B, he will get a payoff of 0, which is clearly worse than 7.

The two equilibria have some interesting properties. The equilibrium $\langle$A,A$\rangle$ gives a higher payoff to both players than any other equilibrium, so it is a *payoff dominant* (or *Pareto dominant*) equilibrium.[3] While the other equilibrium, $\langle$B,B$\rangle$ has a payoff that is strictly worse than the other, it has a special property that the payoff dominant equilibrium lacks: it has less risk associated with it. A player who plays B is guaranteed a payoff of at least 7, and possibly 8, depending on his opponent's choice. It is the *risk dominant* equilibrium (Harsanyi and Selten (1988)). Notice that if I play A in this game, I accept that possibility that my opponent will play B in which case

[2] A strategy is pure when it is played with a probability of 1.

[3] An equilibrium is payoff dominant when it yields a higher payoff than any other equilibrium for at least one player, and no player nets a lower payoff.

I get nothing and my opponent gets a payoff of 8. Thus, by playing A I'm gambling that my opponent will also play A and I risk getting no payoff at all. Hence, the game in Figure 1 is an assurance game in the sense that I should pick an action just in case I'm sure that you're going to pick the same action.

Now we can turn to Aumann's problem. Suppose that the row player makes the following announcement to the column player:

(1) I plan on playing A.

Should the column player believe her and play A as well?

It seems that the answer should be "yes" since, if the row player is truthful then the column player will get a higher payoff, which a rational agent should prefer. A rational agent, though, might not be willing to risk a sure thing; Aumann points out that the column player might reason as follows:

(2) Row player has said she intends to play A, but she really means to play B; she told me she would play A on the chance that I might believe her and play A, giving her a payoff of 8 instead of 7.

In other words, the row player's statement is *cheap talk* Aumann (1990); in the absence of some external constraint that would force the row player to live up to her statement, the column player is well-advised to be skeptical of what she says.

The situation that the column player finds himself in is a familiar one. Someone claims that I can get a fabulous return if I invest my money with him; should I believe him? I'm more inclined to trust if I have some reason to suppose his word is good. In foreign relations, should adversaries take each other's word?[4]

We can get a better sense of the factors involved if we consider the relationship between probabilities and payoffs—the expected utility—for the two pure strategies in the game. The column player's expected utility for playing A is the probability, p, that the row player plays A times the payoff to column for $\langle$A,A$\rangle$ plus the probability that row player plays B—$1 - p$, since it's the only other choice—times the payoff to column for $\langle$A,B$\rangle$. That is:

[4]One might consider the relationships between the U.S. and North Korea or Iran. Should the U.S. take North Korea at its word if it says it's disarming; should Iran trust the U.S. and Israel and Israel and halt development of nuclear weapons?

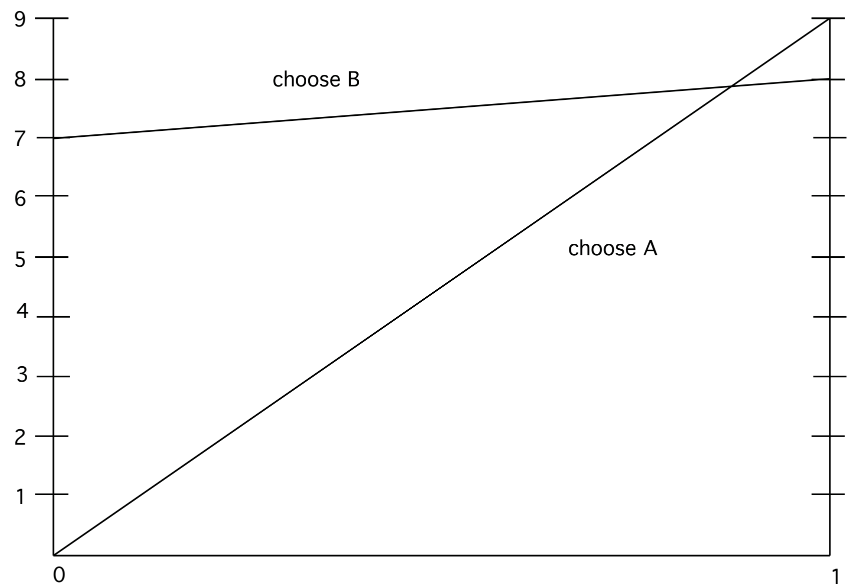


Figure 2: The column player's expected utilities for the pure strategies in Aumann's game

(3) Column Player's expected utility for playing A:
$9p + [0 \times (1-p)] = 9p$

The column player's expected utility for playing B is likewise:

(4) Column Player's expected utility for playing B:
$8p + [7 \times (1-p)] = 7 + p$

I've shown graphed the expected utilities in Figure 2. The x-axis is the probability that row plays A. The y-axis is the resulting payoff to the column player.

The two expected utility lines cross when the expected utility of playing A is equal to the expected utility of playing B:

(5) $9p = 7 + p$
$p = \frac{7}{8} = 0.875$

This is the mixed Nash equilibrium of the game, but we can interpret as the *indifference point* of the game, the point at which column player becomes indifferent between playing A and playing B.[5]

We can now convert the expected utility curves in Figure 2 into a decision rule for the column player. Interpreting p as the column player's subjective probability that the row player will play A, then then column player should also play A if $p > \frac{7}{8}$. If his expectation is that row player will play A with $p < \frac{7}{8}$, then he should play B. At the indifference point, he is indifferent between A and B, so he should just pick one at random. In other words, the column player can simply follow the upper envelope of the expected utilities in Figure 2 and play the corresponding choice.

