A Game-Theoretic Account of Adjective Ordering Restrictions

Lucas Champollion

March 29, 2006

1 Introduction

This paper sets out to provide a game-theoretic explanation of some facts about adjective ordering restrictions (AOR) a phenomenon that has been widely discussed in the literature but is not yet well understood. [Teo06] notes that AOR do not choose between structures that are truth-conditionally distinct. She demonstrates that earlier attempts by [Cin94] to give a syntactic formulation of AOR cannot be extended to incorporate this constraint and concludes that a semantic implementation is needed. I will show that this semantic constraint follows immediately if we assume a game-theoretic framework as described in [Par01]. Moreover, I will use game theory and corpus data to provide a speculative explanation of why AOR should arise in natural language.

2 The Puzzle

Crosslinguistically, sequences of multiple adjectives that modify a noun occur in an unmarked order. Consider the following English examples:

(1) a. a great green dragon
    b. # a green great dragon

(2) a. a short Italian student
    b. # an Italian short student

As surveyed in [Teo06], several proposals have been made that generalize over the semantic classes of adjectives that pattern together with respect to these restrictions. In these scales, > is the relation "must occur farther away from the head noun than".

(3) a. QUALITY > SIZE > SHAPE > COLOR > PROVENANCE [SS91]
    b. POSSESSIVE > SPEAKER-ORIENTED > SUBJECT-ORIENTED > MANNER/THEMATIC [Cin94]

c. VALUE > DIMENSION > PHYSICAL PROPERTY > SPEED > HUMAN PROPENSITY > AGE > COLOR [Dix82]

However, none of these scales are *explanations* for why there should be restrictions on the order of adjectives in the first place. In the case of intersective adjectives such as *great* and *green*, the meaning of the whole noun phrase is the same regardless of the order in which the adjectives combine. So it is not possible to argue that differences in meaning correlate with, or impose, the order restriction.

To make matters even more complicated, non-intersective adjectives, where order does matter for the meaning of the noun phrase, are exempt from AOR:

(4) a. an alleged former thief

b. a former alleged thief

As this example illustrates, it is possible to use non-intersective adjectives in whatever order is necessary to express the intended meaning of the utterance. The fact that semantic distinctions between adjectives (i.e. intersective versus non-intersective) correspond to restrictions on surface order has led [Cin94] to attempt a syntactic explanation. According to him, intersective adjectives are generated in the specifier position of dedicated syntactic categories such as ColorP, SizeP etc, which occur in a fixed order in the syntactic tree. Non-intersective adjectives, by contrast, are generated as adjuncts and as such can be inserted anywhere in the sentence.

As [Teo06] shows, however, this account is empirically flawed, since it predicts that any given adjective is either always fixed or always free with respect to order, no matter what environments it occurs in. Citing the example of nondefinite superlatives [HS06], she shows that the meaning of the whole utterance, not individual lexical properties of the adjectives involved, matters for the free/fixed distinction. Consider the following examples:

(5) Every class in this school has a shortest student.

(6) a. My class has a shortest Italian student.
Meaning: In my class, there is a unique Italian student who is shorter than every other Italian student.

b. My class has an Italian shortest student.
Meaning: In my class, the shortest student is Italian.

As this example shows, the usual restriction on adjective ordering is lifted when the resulting sentences differ in their truth-conditions. Since the adjectives *short* and *Italian* may not normally occur in just any order, this contradicts any account in which adjectives are grouped neatly into fixed-order and free-order. Rather, it is the meaning of the whole sentence that matters: note that once again, the difference in order corresponds to a difference in meaning. [Teo06] generalizes this pattern as follows: AOR do not apply to sequences where different linear orders yield different semantic interpretations, even if the sequences contain adjectives which may in other contexts be subject to ordering restrictions. In other words: *AOR only discriminate between paraphrases.*

This generalization is easy to describe but less easy to implement, at least in a classical framework in the style of [Mon70] or [Cho95]. The restrictions on adjective ordering are determined not just by the meaning of the sentence as a whole, but also by the meaning of alternative potential sentences which result from reordering the adjectives in a different way. If the sentence and its alternatives differ in meaning, then AOR are blocked from applying. Under a "pipeline" view in the Chomskyan/Montegovian tradition, the syntactic structure (LF) of a sentence is sent to the conceptual interface, where it serves as an input to the semantic module which computes the meaning of that sentence. It is not possible for the semantic module to go back and ask for alternative sentences or to cause the sentence to crash except by presupposition failure (and this option is clearly out of the question here). However, AOR apply to sentences, not to meanings, because they make a distinction between candidates that are truth-conditionally equivalent. There is no way to distinguish between "a green great dragon" and "a great green dragon" once they have passed the syntax-semantics interface.

