Ambiguity in Anaphora Workshop Proceedings, pages 73–82
ESSLLI 2006, Málaga, Spain, 7–11 August 2006
Edited by Ron Artstein and Massimo Poesio
http://cswww.essex.ac.uk/anaphora/

# On pronouns in Catalan and game theory[1]

Laia Mayol
University of Pennsylvania
laia [at] babel.ling.upenn.edu

The goal of this paper is to explore whether game theory can account for the distribution of pronouns in Catalan, a null-subject language. I extend the game-theoretical account proposed by Clark and Parikh (2006), which uses games of partial information to account for anaphora production and resolution in English. Their proposal needs to be extended to account for the fact that null-subject languages make extensive use of null pronouns in subject position, apart from overt pronouns, proper names and definite descriptions. The main idea is that speaker and hearer choose the most efficient strategy with the highest payoff, given the type of entity they want to refer to (more or less salient) and the kind of information they want to convey. Game theory provides an elegant framework to capture anaphora production and resolution, by giving payoffs to each option and probabilities to information states. However, anaphora production and resolution in Catalan constitute a challenge for this type of approach. While this approach elegantly captures the use of an overt pronoun to mark a special information structure, its use in intrasentential anaphora situations is more problematic to model.

## 1. Introduction

In Catalan, a pronoun in subject position can be either overt or phonologically null. The goal of this paper is to model the distribution of overt and null pronouns within a game-theoretical framework, extending the account proposed for English by Clark and Parikh (2006) and discussing whether game theory makes the right predictions.

This paper is structured as follows. Section 2 gives an overview of game theory and its applications to linguistics. Section 3 briefly reviews Clark and Parikh's proposal. Section 4 introduces the pronominal system of null-subject languages. Section 5 analyzes some cases of intrasentential anaphora in connection with the choice and interpretation of referring expression. Section 6 deals with the situations in which overt pronouns are mandatory as markers of a particular information structure. Section 7 concludes.

## 2. Game theory and linguistics

Game theory (GT) is the study of mathematical models of conflict and cooperation between intelligent rational decision-makers (Myerson, 1991). In linguistics, GT has mainly been used in semantics and pragmatics, since it provides a good framework to explain why speakers and hearers (that is, rational agents) choose a certain action, i.e. why they utter a

[1] Thanks to Robin Clark, Prashant Parikh and Josh Tauberer for many helpful comments and to Lucas Champollion for many interesting observations and for letting me use his ideas of how to draw a game tree.

sentence or interpret a sentence with a particular meaning in a particular context. Specifically, game theory has been applied to derive the semantics of questions, quantifiers scope, discourse anaphora (Clark and Parikh, in preparation) and implicatures (Parikh, 2000).

The basic idea of GT is that an agent must often decide among several possible actions; in such cases, he will obviously choose the one he prefers. It is possible to translate this preference by giving a numerical value (or payoff) to each option. The option with the highest number will be the preferred action. However, sometimes payoffs are uncertain, so that every possible outcome has a certain probability associated with it. In that case, the agent might choose the action with the highest expected payoff or he might prefer the action with the minimum risk. When there is more than one agent making decisions, the action one agent decides to do might affect the other agents' payoffs and, thus, the other agents' decisions. That is, one agent needs to consider the other agents' actions and payoffs in order to choose his best option. In this sense, there is a strategic interaction among all rational agents. When no agent has an incentive to change his action (given all others agents' actions), an equilibrium (called Nash equilibrium) is reached and the game is solved. In a given game, there may be several Nash equilibria; the one(s) with the highest payoffs for both agents is called Pareto-Nash equilibrium.

A game of partial information is that in which, at some point, one of the agents does not know which state he is in. The usefulness of games of partial information for linguistics is straightforward (Parikh, 2000). Speaker and hearer are rational agents; the speaker is trying to convey some information by uttering a proposition (among the several possible propositions she could utter), and the hearer is trying to correctly interpret this proposition (among several interpretations, given the fact that language is ambiguous). Both agents are trying to minimize production and processing costs (by, for example, avoiding unambiguous but extremely long sentences), while communicating successfully.