This bit of arithmetic can help us understand Aumann's point; the column player should be quite sure of row player's real intentions before playing A. If he is less sure, then he is best advised to reduce his risk and play B. Behaviorally, we might expect cautious players to avoid playing B—once they are aware of the risk—until they are virtually certain of the other player's intentions.

3 Examples

In this section, we will turn to two examples of the application of indifference points in language games. In section 3.1, I will work through an example of a quantity implicature. The example is intended to show how mutual confidence can shape the use of conversational implictures. I have selected a quantity implicature for its interest and to illustrate the use of face considerations in shaping the interpretation of indirectness; I do not claim that this is a general solution to quantity implicatures.

In section 3.2 I return to a game treatment of lexical access, discussed in Clark (2012). This example not only illustrates indifference points, but it also illustrates the use of focal points (Schelling (1960)). I will suggest a method of assigning utility to a focal interpretation of a lexical item based in the frequencies of the sense associated with it.

[5] In Clark (2012) I had ignored this potential use of the mixed Nash equilibrium and dismissed it as of little linguistic interest; I now see that I was wrong and the mixed Nash equilibrium—interpreted as an indifference point—is central to understanding implicature (for example), as will be clear below.

3.1 Example: Quantity Implicature

In this section we'll turn to an application of the reasoning illustrated in section 2 to a concrete example, a conversational implicature. I will use a quantity implicature[6] as an example, developing a game tree, a set of payoffs and a solution to the game.

Let's suppose that I announced to you that I intend to bicycle from Philadelphia to Los Angeles and set off on my bike, with you waving farewell; that is, it is mutual knowledge between that I intend to cycle to Los Angeles and I have started on my voyage. After some weeks, you get a phone call from me and I say:

(6) I made it to Albuquerque.

Given the circumstances, the utterance in (6) implicates that I've gone no farther than Albuquerque and that I am unlikely to make it any closer to my goal.

We will take the view that the availability of the implicature in (6) involves a strategic reasoning on the part of both the speaker and the hearer, and that this reasoning is grounded in the context. In particular, the speaker and the hearer are, in the ideal case, aware of the choices available to both. Taking a strategic stance on these choices allows both the speaker and the hearer to develop an account of what is signaled by uttering (6).

My intention to travel across country establishes a scale, starting in Philadelphia and terminating in Los Angeles. Given that I intend to tell the truth (the maxim of Quality) and that I have no interest in either withholding information or saying more than is required (the maxim of Quantity), you and I should both be aware that by uttering (6) in this context I am potentially signaling:

(7) I made it to Albuquerque, and no farther.

Example (7) explicitly reinforces the implicature of (6).

In principle, I could utter (6) with no intention of signaling anything about where I am on the implicit scale established by my itinerary. For example, (6) is consistent with my arriving in Los Angeles and letting you

[6] See Grice (1989) for the basic account. Since Grice, quantity implicatures have been the subject of intense research Horn (2001), Levinson (2000) and Geurts (2010) for a variety viewpoints. I will take Geurts (2010) as correct in the essential details.

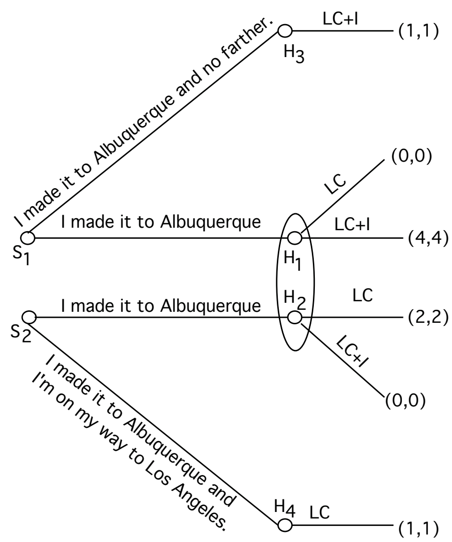


Figure 3: Game tree for a quantity implicature

know that I had at least achieved my ambition of finally seeing Albuquerque; in this case, by uttering (6), I mean only (6). Given this, I might cancel the implicature by uttering, for example, something like:

(8) I made it to Albuquerque and I'm on my way to Los Angeles.

Notice that there are a number of ways that I could cancel the implicature that I got to Albuquerque and no farther, depending on the facts of the matter. I will use (8) as a stand-in for the various ways of accomplishing this.

From the speaker's perspective, there are a number of choices, the outcome of which will signal, more or less faithfully, his intent; given that he intends to signal a particular meaning, is there some expression that is most likely to transmit that meaning to the hearer efficiently? From the other perspective, the hearer can work out the speaker's intentions if she compares his choice with the other potential choices he could have made given the circumstances; given what the speaker has said, the puzzle for the hearer is to work out his intended meaning. The speaker and the hearer are engaged in a joint activity in the sense of Clark (1996). We can represent this joint activity in the game tree shown in Figure 3; both the speaker and hearer are aware of the choices available to both.

The game tree in Figure 3 can be read as follows: Nature places the speaker in one of two information states, S_1 or S_2, according to some probability distribution. In information state S_1, the speaker intends to signal that he made it to a particular point in his traversal of the United States and no farther than that point. I've shown this meaning as "LC+I" in the game tree; that is, "Literal Content" plus "Implied meaning." This shorthand certainly holds for "I made it to Albuquerque" uttered with the intention of implying that Albuquerque is as far as he got. The notation is a bit of an abuse in the case of "I made it to Albuquerque and no farther" since the implied content is made explicit. I have retained the notation for simplicity.