The questions to be answered, then, are the following: why do AOR hold at all; why do they only discriminate between paraphrases of the same meaning; and how can the explanation be implemented in an actual semantic theory? These questions will be addressed in the remainder of this paper. Although it is not possible to answer them exhaustively, at least a partial answer to these questions follows for free if one assumes a game-theoretic framework.

The remainder of this paper is organized as follows. In section 3, I give a short introduction into the [Par01] framework, Game Theoretic Semantics. In section 4, I implement AOR restrictions in this framework and show that the generalization described in [Teo06] (AOR only discriminate between paraphrases) is a direct consequence of the framework's properties. Finally, section 5 offers a speculative, game-theory based answer to the reason for the presence of AOR in natural language. The plausibility of this answer is tested using corpus data.

3 Game Theoretic Semantics

[Par01] views communication as a series of games of partial information between a speaker (A) and a hearer (B).[1] A selects which games to play, but their outcome depends on the common knowledge of A and B about their information states and the language they use. What is crucial for our purposes is that syntactic, semantic and pragmatic factors all influence these games and hence the content of any particular utterance. In contrast to standard generative/Montegovian models, this is not a sequential architecture. That is, there is

[1]The framework and applied here should not be confused with the "game theoretical semantics" described in [HS97], an approach which has very little in common with [Par01].

no "pipeline" in which syntactic structure is constructed first, then sent through an interface and further processed by a semantic module. Rather, the structure and meaning of an utterance are computed simultaneously and can influence each other.

Game-theoretic semantics is at its best when modeling the resolution of ambiguity or vagueness. A situation that lends itself to a game-theoretic treatment typically involves a tradeoff between an unambiguous but overly cumbersome expression and a shorter but ambiguous one. The speaker selects the expression that best fits this tradeoff and makes sure that the hearer would have no reason to misunderstand her in case she selects the ambiguous option.

To take a concrete example from [Par01], consider the situation in which A wants to communicate to B that every ten minutes, a different man gets mugged in New York. The cumbersome option is for A to say unambiguously: "Every ten minutes, a different man gets mugged in New York." However, it would cost her less effort just to use a shorter sentence: "Every ten minutes, a man gets mugged in New York." This sentence is of course scopally ambiguous – it can also mean that there is an unfortunate man who gets mugged in New York every ten minutes. However, it is a priori highly unlikely that this would be the actual state of affairs in the world, and thus that A would want to communicate this. Because A and B mutually know about this, they know that in such a rare case A would reasonably use a cumbersome and unambiguous expression like "There is a particular man who gets mugged in New York every ten minutes." For this reason it is safe to assume that the literally ambiguous expression refers to the more likely state of affairs. Because all this is mutual knowledge, A can utter the literally ambiguous expression safely in order to convey the more likely situation.

This is realized in Parikh's framework in the following way (see Fig. 1). The situation is represented as a pair of intertwined games of which only one is the actual game that A is playing. Each game corresponds to a possible message that A might want to convey to B. B doesn't know which of the games A is playing.

A has the first move – selecting the utterance – followed by B, who interprets the utterance according to her knowledge of language. The games are rooted in two situations s and s', which correspond the two possibilities for what A might want to communicate to B. For the example, assume that s is the situation in which A wants to communicate that a different man gets mugged every ten minutes, and s' is the (counterfactual) situation where A wants to communicate that there is a particular man who gets mugged every ten minutes. While only one of these situations is factual, only A knows which one. Crucially, though, both A and B have exact knowledge of how likely it is a priori that each situation holds. In the example, assume that the probability of s, $\rho(s)$, is 0.9 and the probability of s' is $\rho(s') = 0.1$. Thus, by assumption, both parties know exactly which situation is more likely to be the case, even though only A knows initially what is the case.