**3. Clark and Parikh's approach**

Consider the simple text in (1):

1. A cop saw a hoodlum. He yawned.

There are several issues regarding the choice of referring expression in this small text, both from the speaker and hearer point of view. Namely, how does the hearer (B) know who *he* refers to? Why does the speaker (A) choose to utter *he* instead of a definite description (DD)? Clark and Parikh (2006) view this problem as a game of partial information in which A and B share some knowledge and in which they try to find the most efficient strategy to solve the game of communicating the utterance. On the one hand, A uses particular discourse anaphors when she expects B to be able to correctly identify the referent. On the other hand, B chooses antecedents to discourse entities based on how he expects A to refer to each entity. As both agents, speaker and hearer, are aware of this fact (they both know they are playing this game), they can find a maximally efficient solution, that is they can compute a Pareto-Nash equilibrium as a solution for the game. This equilibrium is maximally efficient in the sense that for A it is the best way to encode the meaning that she wants convey and, for B, given the form he has heard, it is the best way to interpret the referring expression.

The game tree in Figure 1 shows the moves of A and B and the payoffs they get in each situation for each option. There are two trees for each information state s and s', with probabilities p and p'. The tree rooted in s is the one in which A intends to refer to the cop

(Subj, henceforth), and s' is the one in which A intends to refer to the hoodlum (Obj, henceforth). For each tree, A may use a pronoun or a DD. If a pronoun is used, B may resolve the anaphor correctly or may make a mistake. This is precisely what makes this game a game of partial information: when hearing a pronoun, B does not know whether he is in t or in t' (this is indicated by circling both states).

In s, if A uses a DD, B will surely resolve the anaphor correctly. However, the payoffs will not be very high due to production and processing costs (DDs are longer and syntactically more complex than pronouns) and due to the assumption that referring to a prominent element (the subject) with a full description rather than a pronoun entails some cost. In the example, these assumptions are modeled by payoffs of (6)[2]. If A uses a pronoun in s and B correctly resolves the anaphor, the payoffs are higher (10), since the costs are much less. However, if B interprets Obj instead of Subj, the payoffs would be negative (-10) and would lead to an undesirable situation. In s', the situation is very similar. However, if a DD is used, the payoffs are (7) (and not (6) as in s), which represents the fact that using a DD for a less prominent entity (the object) is assumed to be less costly. Also, if A chooses a pronoun and B correctly chooses Obj, the payoffs are (8) and not (10), because it is less efficient to pronominalize a less prominent element. If A chooses a pronoun and B incorrectly chooses Subj, the payoffs are again negative (-10).

To summarize, payoff are assigned following this two principles:
1. Longer expressions are most costly, while pronouns are less expensive
2. It is cheaper to refer to more salient entities with pronouns, and to less salient entities with DDs.

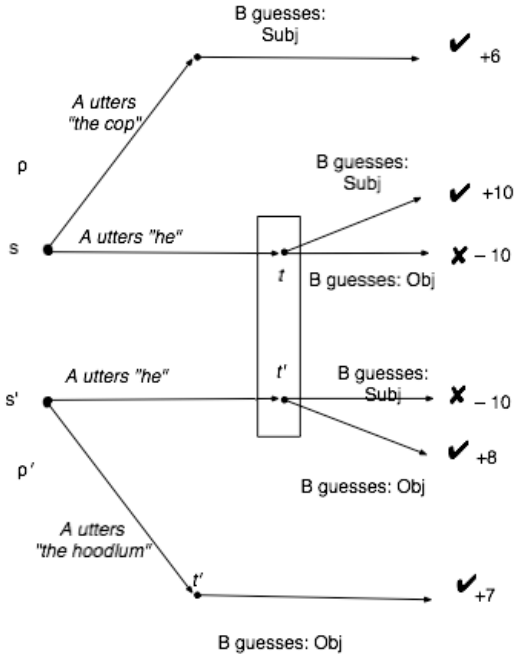


**Figure 1**

[2] They assume payoffs are equal for speaker and hearer.

