



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Pronouns in Catalan: Games of partial information and the use of linguistic resources

Laia Mayol ^{*}, Robin Clark

619 Williams Hall, University of Pennsylvania, Philadelphia, PA 19104, USA

Received 28 June 2008; received in revised form 5 July 2009; accepted 6 July 2009

Abstract

This paper investigates the variation between null and overt subject pronouns in Catalan, a null-subject language. We account for this variation in game-theoretical terms: that is, we analyze the distribution of both overt and null pronouns as a result of the strategic interaction between participants in a communicative exchange.

First, we examine the Position of Antecedent Hypothesis (PAH), as put forward by Carminati (2002). This hypothesis proposes that null and overt pronouns have different biases: null pronouns prefer antecedents in subject positions, while overt pronouns prefer antecedents in non-subject positions. Carminati (2002) tested the PAH for Italian in a variety of intrasentential contexts. In this paper, we show experimentally that the PAH also holds for Catalan even in across-sentence contexts. In the second place, we also show how the PAH can be naturally redefined as a game of partial information, in which speaker and hearer are trying to communicate. This redefinition does not just translate the PAH into a different notation, but it extends the PAH into a model that makes more accurate predictions, since it can account also for the cases in which the biases predicted by the PAH are not obeyed. © 2009 Elsevier B.V. All rights reserved.

Keywords: Anaphora; Game theory; Null-subject languages; Catalan

1. Introduction

A language like English, which lacks phonologically null pronouns, has essentially two ways of evoking discourse-old/hearer-old objects. It can use either a pronoun, as illustrated in 1a or it can use a description, as illustrated in 1b.

- (1) a. A cop was chasing a hoodlum. He stumbled.
b. A cop was chasing a hoodlum. The hoodlum stumbled.

In 1a, the pronoun in the second sentence is most naturally interpreted as denoting the discourse entity introduced by *a cop* in the first sentence. In 1b, a definite description is used to denote the discourse entity introduced by *a hoodlum*. Clark and Parikh (2007) have developed a game-theoretic treatment of discourse anaphors for English which distinguishes between the use of pronouns and definite descriptions on the basis of relative salience.

Languages which allow phonologically null subjects, however, have an additional resource that can be deployed in the service of discourse anaphora. In this paper we investigate the factors that trigger the appearance of null and overt

^{*} Corresponding author.

E-mail addresses: laia@ling.upenn.edu (L. Mayol), rclark@ling.upenn.edu (R. Clark).

pronouns in subject position in Catalan. Following Clark and Parikh (2007), we model the variation between null and overt pronouns as a game-theoretical problem, in which there are, in principle, several resources competing for the same position.

Catalan, as other Romance languages like Italian or Spanish, is a null-subject language and has a double system of pronouns (Rigau, 1986). In subject position, there is an alternation between overt pronouns (*ell* in 2a) and null pronouns (in 2b).¹

- (2) a. Ell és alt.
he is tall
b. Ø És alt.
is tall
“He is tall.”

We will treat the problem of whether to use a null or overt pronoun or a definite description as a strategic problem. Speakers and hearers deploy these resources as a means of achieving efficient communication. In particular, null pronouns provide an extra resource that can be strategically deployed in the service of a more efficient way of referring to a linguistic antecedent.

This paper is concerned with those contexts in which there is, in principle, a choice between two or more forms; we will ignore the contexts in which one of the forms is compulsory.² For example, overt subject pronouns are required when the subject conveys focal information (Vallduví, 1992) or contrast (Luján, 1985).

This paper is structured as follows: Section 2 presents the Position of Antecedent Hypothesis and two psycholinguistic experiments testing this hypothesis. Section 3 introduces games of partial information and uses them to analyze discourse anaphora, with an emphasis on Catalan. Also, Experiment 3, which provides further evidence in favor of a game-theoretical analysis, is presented. Section 4 concludes and discusses some open research questions.

2. The Position of Antecedent Hypothesis

Carminati (2002) studied the distribution of overt and null pronouns in Italian when used as intrasentential anaphors. She proposed that this variation is regulated by the Position of Antecedent Hypothesis (PAH), as defined in 3.

- (3) Position of Antecedent Hypothesis: null pronouns prefer to retrieve an antecedent in the (highest) Spec IP, whereas overt pronouns prefer an antecedent in a lower syntactic position.

Thus, the PAH proposes that null and overt pronouns have different antecedent biases: they prefer to retrieve antecedents in different syntactic positions.

This hypothesis is in accordance with Accessibility Theory (Ariel, 2001), which claims that more marked, informative forms tend to retrieve less salient antecedents, while unmarked, less informative forms tend to retrieve more salient antecedents. Similar proposals for other null-subject languages have been made using the machinery of Centering Theory: such proposals claim that overt pronouns are used to mark that their antecedent is not the maximally salient referent or that the transition between sequences is not the default one. For example, see DiEugenio (1998) for Italian, Dimitriadis (1996) for Greek, Kameyama (1985) for Japanese and Turan (1995) for Turkish. In addition, switch reference has been identified as one of the main constraints regulating the null/overt pronoun variation in the variationist literature (see Cameron (1992) and Silva-Corvalán (1977) for sociolinguistic studies of the phenomenon in Spanish).