In state S_2, the speaker intends simply to signal that at some point he was in Albuquerque and imply nothing else about the journey; attaching the implied content in this case would be a miscommunication. I've noted this intended meaning as "LC."

By uttering "I made it to Albuquerque" the speaker has not fully disambiguated his intention. This places the hearer in the information set $\{H_1, H_2\}$. In order to solve the game, the hearer will need to work out a plan of action in this case; should she pick "LC+I" or "LC"? In other words, should she draw the implicature or simply infer the literal content of the utterance?

The speaker could clarify things for the hearer, if he so desired. He could, from state S_1, be explicit and say "I made it to Albuquerque and no farther." This utterance is longer, but it reinforces the implicature by making it explicit, allowing the hearer to see his intentions with less uncertainty. Equally, from S_2, he could say "I made it to Albuquerque and I'm on my way to Los Angeles." This cancels the quantity implicature explicitly, again helping the hearer see the intended meaning.

The leaves of the game tree in Figure 3 show the payoffs to the speaker and the hearer; the payoffs are a way of arithmetizing their preferences; we assume that both parties are interested in finding a solution to the game that maximizes their preferences, given what the choices made by the other player. In this analysis, both the speaker and the hearer have preferences that coincide exactly, though this need not be the case.

For present purposes, I'll suppose that the players' preferences are captured by the following principles:

- **Communicative success**: Did the speaker and hearer coordinate on the intended meaning? Choices that fail to so coordinate are given a mandatory zero.
- **Brevity**: All else being equal, speakers and hearers both prefer the shortest form for signaling the intended meaning.
- **Face**: Speakers and hearers attempt to maintain the face of others and, to the extent possible, maintain or increase their own face.
- **Focality**: Given a set of choices, if one option stands out as an obvious point to choose, then its utility is augmented.

The principles of communicative success and brevity should be obvious. If the speaker and hearer fail to coordinate on the speaker's intended meaning, then communication has failed and some repair must be made.[7] Communication success says that if the speaker and hearer fail to coordinate around the speaker's intended meaning, then the payoff to both is zero. The principle of brevity simply says that, all else being equal, speakers will choose the shortest form that expresses their intentions; it follows that choice of a more complex form will signal something.

The principles of face and focality require more discussion. Face involves the social presentation of self (see Goffman (1959) and Goffman (1967), among others). Face involves the adoption of a "stance" or "pose" on the part of a participant in a social interaction:

> The term *face* may be defined as the positive social value a person effectively claims for himself by the line he has taken during a particular contact. Face is an image of self delineated in terms of approved social attributes—albeit an image that others may share, as when a person makes a good showing for his profession or religion by making a good showing for himself.
> Goffman (1967)
> page 5

[7] That is, if the miscommunication is noticed! It's possible that speakers and hearers both fail to coordinate around the intended meaning and fail to note that they failed to coordinate; on miscommunication see, in particular, Labov (2010).

The actual reputation of an agent has been much studied in recent work on game theory (see the textbook by Mailath and Samuelson (2006) for extensive discussion of reputation). We can understand face in contrast to reputation, as revealed by repeated game play. Reputations involve actual history, while face is a stance a person adopts as though they had a particular reputation; thus, face and actual reputation might be at odds with one another.

It is a given in social interaction that people are endowed with both "positive" and "negative" face.[8] Positive face endows the individual with the imprimatur of being a person of good standing in the present social circle, it adopts the pose that her ends are ends approved by the community, that her wants are viewed positively, and so on.

Negative face involves freedom of action and non-interference; the individual is free to choose and initiate actions without external constraints. A person with sufficient negative face is not impelled to act against his will; thus, asking or commanding someone to do something is an immediate threat to their negative face (Brown and Levinson (1987)).

As Brown and Levinson (1987)) argue, speakers and hearers respect the positive and negative face of others—seeing someone "lose face" is mortifying for most people, except when framed as comedy. The avoidance of face threatening acts is a cornerstone of politeness theory. Our interest in the present paper is not so much how speakers avoid threatening other individuals' face but on how speakers use indirection to maintain their own face. Turning to the problem at hand, let us apply reasoning about face to the case at hand, a quantity implicature. Recall that my stated intention at the start of my journey was to bike from Philadelphia to Los Angeles. My utterance:

(9) I made it to Albuquerque.

invites the conclusion that I didn't make it to Los Angeles, the implicature being a tacit admission of failure. Now, I could have formulated my failure more explicitly:

(10) I didn't make it to Los Angeles.

(11) I only made it to Albuquerque (not Los Angeles).

[8] This is foundational in work on politeness, see Brown and Levinson (1987). We differ in that their focus is on the avoidance of "face-threatening acts" (FTAs) while I will also be concerned with acts that promote and construct face.

These two examples involve an overt admission of failure, while (9) avoids the explicit admission of failure. As such, it allows me to maintain my positive face while still communicating my defeat, assuming that my interlocutor picks up my intended meaning, of course. Thus, the interpretation of (9) that includes both the literal content and the implied content should get an extra bit of utility for preserving my positive face.