In absence of further information, they assume that the two information states are equally likely, that is p = p' = 0.5. In that case, there are two Nash equilibria. One corresponds to A uttering *he* in s, uttering *the hoodlum* in s' and B choosing Subj whenever he's in {t, t'}. This equilibrium has an expected payoff of p·10 + p'·7 = 8.5. The other equilibrium corresponds to A uttering *the cop* in s, *he* in s' and B choosing Obj. This equilibrium has an expected payoff of p·6 + p'·8 = 7. Since the payoff of the first equilibrium is higher, it is the unique Pareto-Nash equilibrium and it is the strategy A and B will follow.

Clark and Parikh also show how this account can deal with apparent counterexamples. The basic idea is that several factors can influence the probabilities, so that one of the information states becomes more likely. For instance, note the following contrast:

2. a. John called Bill a Republican. Then he insulted him.
   b. John called Bill a Republican. Then HE insulted him.

The partial game just presented correctly predicts that the pronoun in (a) should refer to John. The game for (b) should be identical to the game for (a). However, the contrastive stress in the pronoun has the effect of affecting the probabilities, so that p' > p; that is, it becomes more likely that the speaker wants to refer to the object of the previous sentence.

**4. The pronominal system of null subject languages**

Catalan, as other Romance languages including Italian or Spanish, is a null-subject language and has a double system of pronouns: (a) strong lexical pronouns and (b) clitic and empty pronouns.

| 3 | a. Ell t' estima a tu | b. [∅] T' estima |
|---|---|---|
| | he obj-cl loves to you | obj-cl loves |
| | `He loves you' | `He loves you' |

In subject position, there is an alternation between overt and null pronouns (*ell* in 3a, null in 3b). From a pragmatic point of view, weak pronouns are the unmarked form in Catalan (Vallduví, 2002). It has been suggested that overt pronouns are used when, for some reason, it is difficult to access the referent they denote. For instance, one such case would be when there are two possible antecedents for the pronoun. The use of overt pronoun could help to select the less accessible antecedent. We will explore some of these cases in section 5. However, Vallduví (2002) also notes that in some cases the use of overt pronouns is mandatory. Constructions in which a stressed form is needed (clefts, answers to a wh-question, comparative constructions, focus constructions, constructions with an elliptical verb) will require the use of an overt pronoun, since a null pronoun cannot, of course, be stressed. Such cases will be examined in section 6.

**5. Intrasentential anaphor in Catalan**

Consider the following mini-discourses:

4. a. El Joan va pegar el Pere. [∅] Està enfadat
      The Joan hit-PAST the Peter. Is angry
   b. El Joan va pegar el Pere. Ell està enfadat.
      The Joan hit- PAST the Peter. He is angry
   c. El Joan va pegar el Pere. El Pere/?? El Joan està enfadat
      The Joan hit- PAST the Peter. The John/The Peter is angry

Alonso-Ovalle et. al. (2002) used the Spanish version of 4a and 4b in an experimental setting. They asked subjects the question "Who is angry?". When presented with the text in (4a), the subject of the first sentence was chosen as antecedent for the null pronoun in 73.2%

of the answers. When presented with (4b), the percentage of responses choosing the subject of the first sentences dropped to 50.2%, while the percentage of responses choosing the object of the first sentence increased considerably.

I reproduced the same experiment with thirteen Catalan speakers and the results were fairly similar to the Spanish experiment. For the null pronoun, in 68.75% of the sentences, the speakers said the pronoun referred to the subject, in 18.75%, to the object and in 12.5% they judged the pronoun to be ambiguous. In contrast, with the overt pronoun, 50% of the time the speakers said that the pronoun was referring to the subject and 50% to the object.