Carminati (2002) tested the PAH for Italian in a series of off-line and on-line experiments investigating a variety of antecedents that are standardly assumed to occupy the subject position in the syntactic structure. Overall, her findings supported the Position of Antecedent Hypothesis, as opposed to other hypotheses, such as hypotheses based on an economy principle (favoring the null pronoun generally), or those based on avoidance of ambiguity (favoring the overt pronoun, because it can disambiguate some cases, given that it carries gender information). Her findings also showed that the antecedent bias of the overt pronoun is less stable and more context-dependent than that of the null pronoun. Carminati (2002) used different methods in her research: self-paced reading tasks, questionnaires, and correction tasks.

¹ The empty set is used to represent the null pronoun in the examples.

² See Cameron (1992) for a good overview of the contexts in which one of the two pronominal forms is compulsory.

It is important to note that Carminati is concerned with intrasentential anaphora and her hypothesis is that this kind of anaphora has access to the syntactic representation; this might not be the case for intersentential anaphora. Since our interest is in discourse anaphora, we cannot take Carminati's results at face value, but they should be tested in across-sentence contexts. This is the goal of Experiment 1 and Experiment 2. Experiment 1 is a questionnaire study and Experiment 2 is a self-paced reading study. Both experiments test the PAH in across-sentence contexts in Catalan and, since similar experiments have been carried out for Italian and Spanish, the results can be compared.

2.1. Experiment 1

This experiment tests the PAH with regard to two-sentence discourses. The first sentence introduces two individuals by means of two proper names of the same grammatical gender, one in subject position and the other in an object position. The second sentence, which contains either a null or an overt pronoun, is not pragmatically biased and, in principle, can refer both to the previous subject or the previous object. In order to test the PAH, we did a questionnaire study, in which, after reading the sentence with a null or an overt pronoun (such as 4a and 4b, respectively), subjects had to choose their preferred interpretation for the pronoun, by choosing a paraphrase among the ones in 4c.

Materials. The materials consisted of 16 two-sentence discourses with two conditions (Null or Overt pronoun in the second sentence) and paraphrases for the ambiguous sentences.

- (4) a. Condition 1: Null Pronoun
La Marta escrivia sovint a la Raquel. Ø Vivia als Estats Units.
“Marta wrote frequently to Raquel. Ø Lived in the United States.”
- b. Condition 2: Overt Pronoun
La Marta escrivia sovint a la Raquel. Ella vivia als Estats Units.
“Marta wrote frequently to Raquel. She lived in the United States.”
- c. Paraphrases:
i. La Marta vivia als Estats Units.
“Marta lived in the United States.”
ii. La Raquel vivia als Estats Units.
“Raquel lived in the United States.”

The conditions for each item set were counterbalanced and incorporated into a questionnaire experiment together with 24 filler items and 5 practice items. Four counterbalanced lists were constructed (the last two lists with the items in reverse order), with a single randomization for all lists.

Procedure. The experiment was administered using a laptop, equipped with EPrime software. Before starting the experimental session proper, participants read a set of written instructions, which explained the experimental procedure. Participants went through a practice session, so that they could get familiar with the keyboard and the procedure, and the experiment subsequently began. The discourses were presented on the computer screen. Participants were asked to indicate which interpretation of the second sentence they preferred, i.e., whether they thought it was a statement about the subject of the first sentence, or the object of the first sentence. That is, they had to choose between paraphrases such as the ones in 4c for the example above.

Participants. Thirty-two students from Universitat Pompeu Fabra in Barcelona took part in this experiment.

Predictions. If the PAH holds for Catalan in across-sentence contexts, we would expect that the subject antecedent will be preferred for Condition 1 while the object antecedent will be preferred for Condition 2.

Results. The results in raw percentages can be seen in Table 1.

Table 1
Results for Experiment 1.

	Subject antecedent	Object antecedent
Condition 1 (null)	70.3	29.7
Condition 2 (overt)	35.5	64.5

Table 2

Results for Experiment 2 in Carminati (2002).

	Subject antecedent	Object antecedent
Null pronoun	80.7	19.3
Overt pronoun	16.7	83.3

Table 3

Results for Experiment 1 in Alonso-Ovalle et al. (2002).

	Subject antecedent	Object antecedent
Null pronoun	73.2	26.8
Overt pronoun	50.2	49.8

There was quite a strong preference to interpret null pronouns as having subject antecedents and overt pronouns as having object antecedents. An ANOVA test of the frequency with which the subject antecedent was chosen in the null vs. overt pronoun conditions was performed with subjects and items as random effects. The difference of choosing the subject antecedent (70% vs. 35%) was significant ($F(1, 31) = 64.23$, $p < 0.001$; $F(1, 15) = 26.153$, $p < 0.001$). The difference between the preferred antecedent choices for the two conditions (70% vs. 64%) was not significant ($F(1, 31) = 1.5573$, $p = 0.22$; $F(1, 15) = 0.353$, $p = 0.56$).

2.1.1. Comparison with Italian and Spanish data

A very similar version of this experiment was originally developed by Carminati (2002) to test the PAH in Italian and later replicated by Alonso-Ovalle et al. (2002) for Spanish. An important difference is that the former used within-sentence contexts, while the latter used across-sentence contexts, as do the experiments here for Catalan.

Table 2 shows the results for Experiment 2 in Carminati (2002).

The results clearly give support to the PAH and the interaction between type of antecedent and type of pronoun is even stronger than the one we report for Catalan.