Let us turn, finally, to the principle of *focality.* A focal point (Schelling (1960)) is item of obvious salience in a set; if I am to coordinate my behavior with another person, then our best course of action might be to choose what we think is a focal point in the set. To modify one of Schelling's (1960) examples, suppose that a couple are separated in a department store and one of them has forgotten his cell phone, so they can't communicate directly. If each person goes to what they think is the other person's favorite choice, they are likely not to meet: Person A might go to the shoe department, thinking person B prefers shoes, while person B might go to the perfume department on the hypothesis that person A is particularly interested in perfume. Instead of thinking what the other would do, they are well-advised to think of what *one* would do—is there an obvious place in the department store that anyone would think of?

Indeed, focality has been shown to have a large effect on coordination games (see Mehta et al. (1994)) and focality has been an area of interest in game theory (see Sugden (1995), Bacharach (2006), and Sugden and Zamarrón (2006), among others)and behavioral game theory (Camerer et al. (2004) and Bardsley et al. (2010), for example). Our approach to focality involves adding to the utility associated with a focal item, thus increasing the likelihood that the players will choose this item, as we will see in the next section and in the discussion below. For the moment, I will assume that Focality adds utility to the interpretation of (9) that includes the quantity implicature on the grounds that, given the context and the available interpretations, it is what a reasonable person would choose.

Returning to the game in Figure 3, we can now work out the payoffs associated with each outcome. First notice that in the case where the speaker utters "I made it to Albuquerque" intending the implicature but the hearer simply draws the literal content of the utterance is a miscommunication and both players receive no payoff. Equally, the case where the speaker does not intend the implicature, but the hearer draws it anyway is also a miscommunication with a zero payoff.

Suppose that the speaker utters (9) with the intention of signaling the implicature and the hearer, in fact, takes the uptake and draws the literal content plus the implicature, LC+I. In this case, the speaker and the hearer successfully communicate (1 point), the message was brief (1 point), and it served the face interests of both the speaker, maintaining his positive face against possible damage by an admission of failure, and the hearer who was spared witnessing the speaker's loss of positive face (so 1 point for each). We assume, although little hinges on the assumption right now, that both the speaker and the hearer regard this interpretation as a focal point in the set of available interpretations (again, for 1 point each). This yields a payoff of 4 for both the speaker and the hearer.

We can contrast this case with the one where the speaker utters (9) with no intention of signaling a quantity implicature and, indeed, the hearer draws only the literal content, LC. In this case, the speaker and the hearer communicated successfully (for 1 point) and the shortest form was used to signal the intended meaning (again, for 1 point). The utterance did not do any significant face work and, given our assumption in the previous paragraph, is not regarded as focal by either the speaker or the hearer. So this outcome garners both the speaker and the hearer 2 points each.

Now consider the case where the speaker intends to signal that he has made it only as far as Albuquerque and has, in fact, failed to make it to Los Angeles and does so by explicitly reinforcing the implicature; the speaker does this by saying, for example:

(12) I made it to Albuquerque and no farther.

In this case, the hearer has no doubt about how to interpret the speaker's utterance. Thus, the speaker and hearer successfully communicate (1 point) although the utterance was not particularly brief, it does not maintain the speaker's face and is only trivially focal. We accord the speaker and the hearer one point each. Equally, the speaker could explicitly cancel the quantity implicature:

(13) I made it to Albuquerque and I'm on my way to Los Angeles.

Once again, the speaker and hearer successfully coordinate around the speaker's intended meaning (for 1 point), but the utterance is not particularly brief and is orthogonal to both the speaker and the hearer's face interests. Finally, the focality of the interpretation is trivial given that it is the only available interpretation. This reading once again garners only one point the speaker and the hearer.

We can now relate this language game to Aumann's assurance game, discussed in Section 2. First, recall that Aumann's game had two sorts of equilibrium states in it: A payoff dominant state in which both players played A and a risk dominant equilibrium where both players play B. A player choosing to play A can get the highest possible payoff—if the other player also plays A—but risks getting nothing if the other player plays B. A player playing B will get less, but also risks less; this player is guaranteed a payoff of at least 7, no matter what the other player does.

Games of partial information, like that in Figure 3, were originally solved by associated the root nodes—S_1 and S_2, in this case—with probabilities.[9] Thus, S_1, the state where the speaker intends the literal content plus the implicature might be associated with a probability p, while S_2, the state where the speaker intends only the literal content, would be associated with probability $1-p$. The outcome payoffs would be multiplied by the probability associated with the node that dominates them to yield the expected utilities of each action.

The puzzle for the players is how to maximize their payoffs given their uncertainty about what their opponent would do. A sensible answer to this was to compute the indifference point—the point at which the expected utilities of the pure strategies were equal. At the point, the players become indifferent as to which strategy to choose. This point could be used to construct a decision rule that took the player's confidence about the other player's potential actions into account.

We can apply this method to the language game in Figure 3. The most puzzling feature of this game is what the hearer should do in the information set $\{H_1, H_2\}$, induced by the speaker's utterance of "I made it to Albuquerque." Should the hearer choose "LC+I" (picking up the implicature) or just choose to associate the literal content "LC" instead? Notice that using "I made it to Albuquerque" yields the highest potential payoffs for the speaker and the hearer, but it also carries the highest potential risk.