Now A has two choices per situation as to what to utter. The ambiguous expression "Every ten minutes, a man gets mugged", represented as φ, is available

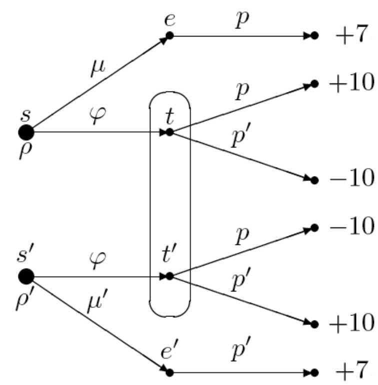

Figure 1: The standard example of a local game

in both situations. In addition, for each situation, there exists an unambiguous way of describing it. For situation s, this is the utterance “Every ten minutes, a different man gets mugged”, represented as μ, while in s’, μ' corresponds to the uttering of “Every ten minutes, the same man gets mugged”. The resulting states of the two games are labeled e, t, e’, and t’. Crucially, there is no obvious way for B to distinguish between states t and t’ since the same (ambiguous) expression leads to them.

B can now interpret A’s utterance. If it was unambiguous, B has by definition only one way to interpret it. This is represented by having only one branch leaving each of the points e and e’ which are the result of A making only unambiguous utterances. In contrast, if the utterance was ambiguous, then B has two ways of disambiguating it, labeled p and p’ in the figure.

The rightmost points corresponds to the different beliefs B may have about what A intended to communicate her. In some of them, B’s beliefs are right, while in others, she has misunderstood A. The payoffs of the different outcomes are indicated next to the points. They are mainly determined by whether successful communication has taken place, but a small amount of points is subtracted in the case of unambiguous utterances as a result of the increased effort the speaker has to undergo to produce them. By assumption, the payoffs are identical for both players, making the game a cooperative one.

Given this structure of the game and given that the game is common knowledge between A and B, the players can in principle opt for different strategies. For example, A might choose to avoid any ambiguity by selecting only e and e’, depending on the initial situation. However, a better choice for her might be to use the ambiguous utterance φ in the more likely situation s and to revert to unambiguous speech only in the less likely case s'. Whether this is actually a better choice depends on B's strategy, which A therefore has to take into account. In this case, the overall payoffs increase only if B's strategy interprets t and t' (which she cannot distinguish and therefore has to treat identically) to mean p.

In general, a pair of strategies such that neither player has reason to unilaterally switch to another strategy is called a *Nash equilibrium*. The expected payoff of a Nash equilibrium is the sum of outcomes of the different situations involved, weighted by the probabilities of these situations. In the case that a game has more than one Nash equilibrium, then the players are assumed to orient their strategies toward the one with the highest expected payoff compared to any other Nash equilibria in the game. This point, if there is one, is called the *Pareto-Nash equilibrium*. By assumption, A can trust B to be always aiming towards the Pareto-Nash equilibrium of the game.

For example, in Fig. 1, there are two Nash equilibria. One corresponds to A choosing φ in situation s and μ' in situation s' and B choosing p whenever she has a choice (i.e. in t and t'). The other one corresponds to A choosing μ in situation s and φ in s' and B choosing p' whenever she has a choice. However, there is only one Pareto-Nash equilibrium. The expected payoff for the first strategy, assuming $\rho = 0.9$ and $\rho' = 0.1$, is $\rho(10) + \rho'(7) = 0.9(10) + 0.1(7) = 9.7$. By contrast, the expected payoff for the second strategy is only $\rho(10) + \rho'(7) = 0.9(7) + 0.1(10) = 7.3$. Since there is a unique Pareto-Nash equilibrium, both players can converge to it and so communication becomes possible.

4 Application to Adjective Ordering

Recall the constraint observed in [Teo06]: *Ordering relations do not choose between structures that are truth-conditionally distinct.* As discussed above, it is not possible to predict out of context whether a given adjective is going to underlie an ordering constraint. This is because if it is intersective, it may occur in different constructions of which only some (the nondefinite superlatives) will mean different things dependent on the order of the adjectives. Therefore, there is no (obvious) way to implement them in the syntax cleanly. However, AOR discriminate between certain sequences of adjectives which are truth-conditionally identical. Therefore, there is no way to implement them in the semantics either.

Under the game-theoretic approach, this conundrum disappears. All we have to assume is that certain orderings of adjectives are in principle possible but disfavored, in the sense that it is a legal move in the game to utter them but that this results in a greater cost (all else being equal) than their "correctly ordered" counterparts. We can remain agnostic as to whether there is a correlation between the meaning of a certain adjective (for example, whether it is intersective or not) and its being affected in that sense by the ordering preference. For example, we might assume for simplicity that all adjectives are subject to ordering preferences across the board. Nor do we have to assume that certain constructions (say, nondefinite superlatives) somehow make this disfavoring disappear.