Table 3 reports the results in Alonso-Ovalle et al. (2002) for Spanish.

Alonso-Ovalle et al. (2002) claim that these results support the PAH in Spanish across-sentences. While it is true that there is still a correlation between antecedent choice and type of pronoun, there are some intriguing differences regarding the overt pronoun in the Italian experiments, on the one hand, and the Spanish experiments, on the other hand: in Italian, the overt pronoun was interpreted as referring to the previous subject in only 16.67% of the cases, while in Spanish it was 50.2%. Thus, the Spanish experiments seem to indicate that, while the null pronoun is clearly biased towards the previous subject, the overt pronoun does not show a clear preference. Although the PAH seems to be in effect also for Spanish in intersentential cases, its effects seem to be milder. In principle, this could be due to the change of language (Italian vs. Spanish) or to the change of type of anaphora tested (intersentential vs. intrasentential). Given that for the Catalan experiments, we also used across-sentence contexts, the Spanish results cannot be just a consequence of the change in the context. More research is needed in determining whether Spanish is qualitatively different or not from Catalan and Italian. For instance, it would be necessary to perform online studies for Spanish to be able to assess the robustness of Alonso-Ovalle et al. (2002) results, such as the one we present in the next section for Catalan.

2.2. Experiment 2

Experiment 2 is a self-paced reading study that aims to test the Position of Antecedent Hypothesis in two-sentence discourses with semantic bias. This kind of design has the advantage of not assuming that the sentences are neutral³ (that is, pragmatically non-biased) and that participants are less aware of the goal of the experiment, since their judgments are not explicitly asked for.

³ We consider dropping the assumption of neutrality to be an advantage in the sense that the neutrality of a sentence is much more subjective and disputable than its non-neutrality.

It has been shown widely in the psycholinguistics literature that readers make use of all the available linguistic cues to arrive at a coherent interpretation. If they encounter explicit information which goes against some of the linguistic cues they had encountered before, this does not have to result in an unacceptable or anomalous sentence, but people do need more time to read the sentence (see Caramazza et al., 1977). For instance, Koornneef and Van Berkum (2006) investigated some verbs conveying implicit causality, such as *apologize*, which has a bias towards a continuation that makes reference to the first NP. They tested sentences such as the ones in example 5 in which the pronoun was either consistent or inconsistent with the bias of the verb. Although both discourses are coherent, they found that bias-consistent sentences were read faster than bias-inconsistent ones.

- (5) a. Bias-consistent pronoun.
Linda and David had an accident. David apologized to Linda because he was the one to blame.
b. Bias-inconsistent pronoun.
Linda and David had an accident. Linda apologized to David because he was not the one to blame.

These results suggest that sentences which obey the biases predicted by the PAH should be read faster than sentences which do not obey them. The goal of experiment 2 is to test this claim.

Materials. The materials consisted of 16 two-sentence discourses with four conditions. In these discourses, the first sentence introduces two individuals by means of two proper names of the same grammatical gender, one in subject position, the other in either object or indirect object position. The second sentence contains either an overt or a null pronoun and is semantically biased so that the pronoun is more likely interpreted as referring either to the previous subject or previous object. The four conditions can be seen in 6.

- (6) a. Condition 1: Null pronoun + bias towards subject antecedent.
El Joan va deixar en ridícul el Dani davant de tothom. Ø Es va excusar repetidament.
“John made fun of Dani in front of everyone. Ø Apologized many times.”
b. Condition 2: Overt pronoun + bias towards subject antecedent.
El Joan va deixar en ridícul el Dani davant de tothom. Ell es va excusar repetidament.
“John made fun of Dani in front of everyone. He apologized many times.”
c. Condition 3: Null pronoun + bias towards object antecedent.
El Joan va deixar en ridícul el Dani davant de tothom. Ø Es va ofendre molt.
“John made fun of Dani in front of everyone. Ø Was very offended.”
d. Condition 4: Overt pronoun + bias towards object antecedent.
El Joan va deixar en ridícul el Dani davant de tothom. Ell es va ofendre molt.
“John made fun of Dani in front of everyone. He was very offended.”

The conditions for each item set were counterbalanced and incorporated into a self-paced reading experiment together with 24 filler items and 5 practice items. Eight counterbalanced lists were constructed (the last four lists with the items in reverse order), with a single randomization for all lists.

Procedure. The experiment was administered using a laptop, equipped with EPrime software. Before starting the experimental session proper, participants read a set of written instructions, which explained the experimental procedure. Participants then went through a practice session, so that they could get familiar with the keyboard and the procedure, and the experiment subsequently began. The discourses were presented on the computer screen. Participants were asked to press the space bar after each sentence and this is how the reading times for each sentence were measured. Comprehension questions, such as the one in 7, probing the resolution of the pronoun, were asked after each item.

- (7) a. Qui es va ofendre?
“Who was offended?”
b. El Joan.
c. El Dani.

Participants. Thirty-two students from Universitat Pompeu Fabra in Barcelona took part in this experiment.

Table 4
Results for Experiment 2.

Condition	Second sentence reading time	Difference (observed – expected)	% correct
Condition 1: subj + null	2464	–24.2	90.5
Condition 2: subj + pron	2928	298.7	90.0
Condition 3: obj + null	2587	45.4	91.1
Condition 4: obj + pron	2700	–1.4	91.2

Predictions. If the PAH holds for Catalan in across-sentence contexts, we would expect that Condition 1 and 4 will be read faster, since they respect the biases the PAH predicts, while Conditions 2 and 3 will be read slower, since they go against the biases predicted by the PAH.