From the Spanish and Catalan experiments, we see the following generalizations:

- The null subject of (4a) is mostly interpreted as unambiguously referring to the subject of the previous sentence, Joan. Being interpreted as the object of the previous sentence is not impossible, but much less likely.
- The facts of (4a) sharply contrast with the situation in (4b). Experimental data from Spanish and Catalan indicates that both the subject and the object of the sentence are equally likely.

As for (4c), the felicity of using a proper noun (or a DD) depends on the grammatical function of the phrase referring to this entity in the previous sentence. Using a proper noun to refer to a previous subject (the most salient entity of the sentence) is highly infelicitous. Using a proper noun to refer to the object of the previous sentence (which is less salient) is perfectly felicitous.

Modeling these facts with game theory is a challenging task. The game between hearer and speaker to resolve the anaphor, shown in Figure 2, looks similar to the game for the English discourse in Figure 1, but the complexity has increased because A now has three choices, instead of two: she can use an overt pronoun, a null pronoun or a proper noun. B has to decide whether A wants to refer to Joan, Subj, or to Pere, Obj.

Let's first assume that A wants to refer to Subj (the actual information state is s). In that case, if A uses a proper noun, B will surely interpret it correctly, but the payoffs will not be high because of the production and processing costs and because it is not efficient to use a proper noun to refer to the subject. We model this with payoffs of (5). If A uses an overt pronoun and B correctly interprets Subj as the referent, the payoffs will be higher than in the case of the proper noun, but not maximally high (not (10)) because there is a better strategy to refer to a maximally salient entity. This strategy is the use of a null pronoun: it is the most efficient and less costly strategy to refer to maximally salient entities and, thus, the payoffs are (10). When using a pronoun (overt or null), there is the chance that B makes a mistake and incorrectly interprets Obj instead of Subj. In such cases, the payoffs are negative (-10).

Let's now assume A wants to refer to Obj. If B incorrectly interprets Subj in the case of both pronouns, the payoffs are again negative (-10). If A uses a null pronoun, the payoffs will be (9), and not (10) as in s, since it is not as efficient the use a null pronoun to refer to a non-maximally salient entity as to refer to a maximally salient one. The use of a proper noun yields a better payoff than in s, (7) versus (5), because we are referring to a less salient entity.

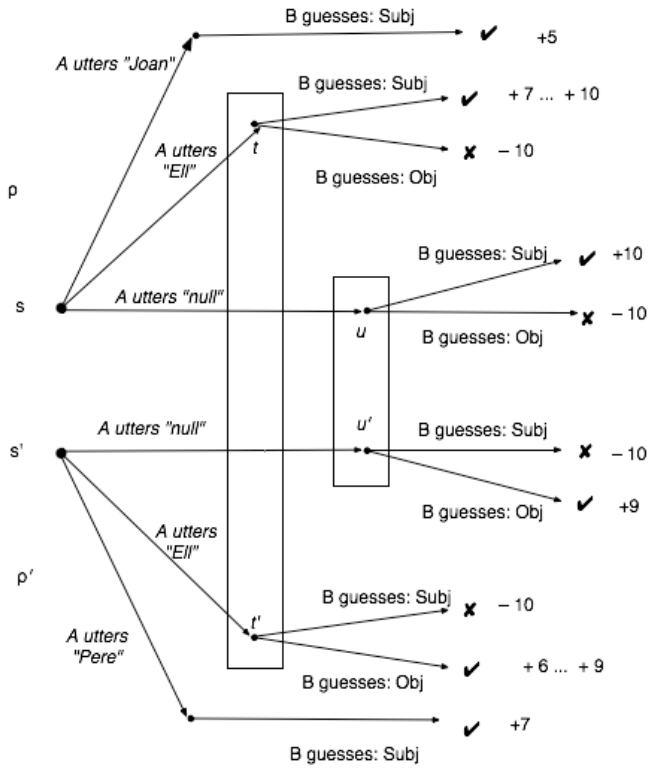


**Figure 2**

The difficult task is to assign a payoff to the overt pronoun. As mentioned, the use of an overt pronoun is considered ambiguous. This suggests that the payoffs of using an overt pronoun should be the same in s and s'. However, no payoff for the overt pronoun will fully capture the data we want to model:

- If the payoffs consist of any fixed value greater than 7, the Pareto-Nash equilibrium will predict that A should choose the null pronoun in s, utter the overt pronoun in s' and B choose Subj whenever he's in {t, t'} and the Obj consistently whenever he's in {u, u'}. The intuitive reasoning behind this Pareto-Nash equilibrium is the following: the hearer knows that the speaker will always use a null pronoun to refer to the subject and, therefore, is able to conclude that the speaker will only use an overt pronoun to refer to the object.
- If the payoffs consist of any fixed value smaller than 7, B is predicted to choose a proper noun to refer to Obj, since it will have a higher payoff than the overt pronoun.

Therefore, neither option is consistent with the data I want to model. This suggests that in order to model the ambiguity of the overt pronoun we need its payoffs to be indeterminate and to fall within a range. To maintain the Subj-Obj asymmetry, the payoff for the subject should be in the range between 7 and 10 and the payoff for the object between 6 and 9. Modeling the overt pronoun's payoff as a range and not a fixed value can capture a number of things:

1. When hearing an overt pronoun, B will not be able to choose a pure strategy, because, whenever the payoff for the overt pronoun is smaller than 7, there are two

Pareto-Nash equilibria: one in which B has to interpret the pronoun referring to the Subj and the other to the Obj. Therefore he will have to choose randomly between the two options.

2. Referring to the Subj with a proper noun is highly infelicitous. Our model predicts exactly this, since there are at least two more efficient alternative strategies: the most efficient and unambiguous use of a null pronoun and the ambiguous use of an overt pronoun. In contrast, the use of a proper noun to refer to the Obj of the previous sentence is fully acceptable. Given that the overt pronoun strategy was ambiguous and its payoffs indeterminate, the proper noun strategy is the most efficient unambiguous strategy.

3. A change of probabilities, by biasing towards one of the information states, should be able to affect the choice of pronouns. Namely, when s' is more likely than s, the speaker will be able to use the null pronoun to refer to Obj. This biasing can be caused by lexical semantics, by choice of discourse markers, etc. For example, in (5) the null pronoun unambiguously refers to the previous object. The biasing comes from the lexical semantics of the verbs (if X hits Y, Y is likely to spend time in hospital and not the other way round) and from the consequence connective "per això" '*therefore'*.

   5. El Pere va apallissar el Joan. Per això, va haver-se d' estar-se un mes a l' hospital
      The Peter beat-pst the John. So, had-pst of stay-pst one month in the hospital
      "Peter beat John. So, he had to stay one month in the hospital"

An obvious question to ask is why should the overt pronoun be associated with a range in its payoffs. Although I don't have a conclusive answer, I would like to suggest that it is related to the fact that the overt pronoun covers a middle ground in a scale of efficiency. The null pronoun is clearly the most efficient way of referring to maximally salient entity and a proper noun is the less efficient, but more safe way. The overt pronoun falls in the middle and it has a mixed status: it is a pronoun, but, unlike most pronouns, is not the most economical way of referring to some entity. A related question would be how this mixed status of the overt pronoun affects the choice of a speaker. The data I have been examining has to do with interpretation and not with production. A plausible consequence of the indeterminacy of overt pronouns would be that speakers will avoid them unless they have an extra reason to use them. Section 6 deals with what these extra reasons may be.

To conclude this section, the Catalan data presented is a challenging case for game theory:

- The use of a null pronoun unambiguously refers to the subject of the previous sentence and these facts are translated in high payoffs for this strategy. Reference to the previous object only arises in case of biasing.
- The use of a proper noun is unacceptable for the subject of the previous sentence and acceptable for the object. Both facts are predicted by our model
- The use of an overt pronoun seems to be ambiguous for the hearer as for the referent choice. We have modeled this by proposing that the payoffs for the overt pronoun are not fixed, but fall within a range. The hearer when trying to interpret an overt pronoun will not be able to choose a pure strategy and will have to randomly choose between Obj and Subj.