In order to decide what to do at this point, the hearer should reason as follows: There are two pure strategies—pick "LC+I" and pick "LC"—suppose that there is a probability p that the speaker intends me to infer "LC+I" and, hence, a probability $(1 - p)$ that the speaker merely intends

[9] See Parikh (2001), Parikh (2010), Clark (2012)

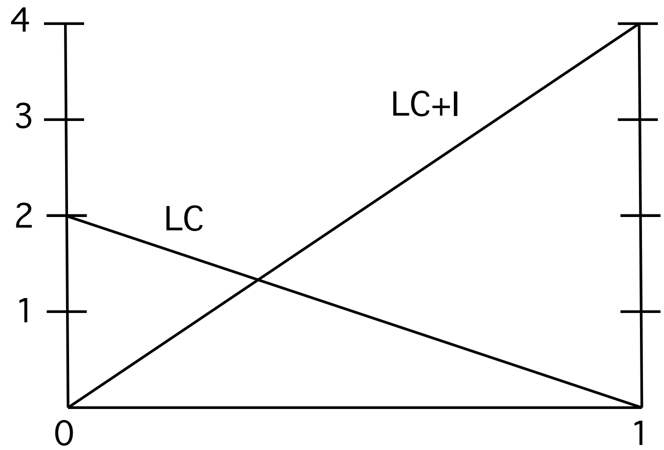


Figure 4: Expected utilities for pure strategies in a language game

the literal content, "LC." My expected utility for playing the pure strategy "LC+I" is:

$$4p + [0 \times (1-p)] = 4p$$

and my expected utility for playing the pure strategy "LC" is

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

The indifference point is given by setting the expected utilities of the two strategies equal:

$$4p = 2 - 2p$$

which is $p = \frac{1}{3}$.

Translating this into action, the hearer should pick "LC+I" if her confidence that the speaker intends "LC+I" is greater than $\frac{1}{3}$, that is, if $p > \frac{1}{3}$; if her confidence in the speaker's intentions is less than that, if $p < \frac{1}{3}$, then she should pick the literal content only. If she is completely uncertain, if $p = \frac{1}{3}$, then she can do better than to randomize her guess.

Let us now consider this result from the point of view of the speaker, who must make a judgment about the hearer's possible actions. He might suppose that the probability that the hearer will choose "LC+I" when he utters *I made it to Albuquerque* is greater than $\frac{1}{3}$; if he indeed intends "LC+I" then that is his best option. Equally, if he estimates the hearer's likelihood of choosing "LC+I" is less than $\frac{1}{3}$ and he intends this meaning, then he is better off choosing the paraphrase *I made it to Albuquerque and no farther* in order to get his meaning across. That is, depending on his assessment of the hearer's likely behavior the speaker can tune his behavior, using the briefest form possible to signal his intended meaning, if he thinks that the hearer will catch his meaning, switching to a more explicit formulation otherwise.

We can now relate this game to payoff dominance and risk dominance; it is clear the strategy profile:

(14) (Speaker: "I made it to Albuquerque and no farther.", Hearer: LC+I)
(Speaker: "I made it to Albuquerque and I'm on my way to Los Angeles.", Hearer: LC)

is an equilibrium since no player has any incentive to unilaterally defect from it and, furthermore, it is a risk dominant equilibrium since neither the speaker nor the hearer risk misunderstanding each other and, therefore, receiving a payoff of zero. Notice, though, that if the speaker elects to follow this profile, he chooses not to do any face work; he does more work for less result, even though he minimizes the risk of being misunderstood.

Suppose that the Speaker is completely uncertain about what the Hearer will do; that is, he supposes that the Hearer estimates the Speaker's probability of using "I made it to Albuquerque" to signal LC+I as $\frac{1}{3}$. In this case, the Hearer is indifferent as to whether to choose "LC+I" or "LC" and there is a significant chance that the Speaker will be misunderstood. In this case, when $p = \frac{1}{3}$, the risk dominant strategy profile in (14) might be preferred over the more dangerous payoff dominant profiles.

Suppose, now, that Speaker and Hearer both judge the probability of using "I made it to Albuquerque" with the intent of signaling a quantity implicature to be greater than $\frac{1}{3}$; the payoff dominant strategy profile in this case is:

(15) ((Speaker: "I made it to Albuquerque.", Hearer: LC+I),
(Speaker: "I made it to Albuquerque and I'm on my way to Los Angeles.", Hearer: LC))

In this case, the speaker is sufficiently confident that the hearer will get the quantity implicature that he can signal it by using the shortest available expression, "I made it to Albuquerque." This sense of the expression blocks the purely literal interpretation, LC, and forces the Speaker to encode this meaning by explicitly canceling the implicature.

Given the confidence of both players in their estimate of the others behavior, there is no reason for them not to select the payoff dominant profile in (15). Thus, when the players estimate that $p > \frac{1}{3}$, the optimal strategy is the payoff dominant strategy that encodes "LC+I" as "I made it to Albuquerque." This is wholly analogous to the treatment of Aumann's Assurance Game discussed in section 2.

We make another prediction: if the players judge that the probability that the Speaker will intend the quantity implicature to be less than one-third, if $p < \frac{1}{3}$, then a different payoff dominant strategy emerges, shown in (16):

(16) ((Speaker: "I made it to Albuquerque.", Hearer: LC),
(Speaker: "I made it to Albuquerque and no farther.", Hearer: LC+I))

That is, the interpretation of "I made it to Albuquerque" is the literal one and the Speaker must use explicit reinforcement of the quantity implicature to get his point across. This change in interpretation is similar to the "pop-out effects" noted in the interpretation of discourse anaphora (Clark and Parikh (2007)):

(17) Mary$_i$ insulted Susan$_j$, so she$_j$ slapped her$_i$.

(18) Mary$_i$ insulted Susan$_j$, then she$_i$ slapped her$_j$.