First consider the classical case. Suppose A sees a great green dragon over there and wants to communicate this to B. There are only two (relevant) potential utterances that A could use in order to do this: Either "There's a great green dragon over there" or "There's a green great dragon over there." Since the adjectives *great* and *green* are both intersective, these two utterances are truth-conditionally unambiguous, and so B has only one way of interpreting each of them (in fact, they are paraphrases, so B's interpretation of each of them is identical). However, one of them has a disfavored order and therefore results in a somewhat diminished payoff. (I follow [Par01] in assuming that the payoffs for both players are equal. However, it does not matter for this analysis whether the payoff diminishes only for A when she uses a disfavored sequence.) For this reason, the Pareto-Nash equilibrium of this game is the strategy in which A chooses the branch with no order violations (see Fig. 2).

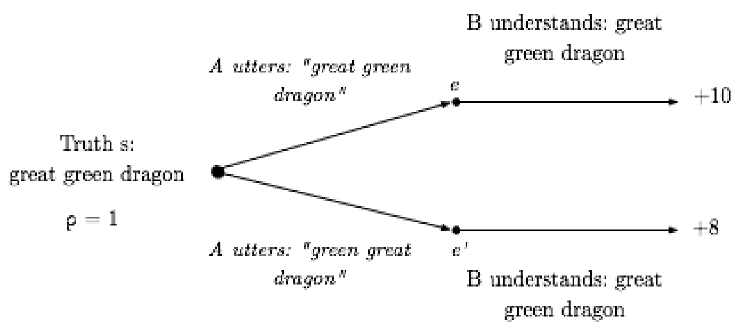

Figure 2: Intersective adjectives: AOR as a payoff malus.

Now consider the case of nonintersective adjectives. Assume that A wants to communicate to B that there is an alleged former thief over there. The only available way to express this information is the (unambiguous) utterance "There's an alleged former thief over there." It is not possible to use the utterance "There's a former alleged thief over there" because this would unambiguously describe a different situation.

Therefore, A and B play a trivial game here, in that A has exactly one available (relevant) way of communicating what she intends to, and because the utterance A uses is unambiguous, no confusion can arise for B as to which game is played. (See Fig. 3.) The payoffs of this game don't matter for determining its solution. Therefore we can remain agnostic as to whether any ordering constraints apply to the particular order that the adjectives in A's utterance occur in. Since we are modeling these constraints by subtracting from payoffs of the relevant sentences, what will result is that the only solution to the game will have a lower payoff. This, however, doesn't affect its status as the Pareto-Nash equilibrium.

The constraint that AOR only choose between truth-conditionally equivalent orderings is a direct consequence of the game-theoretic framework. Since the

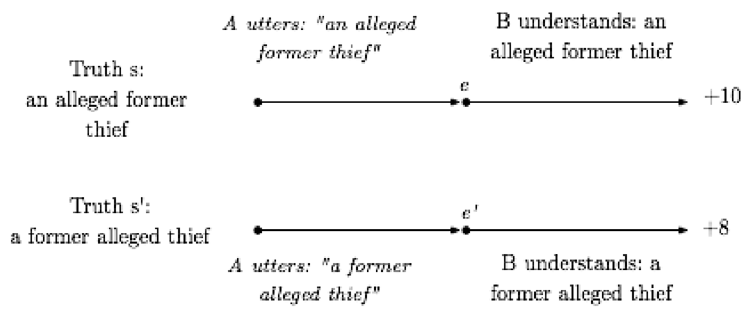

Figure 3: Truth-conditionally different orderings: AOR payoff mali apply vacuously.

choice of what to say (i.e. which information to convey by some utterance or other) occurs prior to the game, the only candidates that are even available for a comparison with regard to AOR are the truth-conditionally equivalent ones.

The second puzzle that [Teo06] observes, and that causes a problem for the syntactic account in [Cin94], is that in the nondefinite superlative construction, even intersective adjectives, which are otherwise subject to AOR, are exempt from them. This too follows from a game-theoretic perspective: There is nothing distinctive about any particular adjective. AOR apply across the board, but they only have an effect if there is more than one way to express a given content. This is not the case with the nondefinite superlative construction, since different orderings of adjectives are truth-conditionally distinct. Therefore, the game in the case of A wanting to convey the information that this class has a shortest Italian student is going to look like in Fig. 4. As can be seen, it is completely separated from the game involving A wanting to convey that this class has an Italian shortest student. Both meanings are quite separate, and there is no ambiguity in the phrases used to convey them, so the two games don't interfere with one another.