Results. The results for this experiment can be seen in Table 4.

The average reading times for the second sentence were computed, after eliminating times that were longer than 6000 ms and shorter than 200 ms (about 3.5% of the total number of trials). The number in the “% correct” column refers to the percentage of answers in which subjects understood the pronoun as referring to the pragmatically biased antecedent.

For both types of biases (sentences pragmatically biased to the subject and to the object), the sentence with the null pronoun was read faster than the sentence with the overt pronoun. Thus, the data of the raw reading times for the overt pronoun does not follow the pattern predicted by the PAH. In fact, an ANOVA analysis gives pronoun type (null vs. overt) as the only significant factor ($F(1, 31) = 23.86$, $p < 0.001$; $F(1, 15) = 27.66$, $p < 0.001$), while the interaction Pronoun by Bias is not significant ($F(1, 31) = 3.93$, $p = 0.059$; $F(1, 15) = 3.09$, $p = 0.098$).

However, conditions (2) and (4) are systematically longer because they contain the pronoun, and this may be masking the effect of the PAH.⁴ Thus, deviations from regressions were computed to account for length differences. These are the numbers in the Difference column. They were calculated as follows. The predicted reading time for each segment was computed by a regression equation, calculated on a subject by subject basis by correlating the reading time and segment length over all times and conditions of the experiment. Expected times were then calculated for each segment and each subject and subtracted from the observed reading times. This procedure to adjust for length differences is quite standard in the processing literature (see Ferreira and Clifton (1986) for one of the first papers to use it and Trueswell et al. (1994) for some additional discussion of this technique).

The average for each condition of the difference between observed and expected reading times is shown in the Difference column in Table 4. Positive numbers mean that reading times were slower than expected and negative numbers that they were faster than expected. Conditions 1 and 4 were faster than expected, which signals that they were easy to process. In contrast, Condition 2 and 3 were slower than expected, which signals that there was some difficulty to process these conditions. This is, of course, consistent with the PAH.

The data regarding the difference between observed and expected reading times was submitted to an ANOVA analysis. As with the raw reading times, the effect of the type of pronoun was significant ($F(1, 31) = 4.78$, $p = 0.03$; $F(1, 15) = 4.58$, $p < 0.049$). In addition, in this case there was also a significant bias by pronoun interaction ($F(1, 31) = 4.68$, $p = 0.038$; $F(1, 15) = 11.04$, $p < 0.001$). This interaction confirms that sentences with null pronouns are read faster in the subject bias condition, while they are read slower in the object bias conditions; the opposite holds of sentences with overt pronouns.

2.2.1. Comparison with Italian

Carminati (2002) carried out the same experiment we have just described for Italian, intrasententially, using complex sentences, with a subordinate and a main clause. The results can be seen in Table 5, which reports the raw reading times, the corrected times (i.e., difference between observed and expected times) and the percentage of answers in which subjects understood the pronoun as referring to the pragmatically biased antecedent.

The average reading times for the main clause were computed, after eliminating times that were longer than 6000 ms and shorter than 200 ms (about 4% of the total number of trials).

⁴ Although the stimuli were designed so that the length differences would be kept to a minimum for all conditions, the differences did exist. These are the mean lengths (in characters) of the items in each condition: Condition 1 = 59.25; Condition 2 = 62.75; Condition 3 = 59.56; Condition 4 = 63.0.

Table 5
Results for Experiment 2 in Carminati (2002).

Condition	Main clause reading time	Difference (observed – expected)	% correct
Condition 1: subj + null	1844	–162	88.7
Condition 2: subj + pron	2666	499	80.4
Condition 3: obj + null	2352	349	70.4
Condition 4: obj + pron	2236	41	89.1

Reading times were significantly faster for main clauses with null pronouns than clauses with overt pronouns, which is expected because, as we mentioned before, clauses with null pronouns are always shorter than clauses with overt pronouns. The effect of bias was not significant. In addition, there was a significant antecedent by pronoun interaction, both using the raw reading times, unlike the Catalan experiment, and the Difference times, like the Catalan experiments (the ANOVA results for the raw reading times are the following ($F1(1, 39) = 28.16$, $p < 0.001$; $F2(1, 15) = 23.68$, $p < 0.001$)).

Main clauses biased towards a subject antecedent were read faster than sentences biased towards the object, when they contained a null subject pronoun. The opposite was true for sentences with an overt subject pronoun. Again, Carminati's (2002) results are stronger, since her data presents the interaction both for the raw reading times and the corrected times.

2.3. Conclusion

In this section, we have presented some processing evidence that the PAH holds for Catalan intersententially: null pronouns have a subject preference and overt pronouns have a non-subject preference. However, this tendency is milder at discourse level than at sentence level. We point to possible reasons for this difference in section 4.

3. A game-theoretical analysis of discourse anaphora

The use of an anaphor like a pronoun in a discourse involves strategic decision making on the part of the speaker and on the part of the hearer. The speaker must decide whether the hearer will select the discourse object she intends as the target of the pronoun. The hearer must decide which discourse object the speaker intends him to pick out. Each agent's choice is strategically dependent on the other's choice. Game theory provides exactly the mathematics we need to analyze this sort of strategic problem.