In example (3.1), the discourse connective *so* increases the probability that *she* is interpreted as Susan (and, hence, that *her* is Mary.) In example (17), the discourse connective *then* has the opposite effect, biasing the interpretation of *she* toward Mary. The presence of the discourse connectives, in other words, conditions the interpretation the pronouns, making one interpretation or the other "pop out" (Clark and Parikh (2007), Clark (2012)).

The treatment of conversational implicature outlined here relies on mutual knowledge, expressed in terms of the speaker and hearer's assessment of the likelihood of potential behaviors. This captures the intuition, discussed in Sally (2003), that speakers opt to use indirect speech when they are confident that their interlocutors will get the uptake. Notice that we have

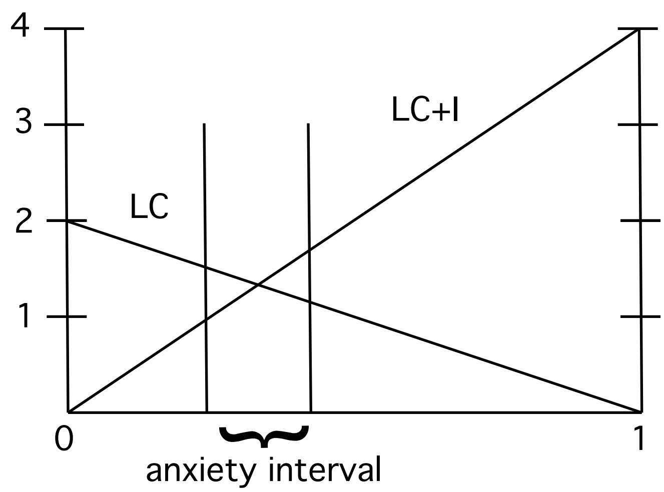


Figure 5: Indifference plus an anxiety interval

generalized the treatment of implicature found in Parikh (2001); this treatment focuses on the payoff dominant nature of conversational implicature. While we agree that indirection is opportunistic and usually involves payoff dominance, our analysis suggests that speakers will sometimes rationally choose the risk-dominant option when they are not confident about hearer's potential behavior.

Our analysis, though, is rather like the analysis of motion assuming a frictionless world. In particular, it seems unlikely that agents are so confident in their assessments of each other that their uncertainty is captured by a single indifference point. Rather, as their probability assessment approaches this point we would expect their uncertainty to grow so that they would become more likely to resort to a risk-dominant strategy profile. We can model this using an "anxiety" constant, δ, that can be used to establish an interval around the indifference point, the anxiety interval. I've shown an example of this in figure 5. We interpret it as follows: if the probability that the speaker intends LC+I in using a signal is greater than $\frac{1}{3} + \delta$ then the hearer should choose "LC+I" in response to the speaker uttering "I made it to Albuquerque" and the speaker should feel free to use this utterance to signal the quantity implicature; that is, the players should use the strategy profile in

(15). If the probability that the speaker intends LC+I is less than $\frac{1}{3} - \delta$, then the hearer should choose "LC" in response to the speaker's utterance. The speaker knows this, so if he agrees in her assessment then he should signal "LC+I" by reinforcing the implicature; that is, they should use the strategy profile in (16). Otherwise, if their assessment is that the probability that the speaker intends the interpretation LC+I is between $\frac{1}{3} \pm \delta$ (the indifference point plus or minus δ), then the players should use the risk-dominant profile in (14).

More generally, suppose that i is the indifference point for a language game, as above. A is an action in the language game and p is the probability of the speaker selecting A. Equally, B is also an action in the game, played with probability $1-p$. For $p > i$, A is the payoff dominant choice, while B is the payoff dominant choice for $p < i$. Finally, suppose C is a risk-dominant profile. We can construct the following decision rule for the language game:

(19) If $p > i + \delta$, play according to A; if $p < i - \delta$, play according to B. Otherwise, if $i - \delta < p < i + \delta$, play according to C.

Thus, if the speaker has confidence that the hearer will pick up on A as an order of play, he should use A; if he has confidence that the hearer will pick up on B, then he should use B. Finally, if the speaker is uncertain about how the hearer will behave, then he is well-advised to pick the risk-dominant profile. This approach to play relies very much on mutual knowledge (Clark (1996)) and accords well with the intuitions of Sally (2003); in planning a joint action, both the speaker and the hearer use mutual information to guide their choices.

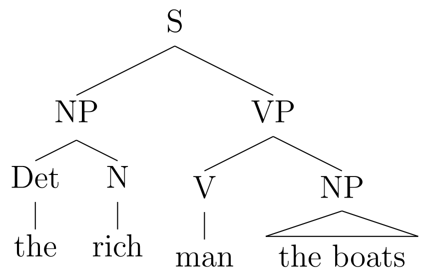
3.2 Example: Lexical Access

Consider a sentence like:

(20) The rich man the boats.

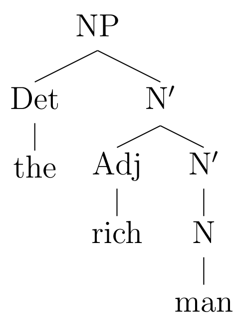
Speakers report that the example in (20) is difficult to process. The intuitive explanation is that the lexical items *rich* and *man* reinforce each other in their most frequent interpretations—*rich* in its adjectival use and *man* as a noun'—which happens to be incompatible with the structure at hand:

(21)



The most frequent interpretations of *rich* and *man* drive the structure:

(22)



which cannot be integrated with the following noun phrase, *the boats*.