5 The Emergence of AOR

In the previous section, I have shown that a game-theoretic framework can account naturally for [Teo06]'s generalization that AOR choose only between truth-conditionally distinct cases. However, a crucial assumption was that there is something which causes different orderings of adjectives (truth-conditionally distinct or not) to have different payoffs. Although no further assumptions were needed to explain [Teo06]'s generalization, the question of why AOR should be present in the first place has of course not been answered yet. Nor has it in general, and I do not claim to have found a solution to that puzzle. All that

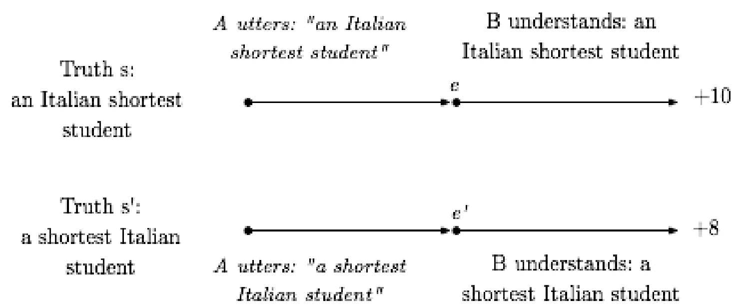

Figure 4: Nondefinite superlatives work just like nonintersective adjectives.

will be done here is to make a suggestion as to why AOR *might* have arisen.

Recall that in the discussion so far, we have remained agnostic as to whether AOR apply only in cases that are truth-conditionally equivalent or in all cases, including truth-conditionally distinct ones such as nonintersective adjectives. In the latter case, they apply vacuously in the sense that they diminish the payoff of the only solution to a trivial game, and therefore do not affect its outcome. While it is not possible to distinguish empirically between the two cases, the possibility that they apply across the board is conceptually simpler. Let us assume therefore that AOR can apply *only* (1) not at all or (2) across the board, but they cannot apply selectively. The following argument intends to show that any language of type (1) is bound to change into a language of type (2) in order for conversation to become possible.

Consider a hypothetical language in which AOR *never* apply, to the extent that adjectives can be freely ordered even if some of them are nonintersective. The idea is that if any such language existed, it must necessarily have developed restrictions on adjectival ordering. For convenience, let us pretend that this language is like English in all other respects, though nothing in the argument hinges on this assumption. Thus, our hypothetical language allows both ”a great green dragon” and ”a green great dragon”. Moreover, in this language the two expressions ”an Italian shortest student” and ”a shortest Italian student” each have two readings and these readings are the same for the two expressions. That is, it is possible to describe a situation involving an Italian shortest student both by uttering ”an Italian shortest student” and by uttering ”a shortest Italian student”; and either of these utterances could also refer to a situation involving a shortest Italian student.

Now consider the game that corresponds to this state of affairs (Fig. 5). There is no Pareto-Nash equilibrium, because there is nothing available in the game to break the symmetry. It is easy to see that the only strategies that would be potential candidates for a Pareto-Nash equilibrium are the following:

- In the first strategy, A refers to situation s (the one that involves an Italian shortest student) as "an Italian shortest student" and to s' (the situation that involves a shortest Italian student) as "a shortest Italian student", while B chooses to interpret any utterance of the form "an Italian shortest student" as referring to situation s and any utterance of the form "a shortest Italian student" as referring to s'. In other words, this strategy happens to work just like English. Call it N_1.
- Competing with this strategy is its mirror image, in which A refers to s by the expression "a shortest Italian student", and B interprets this as referring to s; and A refers to s' by the expression "an Italian shortest student" and B interprets this as referring to s. Call this strategy N_2.

If the payoffs and probabilities are identical, then neither of these strategies Pareto-dominates the other, though both are Nash equilibria. This holds even if we assume (as in the figure) that the probabilities for s and s' are different and by this introduce an element of asymmetry into the game. (In the example in the figure, the expected payoff to both players from N_1 is $\rho(10) + \rho'(10) = .4(10) + .6(10) = 10$. The expected payoff to both players from N_2 is again $\rho(10) + \rho'(10) = .4(10) + .6(10) = 10$.) The players still have no way to know which strategy they will choose in order to maximize payoffs, and so no successful communication is possible.

Consider now what happens to our game as a result of imposing some restriction on adjective order, that is, penalizing any utterance that goes against an arbitrarily chosen ordering constraint (Fig. 6). Unless the probabilities for s and s' happen to be identical, the symmetry between the Nash equilibria is now broken, and there is now a Pareto-Nash equilibrium: In the example, the expected payoff to both players from N_1 is $\rho(8)+\rho'(10) = .4(8)+.6(10) = .92$, while the expected payoff to both players from N_2 is only $\rho(10)+\rho'(8) = .4(10)+.6(8) = .88$. Thus the two strategies result in different payoffs, and so the players can communicate successfully.