3.1. A game-theoretical approach to discourse anaphora

In this section, we will consider the distribution and interpretation of discourse anaphors, specifically overt pronouns in English. The game-theoretical analysis is adapted from Clark and Parikh (2007) and forms the basis for our discussion of overt and non-overt subject pronouns in Catalan. Let us begin with the simple text in 8:

(8) A cop was chasing a hoodlum. He stumbled.

The first sentence in 8 establishes two discourse objects, a cop and a hoodlum. All else being equal, either discourse object is a suitable target. Thus, as noted above, both the speaker and the hearer are faced with a number of decisions. In the second sentence, the speaker must decide whether to use a description—for example, *the cop* or *the hoodlum*—or a pronoun in order to denote the intended discourse object. Given that the speaker has decided to use a pronoun, the hearer must decide which potential target the speaker intends the pronoun to denote.

The various choices faced by the speaker and hearer are summarized in Fig. 1. The root node has two daughters, labeled s_1 and s_2 , corresponding to the speaker's intention to focus on the cop (in the case of s_1) or the hoodlum (in the case of s_2). If the speaker (S, in the figure) is in state s_1 , then she must decide whether to refer to the cop either with the

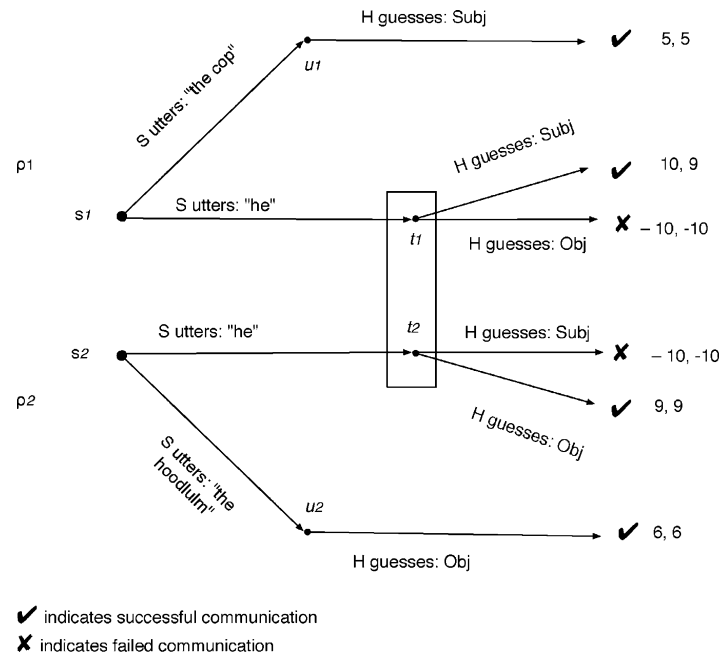


Fig. 1. Game of partial information for English anaphora.

definite description *the cop* or with the pronoun *he*. If the speaker is in state s_2 , then she intends to talk about the hoodlum and must decide whether to use the definite description *the hoodlum* or a pronoun.

If the speaker uses the definite description *the cop* then her⁵ intentions are clear and the hearer's choice is trivial: he will interpret the definite description as referring to the previous subject (Subj, henceforth). The same applies if the speaker uses *the hoodlum*: the hearer (H, in the figure) will interpret it as referring to the previous object (Obj, henceforth). Suppose, however, she uses a pronoun. This places the hearer in the information set $\{t1, t2\}$ in Fig. 1; the information set is indicated by circling the nodes. That is, although the speaker has announced her choice, the choice does not completely disambiguate which discourse entity the speaker intended to denote. This is the distinguishing property of games of partial information; although the speaker has announced her choice, the choice is ambiguous.

Speaker and hearer have preferences about their choices. Clark and Parikh (2007) propose that the following factors figure into the computation of cost for both the speaker and the hearer when producing or interpreting a referring expression. Two of the factors—Encoding Costs and Choice Costs—are explicitly economic since they are based on the cost of production. Aspects of the local context figure in as well.⁶ Finally, a certain value will be associated with successful communication; equally miscommunication will be penalized:

(9) a. Encoding Costs

It is generally more costly to use longer expressions. This costs the speaker in terms of production effort and the hearer in terms of processing effort.

b. Choice Costs

It is generally more costly to use expressions with “high” conventional content; names and descriptions are thus costlier than pronouns. We assume that the cost here involves the effort required to select among a host of possible names and descriptions that could be applied to an object. Here, the speaker and hearer must be engaged in a strategic choice which not only involves some direct costs, but potential risk if the selected expression is misunderstood.

c. Context Factors

It is cheaper to refer to a more prominent element with a pronoun; it is correspondingly more marked (hence, costlier) to refer to a more prominent element with a description or name when a pronoun could be used. Here the speaker and hearer are presumably using some notion of *salience* to guide their choice.

⁵ We use the convention to refer to the speaker with feminine pronouns and to the hearer with masculine pronouns.

⁶ In section 3.2 we eliminate contextual factors from the computation of payoffs and we build them into the probabilities instead.

d. Communication Factors

Successful communication—communication where the speaker transmits the intended message successfully—is highly preferred. Correspondingly, miscommunication—where the hearer misinterprets the speaker's intended message—is strongly penalized.