Figure 6 shows a game-tree of the sort used in Clark (2012); utility was apportioned according to the following two principles:

- **Communicative success**: Did the speaker and hearer coordinate on the intended meaning? Choices that fail to so coordinate are given a mandatory zero.
- **Brevity**: All else being equal, speakers and hearers both prefer the shortest form for signaling the intended meaning.

Thus, using *man* in its noun sense and having the hearer successfully cooridinate with the speaker around this sense garnered one point and the brevity of using *man* instead of *adult male human* garnered another point. The expected utility was derived by multiplying the utilities by the word frequencies. Thus, this sense of *man* had the associated frequency f_{noun}, so the expected utility of using *man* in this sense would be:

$$2f_{\text{noun}}$$

Equally, the expected utility of using *man* in the verb sense would be:

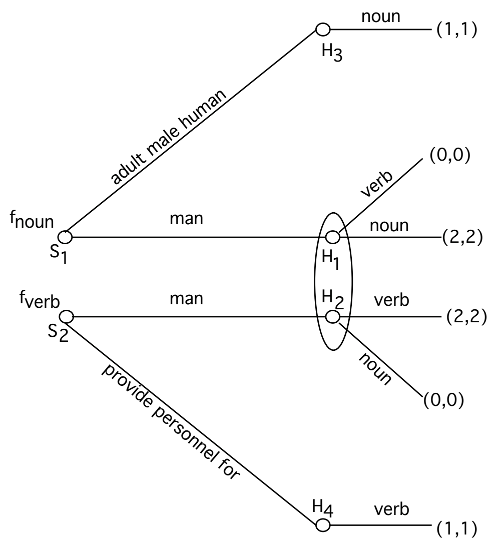


Figure 6: One kind of game for a lexical ambiguity

$$2f_{\text{verb}}$$

Since $f_{\text{noun}} \gg f_{\text{verb}}$, the expected utility of using *man* as a noun would dominate. In addition, the context adds some subjective probability to this use of *man*. If the context were strong enough, then the verb sense of *man* might come to dominate in a pop-out effect.

This analysis is peculiar in the way that it uses frequency to condition both the expected utility of the single word and the periphrastic construction. Thus, the expected utility of *adult male human* would be greater than the expected utility of *provide personnel for* simply because the frequency of *man* in the noun sense is greater than the frequency of the verb sense. This is a peculiar state of affairs, although it had no practical impact on the outcome of the game.

We can go some distance in developing a more sensible account if we observe that frequency and local context can make one interpretation or the other more focal, in the sense discussed above:

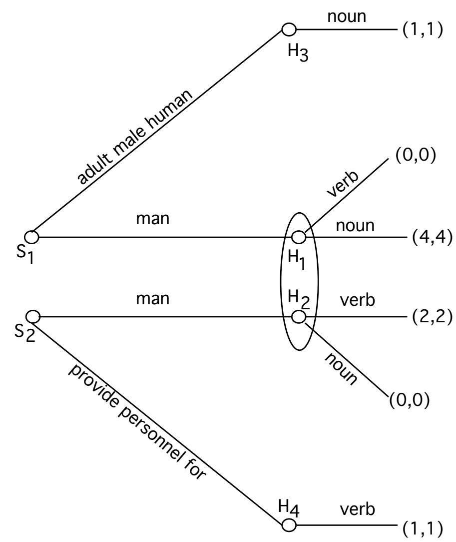


Figure 7: A lexical game with focality

(23) **Focality**: Given a set of choices, if one option stands out as an obvious point to choose, then its utility is augmented.

Thus, highly frequent words and word senses stand out and are more likely to be the center of attention. Equally, the context may cause one or another sense to stand out, as is the case with the adjective sense of *rich* and the noun sense of *man* reinforcing each other in the string *rich man*.

The game in Figure 7 shows how the alternative assignment of utility would work. Successful use of *man* in the noun sense garners 4 points since it communicates successfully (1 point), is brief (1 point), and is focal (2 points). The 2 points for focality follow from the fact that this sense is the most frequent (1 point), and it is reinforced by the local context—the preceding determiner, *the*, and the (apparent) adjective, *rich*. These factors reinforce each other, making the noun intepretation of *man* a stand out.

We should note that the factors that contribute to focality—frequency, syntactic and semantic context, and so on—do not always reinforce each other and can sometimes come close to canceling each other out. Thus, an example like:

(24) The moose head north.

is only a mild garden-path sentence since the compound *moose head* is not significantly more frequent than the head noun *moose* so that the imbalance between *head* as a noun and *head* as a verb is only moderately reinforced.[10]

The analysis of the game in Figure 7 in terms of indifference is essentially identical in form to the quantity implicature analyzed in section 3.1 with one important twist: If the hearer indeed garden-paths on the example in (20), then the speaker has miscalculated in adopting a payoff-dominant strategy when she should have adopted a risk-dominant strategy.

Finally, I have treated frequency as a non-gradable property of lexical items in the sense that the more frequent item or sense is simply accorded one more unit of payoff. It is true, however, that lexical access times are gradable and correlate closely with lexical frequency. While there is no reason to suppose that access time is a direct reflection of expected utility, one might, following recent work in neuroeconomics (see, for example, Glimcher (2003), Glimcher et al. (2005), among many others), suppose that expected utility is directly reflected in the workings of the neuroanatomy. A precise discussion of this issue awaits a more articulated model of the neuroeconomics of lexical access, an issue that would take us far afield. Nevertheless, it is interesting to note how focality and strategic reasoning can partially illuminate these effects.