Since this reasoning does not depend on knowing whether s or s' is the case but only on the other aspects of the game, and since these are mutual knowledge, then this reasoning is also mutual knowledge. In other words, A and B both know that if they agree once and for all on an arbitrary ordering (in the sense of diminished payoffs), then in all situations where the probabilities don't happen to be exactly equal, they can communicate successfully if only it is the same ordering they choose. But how should they choose the same ordering if they have no way to communicate with each other on which one they should choose? (After all, this game is supposed to model communication, so as [Par01, p. 36] remarks, we may not allow the players to communicate outside the game if we are not to run into an infinite regress.)

Clearly, the players have to agree on some ordering that each of them thinks the other is more likely to choose. The solution is for both players to agree on what [Sch63] calls a *focal point*: what players will select is what seems natural or salient to them. As an example, consider Schelling's pseudo-experiment carried out on a sample of 41 residents of New Haven, Connecticut. They were

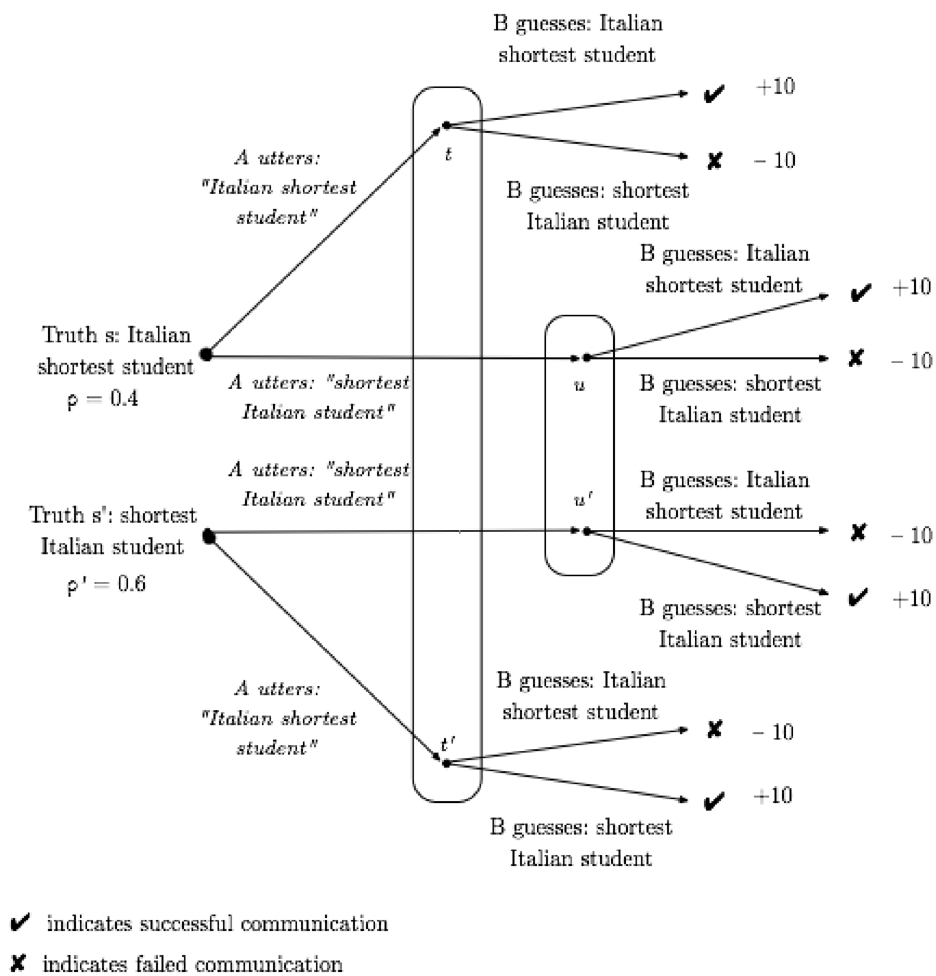

Figure 5: In a language with no restrictions on adjective ordering, the lack of Pareto-Nash equilibria makes communication impossible.