## 6. Overt pronouns as markers of a particular information structure.

In this section, I discuss situations in which an overt pronoun is obligatory in subject position in Catalan, in spite of being the marked option and of there being always the hypothetical alternative of using a null pronoun. Overt pronouns in Catalan can be mandatory even when there is no ambiguity as to which entity is its referent. Catalan is a language with rich verbal inflection that indicates the grammatical person of the subject of the sentence. For instance, in (6) the verb (*nego*, 'deny') is unambiguously first person singular. Therefore, the pronoun does not play any disambiguation role. However, using a null pronoun is not an option and the discourse would be unfelicitous if such a pronoun was used. Why is it so? All examples in which this is the case have a special informational structure, which involve focusing or contrast and which require a segment that may be stressed. Obviously, a null pronoun cannot be stressed and, therefore, an overt pronoun must be used.

> 6. "Hom ha arribat a afirmar que Bernat Mies és el conegut pintor berguedà Joseph Maria de Martín. **Jo** ho nego rodonament [..]"
> "One has even claimed that Bernat Mies is the famous painter Joseph Maria de Martin. **I** deny it vigorously

In (6), there is a clear contrast between what people claim and what the speaker claims. Note that in English, a special focus accent would be used on the pronoun. In Catalan, actually pronouncing the pronoun is enough to convey the contrast.

We can model this situation using partial games[3]. For B, the ambiguity is not between two possible antecedents for the pronoun, but between two possible interpretations: the literal interpretation (*l*), in which the pronoun just refers to some discourse antecedent, and the informational-structure interpretation (*i.s.*), in which apart from the referential use of the pronoun it also conveys some other information (contrast, focusing, etc.). That interpretation *i.s.* contains more information that *l*, and this fact will be encode in the payoffs.

Figure 3 captures all the relevant facts. The two initial nodes s and s' represent two different situations in which A wants to convey either *i.s.* is *l* respectively. B, when hearing the pronoun (overt or null), will not be sure in which state he is in: t or t' in the case of an overt pronoun, q or q' in the case of a null pronoun. In addition, we also need to consider other propositions that A might have said but chose not to. In s, A could have said a proposition *ex* in which the information structure is explicit (for instance, "In contrast, I"). In s', where the speaker does not want to convey any extra information structure, A could have chosen to remain silent, proposition υ.

As for the payoffs, in s, if A uses an expression that explicitly conveys the information structure, B will correctly interpret it, but the payoffs will not be very high due to processing and production costs. If A uses a null pronoun in s, the payoffs will be negative (-10), since a null pronoun cannot express this information structure meaning and the hearer will surely misinterpret. If A uses an overt pronoun, B can interpret *l* (and, therefore, the payoffs will also be negative) or can correctly interpret *i.s.*. In that, case the payoffs will be (8), since there are some processing costs but less than in the case of the explicit proposition.

[3] My analysis follows very closely Parikh's (2000) analysis of implicatures.