The factors in 9a and 9b suggest that pronouns should be preferred over other expressions. The factor in 9c suggests that, all else being equal, salient elements, elements that are particularly prominent in the context, will tend to “catch” pronouns, leaving less prominent elements to be associated with definite descriptions.

Payoffs are used to represent preferences. So if an agent prefers action A to action B, the payoff for the first action must be greater than the payoff for the second one, but the particular values assigned to them are not important. That is, payoff are basically indices and their relationship and ordering is what is important, not their particular value. Keeping this in mind, a basic payoff of 10 points is assigned for getting everything right in an optimal fashion. Points are subtracted from this base in accordance with the production costs in 9. Miscommunication is penalized by assigning this outcome –10 points, in accordance with 9d. Both the speaker and the hearer have a shared interest in communicating, although their interests may diverge on other points. In particular, we can assume that speakers will seek to minimize their own production costs while barely taking into account the hearer's interests in the facility of comprehension.

Consider the tree rooted at s_1 . The speaker must decide whether to use the definite description *the cop* or a pronoun to denote Subj. This choice is represented by the branches labeled “S utters: *the cop*” and “S utters: *he*”. Suppose that the speaker selects the pronoun. The hearer must decide, then, whether the speaker intended to signal Subj or Obj. If he picks Subj, both the speaker and the hearer are content. The speaker gets a payoff of 10, since she has used an expression—a pronoun, which has minimal production costs. The hearer gets a payoff of 9; he might have preferred the speaker to be more explicit about whom she meant, but he's content that she used a simple expression. Payoffs given the sequence of moves that players have chosen are represented at the leaves of the game in Fig. 1: the first number represents the payoff for the speaker and the second number for the hearer.

Suppose, on the other hand, that the hearer had selected Obj where the speaker had intended Subj. In this case, communication has failed and no one is happy. We have shown each player as getting a payoff of –10. A risk averse speaker might choose to use a definite description, that is “*the cop*”. Choosing an appropriate definite description entails costs on the part of the speaker; producing the resulting description involves still more costs. Equally, the hearer must expend effort processing the definite description which is presumably somewhat greater than that expended in processing a pronoun. If the speaker selects *the cop*, which unambiguously denotes Subj, communication is successful albeit at greater costs (and less risk) to both the speaker and the hearer; both receive a payoff of 5. Compare this payoff with the payoff associated with s_2 (the intention to refer to the hoodlum), where the speaker selects the definite description *the hoodlum* and both the speaker and the hearer receive a payoff of 6. In the former case, there is a small penalty for using a definite description to refer to a more prominent element; this is in accord with the Context Factors in 9c. Since the hoodlum is less prominent than the cop, the penalty does not apply to *the hoodlum*.

Now consider the tree rooted at s_2 . In this case the speaker intends to signal discourse object Obj. As above, the speaker can either choose a pronoun or a definite description. If she chooses the pronoun and the hearer correctly selects Obj, then the payoff is (9,9); this is slightly lower than in the sub-tree rooted at s_1 because Obj corresponds to the hoodlum, which is less prominent than the cop; this is in accord with the context factors discussed in 9c which give a preference to using pronouns for more prominent discourse elements. Clark and Parikh (2007) assume that prominence correlates with the grammatical function of the expression used to evoke the discourse element in question in the previous sentence. Since Subj, the cop, was evoked by the subject in the previous sentence and Obj, the hoodlum, was evoked by the object, Subj is more prominent than Obj and there is a higher cost associated with using a pronoun for Obj. Following work in Centering theory (including, among many others, Grosz et al. (1995)), the prominence ranking in 10 determines prominence:

- (10) Subject > Indirect Object > Direct Object > Others

That is, the subject of the preceding sentence is the most prominent discourse element, followed by the indirect object, then the direct object.

Table 6

Expected utility calculations.

Choice	Probability	Utility	Expected utility
$\{(s_1, \text{the cop}), (u_1, \text{Subj})\}$	$\rho_1 = 0.5$	(5, 5)	(2.5, 2.5)
$\{(s_1, \text{pronoun}), (\{t_1, t_2\}, \text{Subj})\}$	$\rho_1 = 0.5$	(10, 9)	(5, 4.5)
$\{(s_1, \text{pronoun}), (\{t_1, t_2\}, \text{Obj})\}$	$\rho_1 = 0.5$	(-10, -10)	(-5, -5)
$\{(s_2, \text{pronoun}), (\{t_1, t_2\}, \text{Subj})\}$	$\rho_2 = 1 - \rho_1 = 0.5$	(-10, -10)	(-5, -5)
$\{(s_2, \text{pronoun}), (\{t_1, t_2\}, \text{Obj})\}$	$\rho_2 = 1 - \rho_1 = 0.5$	(9, 9)	(4.5, 4.5)
$\{(s_2, \text{the hoodlum}), (u_2, \text{Obj})\}$	$\rho_2 = 1 - \rho_1 = 0.5$	(6, 6)	(3, 3)

As we noted, above, the speaker could have chosen to denote Obj with a definite description, as indicated by the other branch emanating from state s_2 . If this course of action is selected, then the production costs are higher (in accord with the cost factor in 9b), but the speaker and hearer are once again guaranteed success with a payoff of (6,6).