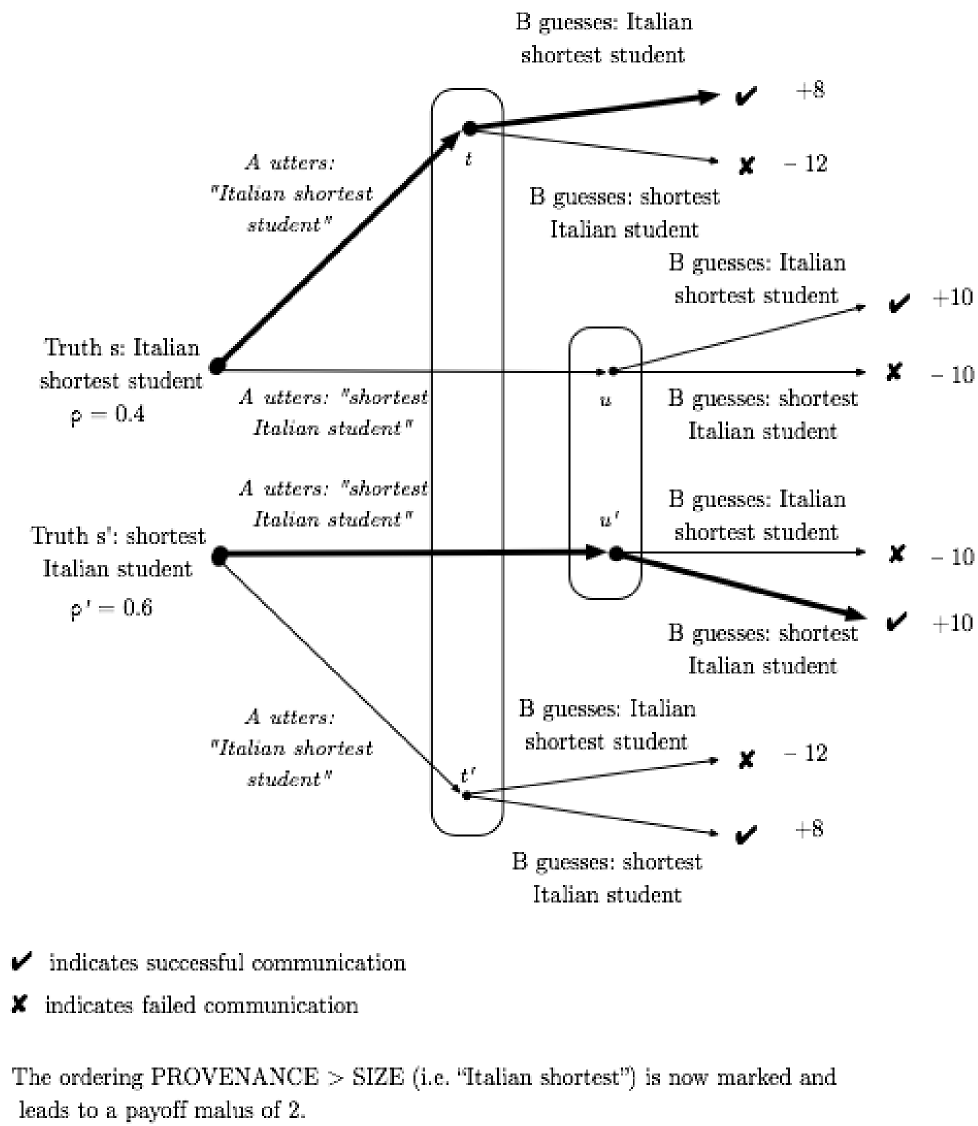

Figure 6: Arbitrarily imposing payoff penalties on some ordering sequences leads to the emergence of a Pareto-Nash equilibrium strategy (shown in bold).

given the instruction (as mutual knowledge) that they were to meet somebody in New York City at a given day with no prior understanding or further instruction as to the exact time and place, and with no chance to communicate. An absolute majority managed to get together at the information booth of the Grand Central Station, and virtually all of them succeeded in meeting at noon [Sch63, p. 55]. Similarly, a focal point for adjective orderings would arguably be to take some cognitively imposed and therefore mutually known breakdown of adjectives into classes, then start with the most common class (i.e. the class whose prototype/most prominent members occur most frequently in language use), followed by the second most common, etc.

There is some empirical justification for this speculation. Consider the scale (3a) proposed by [SS91] as a descriptive generalization on adjective order and repeated here as (7) for convenience:

(7) Quality > Size > Shape > Color > Provenance

If we compare this scale with a frequency-ranked list of adjectival lemmas taken from the British National Corpus [LRW], we find that it corresponds almost exactly with the order in which the most frequent member of each class appears in that list. That is, the most frequent quality adjective is more frequent than the most frequent size adjective and so on. This is what we would expect if quality adjectives as a whole were more frequently used than size adjectives as a whole.