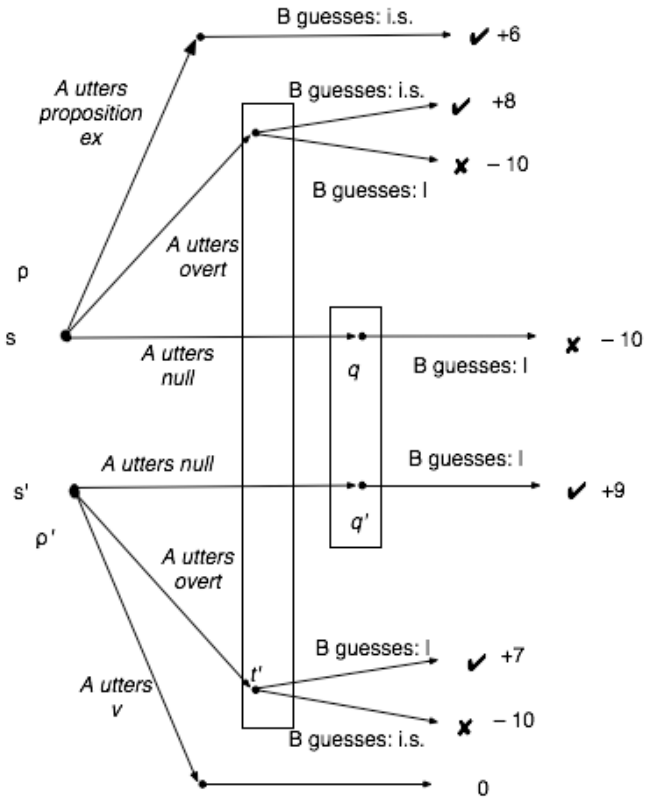


**Figure 3**

In s', if the null pronoun is used, B will correctly interpret *l* and the payoffs will be positive. If an overt pronoun is used, B may incorrectly interpret *i.s.* and the payoffs will be negative or the hearer may successfully interpret *p*, in which case the payoffs will be positive. However, the payoffs will be lower than the same situation in s, since in s' the information conveyed is less and, therefore, has less value. Thus, we model this by assigning payoffs of (7). Finally, probabilities will be assigned to the initial nodes, s and s'. In this case, we can assume both information states have equal probabilities.

There is a unique Pareto-Nash equilibrium with an expected payoff of 8.5 ($p \cdot 8 + p' \cdot 9$), which is the following: {(s, overt), (s', null); ({t, t'}, i.s.), ({q, q'}, l)}. In information state s, A should utter an overt pronoun; in information state s', a null pronoun. In the information set {t, t'}, A should interpret *i.s.* (the literal interpretation of a pronoun plus the extra information structure meaning) and in the information set {q, q'}, B should interpret just the literal meaning *l*. That is, when a null pronoun is used, no information structure meaning is conveyed. In contrast, using an overt pronoun is a possible way to convey such information structure. Therefore, even if the absence of ambiguity or biasing, an overt pronoun will be used if the speaker needs to convey meaning related to contrast or focusing.

In this case, the game-theoretical approach provides an elegant account of why a speaker should use an overt pronoun in absence of ambiguity if she wants to convey a special information structure.

## 7. Conclusion

In this paper, I have examined whether the game-theoretical approach to pronoun choice and anaphora resolution that Clark and Parikh (2006) have proposed for English can be extended to a null-subject language like Catalan.

Catalan data for intrasentential anaphora pose a challenge for this game-theoretical account specially because of the behavior of an overt pronoun, which speakers judged could equally refer to the previous subject or the previous object. This cannot be modeled be assigning a fixed payoff to the overt pronoun. In contrast, assigning an indeterminate payoff within a range could account for the empirical data. In contrast, game theory can very easily and elegantly account of why overt pronouns are mandatory in cases where the speakers want to convey a particular information structure.

## 8. References

**Alonso-Ovalle, Luis; Charles Clifton, Lyn Frazier and Susana Fernández Solera (**2002). Null vs. Overt Pronouns and the Topic-Focus Articulation in Spanish. *Journal of Italian Linguistics*, 14:2.

**Clark, Robin and Prashant Parikh** (2006). "Game Theory and Discourse Anaphora". Submitted to the *Journal of Logic, Language and Information*.

**Myerson, Robin** (1991). *Game Theory. Analysis of Conflict.* Cambridge: Harvard University Press.

**Prashant Parikh** (2000). *Language in Use*. CSLI Publications, Standford, CA.

**Sally, David** (2003). "Risky speech: behavioral game theory and pragmatics" *Journal of Pragmatics*, Volume 35, Number 8, pp. 1223-1245(23).

**Vallduví, Enric** (2002). "L'oració com a unitat informativa". Dins: Solà, Joan et al. (ed.). *Gramàtica del català contemporani*. Barcelona: Empúries. Cap. 4.