Each of root information states, s_1 and s_2 , of the sub-trees in Fig. 1 is associated with a prior probability, ρ_1 and ρ_2 . Clark and Parikh (2007) assume that $\rho_1 = \rho_2 = 0.5$. This need not always be the case; in fact, we will drop this assumption and explore an alternative in the next section (see also Clark and Parikh, 2007, where changes in probabilities are exploited to account for a variety of phenomena associated with discourse anaphora).

We will assume that both players seek to maximize their own payoffs. In particular, each player will seek to maximize his or her *expected utility* given that the other player also seeks to maximize his or her expected utility. Expected utility is given by the payoff (or utility) of an outcome times its probability:

$$(11) \quad \text{Expected Utility} = \text{probability} \times \text{utility}$$

Both the speaker and the hearer have an interest in maximizing their expected utility, which means that they have an interest in using a strategy—a plan of play—that gives the highest expected payoff. This, in turn, implies that they have an interest in discovering an *equilibrium* strategy; that is, a strategy that has the property that no competing play can do better, given the way that the other agent plays. In an equilibrium state a player has no reason to change his or her strategy since any other move will result in a lower payoff. This is the intuition behind the mathematics of the Nash equilibrium.⁷

Table 6 shows the expected utility calculations for the game in Fig. 1. The choice columns show the individual choices for the two players. For example, the choice entry in the second row can be read as follows:

$$(12) \quad \{(s_1, \text{pronoun}), (\{t_1, t_2\}, \text{Subj})\}$$

The speaker, in state s_1 , chooses to produce a pronoun. The hearer, in the information set $\{t_1, t_2\}$, selects discourse entity Subj.

We have grouped together the various branches of the game tree in Fig. 1. In particular, the possible choices for the information set $\{t_1, t_2\}$ are grouped together. Within that grouping, there is a clear equilibrium strategy, which is, happily, the strategy in 12. If the speaker is in s_1 —that is, she intends to refer to Subj (the cop)—then she should pick the pronoun. This drives the hearer to select Subj since choosing Obj results in a (severely) lower payoff. Thus, the hearer has no reason to defect from Subj. Equally, the speaker has no reason to produce the definite description *the cop* since the expected utility for this choice is lower than the expected utility for producing a pronoun. Thus, neither player has any reason to defect from this strategy.

We need, now, to assemble *strategies* from the individual choices shown in Table 6. A strategy will tell the individual players what to do given the various conditions of play. We are, however, interested in finding *equilibrium strategies*; strategies that produce optimal expected utility in the sense that neither player can increase his or her expected utility by defecting from the strategy. This is, in fact, the idea behind *Nash equilibria*. Nash gave a formal

⁷ Any introduction to game theory, for example Myerson (1991) or Osborne and Rubinstein (1994), will contain a lengthy discussion of the mathematics of Nash Equilibrium, so we will not go into the mathematics here, beyond an intuitive discussion. See also van Rooij (2006) for an introduction to game theory for linguists.

method for finding equilibria in a game. In fact, the game in Fig. 1 contains two Nash equilibria. We show the equilibrium strategies in 13; underneath each strategy, we compute the players' expected utility:

- (13) a. $\{(s_1, \text{pronoun}), (s_2, \text{the hoodlum}), (\{t_1, t_2\}, \text{Subj})\}$
 Speaker: $(\rho_1 \times 10) + (\rho_2 \times 6) = (0.5 \times 10) + (0.5 \times 6) = 8$
 Hearer: $(\rho_1 \times 9) + (\rho_2 \times 6) = (0.5 \times 9) + (0.5 \times 6) = 7.5$
 Both: (8, 7.5)
- b. $\{(s_1, \text{the cop}), (s_2, \text{pronoun}), (\{t_1, t_2\}, \text{Obj})\}$
 Speaker: $(\rho_1 \times 5) + (\rho_2 \times 9) = (0.5 \times 5) + (0.5 \times 9) = 7$
 Hearer: $(\rho_1 \times 5) + (\rho_2 \times 9) = (0.5 \times 5) + (0.5 \times 9) = 7$
 Both: (7, 7)

We will discuss the strategies in 13 in more detail below. For now, consider the expected utility calculations shown for both strategies. The strategy in 13a yields a higher expected utility—(8, 7.5)—than the strategy in 13b, which has an expected utility of (7, 7). Following Parikh (2001), we will suppose that we can choose among these by selecting the Nash equilibrium that has the highest expected utility, the Pareto-dominant Nash equilibrium. A Nash equilibrium is Pareto-dominant if its payoff is at least as high as any other Nash equilibrium. In other words, no other Nash equilibrium will yield a better payoff. This suggests that, all else being equal, the players will tend to play by the strategy in 13a, rather than the one in 13b. Let us turn, now, to what the strategies in 13 say.

Consider, first, the strategy in 13a: it states that if the speaker is in the state s_1 , where she wishes to invoke discourse element Subj, she should use a pronoun. If she wishes to invoke discourse element Obj, she should use the definite description *the hoodlum*. If the hearer finds himself in the information set $\{t_1, t_2\}$ then he should pick discourse element Subj. It goes without saying that *the hoodlum* invokes Obj, so this has not been included in the strategy profile in 13a.

The other strategy in 13 says that if the speaker is in s_1 —the state where she intends to denote Subj, she should use a definite description. Otherwise, when she wants to denote Obj, she should use a pronoun. As we saw, though, this strategy has a lower expected utility than the strategy in 13a.