- Quality: `good` (rank in frequency list: 2)
- Size: `high` (6)
- Shape: `long` (13)
- Color: `black` (40)
- Provenance: `British` (16)

As can be seen, the only outlier is the word *British*, which has its place at the end of the adjective order but which occurs more frequently than the most frequent color adjective, *black*. It might be an artifact of the data (this corpus has been collected in a single country, at a historical time where the concept of nation state is more salient than ever before; in contrast, it is unlikely that the frequency of adjectives such as *good* or *black* would change over even extended periods of time).

In any case, a small number of words can't be considered enough to corroborate the focal point theory. Certainly, a more extended analysis, including historical corpora, would be needed in order to determine if the order of adjective classes is indeed based on frequency. However, as a first result it seems that there is a correlation here. This seems especially likely given that the most common adjective that occurs in the BNC, *other*, always comes first in sequences of adjectives.

As a reminder, we have assumed that AOR may only occur across the board, or in other words, players do not distinguish between cases in which AOR apply and others where it doesn't. Note that this assumption makes sense because players have no incentive to apply AOR selectively. In cases like the one just discussed, AOR make communication possible in the first place; in cases where AOR would eliminate one of several truth-conditionally equivalent candidates, communication is not impaired because A can simply select the utterance that conforms to the ordering restriction. Speakers might remember to apply AOR only to those cases in which it is necessary for successful communication, but they might find it easier to apply it across the board. This, then, is the explanation for AOR: it is an epiphenomenon of the players' need to agree on differentiated payoffs in those cases where the order in which adjectives modify each other matters for the meaning, and of the players' convergence towards focal points.

6 Conclusion

Adjective ordering restrictions have been implemented in a game-theoretic framework. Their interplay with the semantics of their host sentence and of alternative sentences can be easily modeled in this framework, and certain generalizations about AOR fall out directly from the choice of framework. While the ultimate cause of AOR remains unexplained, a possible explanation of the emergence of AOR as a focal point has been given. Corpus data shows a possible correlation between word frequency in language use and place in the adjective sequence that would need to be explored further. If confirmed, such a correlation could be made sense of in a game-theoretic framework, while an explanation in models of language that make no reference to language use appears more difficult.

References

[Cho95] Noam Chomsky. *The Minimalist Program.* Cambridge, MA: MIT Press, 1995.

[Cin94] Guglielmo Cinque. On the evidence for partial N-movement in the Romance DP. In Guglielmo Cinque et al., editor, *Paths towards Universal Grammar. Studies in honor of Richard S. Kayne*, Georgetown Studies in Romance Linguistics, pages 85–110. Washington, DC: Georgetown UP, 1994.

[Dix82] R.M.W. Dixon. *Where have all the adjectives gone? and other essays in semantics and syntax.* Berlin: Mouton, 1982.

[HS97] J. Hintikka and G. Sandu. Game-theoretical semantics. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language.* Elsevier, Amsterdam, 1997.

[HS06] Simona Herdan and Yael Sharvit. Definite and non-definite superlatives and NPI licensing. Manuscript, University of Connecticut, 2006.

[LRW] Geoffrey Leech, Paul Rayson, and Andrew Wilson. Frequency list of adjectives (by lemma) in the British National Corpus. Online Mar 29, 2006: http://www.comp.lancs.ac.uk/ucrel/bncfreq/lists/5_3_all_rank_adjective.txt.

[Mon70] Richard Montague. English as a formal language. In Bruno Visentini et al., editor, *Linguaggi nella societa e nella tecnica*. Milan: Edizioni di Comunita, 1970.

[Par01] Prashant Parikh. *The Use of Language*. CSLI Publications, 2001.

[Sch63] Thomas C. Schelling. *The Strategy of Conflict*. Harvard UP, 1963.

[SS91] Richard Sproat and Chilin Shih. The cross-linguistic distribution of adjective ordering restrictions. In C. Georgopoulos and R. Ishihara, editors, *Interdiscplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*. Dordrecht: Kluwer Academic Publishers, 1991.

[Teo06] Alexandra Teodorescu. Adjective ordering restrictions and the interface between meaning and form. Handout. 30th Penn Linguistics Colloquium, University of Pennsylvania, 2006.