4 Discussion

We began by considering a well-known version of an assurance game, discussed in Aumann (1990), which illustrates the tension between payoff dominance and risk dominance; while players might be drawn toward playing for the payoff dominant equilibrium, they also run risks. A risk-averse player might prefer a slightly lower payoff associated with risk-dominant play simply for the peace of mind it brings.

[10]Equally, strategic considerations around the misleading focality of certain interpretations might serve to guide speakers' intonation and provide cues to the hearer about the intended interpretation. This point raises interesting question about cooperation—will speakers work harder to ease the processing burden of hearers? A full consideration of this point should be subject to experimental testing, which I have not yet carried out; I will therefore raise the issue only as an aside.

One way to think about the problem is in terms of indifference. The mixed Nash equilibrium for the strategies available to a player, what we have been calling the indifference equilibrium, is the point at which the player is indifferent as to which action to choose; he can do no better than to play the actions available to him at random. If the player, however, has reason to believe that his opponent will play an action with a probability that is greater than the probability at the indifference equilibrium, then he should choose a best response to that strategy.

Suppose you are put in a T-maze and you discover that a reward is placed on the right arm of the T 75% of the time and on the left arm 25% of the time. In order to maximize your reward, your best option is to go right 100% of the time. This guarantees that you will get the reward 75% of the time. Playing a mix of going right 75% of the time and going left 25% of the time—frequency matching—will yield a payoff only 62.5% of the time. Nevertheless, for many species (including our own) frequency matching is a common solution to decision problems of the above form (Herrnstein (1961), McDowell (1988)). Suppose, for example, that a pigeon can peck either a right button or a left button to receive a reward; its choice of which button to peck will be determined according to the rate of reinforcement for the buttons. Thus, if the rate is 25%/75%, as described above, then the pigeon will peck the right button 75% of the time and the left button 25% of the time. While this strategy is less than optimal in a T-maze (or a Skinner box), it actually works quite well when foraging for food in a dynamic world (Gallistel (1990)).[11]

References

Aumann, Robert. 1990. Nash equilibria are not self-enforcing. In *Economic decision making: Games, econometrics, and optimization: Essays in honor of Jacques Dreze*, ed. J.J. Gaszewicz, J.-F. Richard, and L. Wolsey, chap. 34, 201–206. Amsterdam: Elsevier Science Publishers.

Bacharach, Michael. 2006. *Beyond individual choice: Teams and frames in game theory*. Princeton, NJ: Princeton University Press. Edited by Natalie Gold and Robert Sugden.

[11]Thanks to Steve Kimbrough for reminding me of this.

Bardsley, Nicholas, Judith Mehta, Chris Starmer, and Robert Sugden. 2010. Explaining focal points: Cognitive hierarchy theory versus team reasoning. *The Economic Journal* 120(543):40–79.

Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some universals in language usage*. 2nd ed. Studies in Interactional Sociolinguistics 4, Cambridge, U.K.: Cambridge University Presss.

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. 2004. A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119(3):861–898.

Clark, Herbert H. 1996. *Using language*. Cambridge, U.K.: Cambridge University Press.

Clark, Robin. 2012. *Meaningful games: Exploring language with game theory*. Cambridge, MA: The MIT Press.

Clark, Robin, and Prashant Parikh. 2007. Game theory and discourse anaphora. *Journal of Language, Logic and Information* 16:265–282.

Dixit, Avinash, Susan Skeath, and David Reiley. 2009. *Games of strategy*. 3rd ed. New York, NY: W. W. Norton & Company.

Gallistel, Charles R. 1990. *The organization of learning*. Cambridge, MA.: The MIT Press.

Geurts, Bart. 2010. *Quantity implicatures*. Cambridge, U.K.: Cambridge University Press.

Glimcher, Paul W. 2003. *Decisions, uncertainty, and the brain: The science of neuroeconomics*. Cambridge, MA.: The MIT Press.

Glimcher, Paul W., Michael C. Dorris, and Hannah M. Bayer. 2005. Physiological utility theory and the neuroeconomics of choice. *Games and Economic Behavior* 52:213–256.

Goffman, Erving. 1959. *The presentation of self in everyday life*. New York: Anchor Books.

———. 1967. *Interaction ritual: Essays on face-to-face behavior*. New York: Pantheon.

Grice, H. Paul. 1989. *Studies in the ways of words*. Cambridge, MA: Harvard University Press.

Harsanyi, John C., and Reinhard Selten. 1988. *A general theory of equilibrium selection in games*. Cambridge, MA: The MIT Press.

Herrnstein, Richard J. 1961. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior* 4:267–272.

Horn, Laurence R. 2001. *A natural history of negation*. Stanford, CA: CSLI Publications.

Labov, William. 2010. *Principles of linguistic change: Cognitive and cultural factors*, vol. III. New York: Wiley-Blackwell.

Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: The MIT Press.

Mailath, George J., and Larry Samuelson. 2006. *Repeated games and reputations: Long-run relationships*. New York, NY: Oxford University Press.

McDowell, J. J. 1988. Mathcing theory in natural human environments. *The Behavior Analyst* 11(2):95–109.

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

Parikh, Prashant. 2001. *The use of language*. Stanford, CA: CSLI Publications.

———. 2010. *Equilibrium semantics*. Cambridge, MA: The MIT Press.

Sally, David. 2003. Risky speech: Behavioral game theory and pragmatics. *Journal of Pragmatics* 35:1223–1245.

Schelling, Thomas C. 1960. *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Skyrms, Brian. 2004. *The stag hunt and the evolution of social structure*. Cambridge, U.K.: Cambridge University Press.

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

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