None of the other strategies are Nash equilibria. For example, the strategy of the speaker uttering pronouns both in s_1 and s_2 and the hearer interpreting a pronoun as referring to the subject is not a Nash equilibrium because there will be misunderstanding whenever the speaker wants to denote Obj and, thus, their expected payoffs will be very low, as shown in 14. Thus, given the strategy the hearer is playing, the speaker has a reason to deviate to strategy 13a in order to increase her payoffs.

- (14) $\{(s_1, \text{pronoun}), (s_2, \text{pronoun}), (\{t_1, t_2\}, \text{Subj})\}$
 Speaker: $(\rho_1 \times 10) + (\rho_2 \times -10) = (0.5 \times 10) + (0.5 \times -10) = 0$
 Hearer: $(\rho_1 \times 9) + (\rho_2 \times -10) = (0.5 \times 9) + (0.5 \times -10) = -0.5$
 Both: (0, -0.5)

In terms of our small discourse, we get the texts in 15 as “natural”. On the other hand, the texts in 16 are relatively “unnatural”, where *natural* and *unnatural* are taken as obeying or violating, respectively, the strategy in 13a.

- (15) a. A cop_i was chasing a hoodlum. He_i stumbled.
 b. A cop was chasing a hoodlum. The hoodlum stumbled.
- (16) a. A cop was chasing a hoodlum_i. He_i stumbled.
 b. A cop was chasing a hoodlum. The cop stumbled.

The game in Fig. 1 along with its associated Pareto-dominant Nash equilibrium strategy in 13a, neatly summarize the way in which form, prominence and grammatical function interact to condition the distribution of anaphoric definite descriptions and pronouns. In particular, given the form of the game in Fig. 1, we expect that rational players would behave as predicted by the strategy in 13a. This approach cleanly summarizes the relationship between general rationality and linguistic behavior.

One reviewer points out that there is psycholinguistic evidence that speakers choose the form of their utterance to ease their own processing burden, and not to facilitate comprehension on the part of the addressee. For instance, Ferreira and Dell (2000) found that, in sentences with a temporal ambiguity ('I knew (that) you would come' vs. 'I knew you'), speakers do not consistently use the optional complementizer to eliminate the temporal ambiguity. However, speakers do omit the complementizer when the following embedded material has been activated (because it has already been mentioned or prompted). Arnold et al. (2000) found that the use of double object constructions (as opposed to prepositional constructions) correlate with speakers' disfluencies and they concluded that, when speakers are having difficulty producing an utterance, they are more likely to choose a word order that eases their production process. We believe that it is clear that there is a tension between the wishes of the speaker and the wishes of the hearer, between ease of production and ease of processing, and that these studies show that speakers take into account primarily their own needs. That is, they want to maximize their payoffs and not the hearers' payoffs. However, we believe that there is a fundamental difference between the studies just mentioned and our study of pronouns. In the case of pronouns, the speaker cares about the hearer interpretation because they have a common interest in communicating and miscommunication would carry negative consequences for both players. In contrast, omission of a complementizer or choice of a particular word order do not result in miscommunication which would hurt the speaker's interest, but only in a temporal ambiguity which can create difficulty for the hearer but not to the point of resulting in miscommunication. It is an interesting question how a game-theoretical analysis should deal with these alternations that cannot cause miscommunication, but we will not pursue this issue further in this paper.

The linguistic phenomenon analyzed so far presents one ambiguous linguistic form competing against other less economical, non-ambiguous linguistic forms. Discourse anaphora in Catalan is different in this respect since there are two ambiguous forms competing against each other and competing against non-ambiguous forms. The next subsection presents an account of these cases and captures the data presented in section 2.

3.2. An analysis for null-subject languages

In this section, we extend the analysis presented in section 3.1 in order to derive the asymmetry described by the PAH, which as was shown in section 2 holds for Catalan. Consider sentence 17, which was one of the items in Experiment 1.

- (17) La Marta escrivia sovint a la Raquel. Ø Vivia als Estats Units.
 "Marta wrote frequently to Raquel. Ø Lived in the United States."

The game between hearer and speaker to resolve the anaphora in sentence 17 is shown in Fig. 2.⁸ It looks similar to the game for the English discourse in Fig. 1, but the complexity has increased because the speaker has now three choices instead of two: she can use an overt pronoun, a null pronoun or a proper noun/definite description (DD). When hearing a sentence with either of the two pronominal forms, the hearer will have to decide whether the speaker wants to refer to Marta, the previous subject, or to Raquel, the previous object. We call the information state the hearer is in after hearing a sentence with a null pronoun $\{u_1, u_2, \}$ and the one after hearing a sentence with an overt pronoun $\{t_1, t_2, \}$.

The payoffs will be assigned following a very simple criteria, based solely on the complexity of the forms. Referring expressions are ranked, so that the referring expressions which are shorter and more context-dependent receive higher payoffs, while the ones which are longer and more conventional receive lower payoffs. The hierarchy we use is shown in 18 and summarizes factors 9a and 9b.

- (18) Null Pronoun > Overt Pronoun > Proper Name/Definite Description.

The hierarchy in 18 states that nulls are the more economical form, followed by overt pronouns, followed by proper names and definite descriptions; they are assigned payoffs of 10, 8 and 5, respectively. As we mentioned before, payoffs should be taken as indices which indicate preferences and, therefore, only the relationships between them, and not their particular value, are important.

⁸ Since we assume here that the same utility is accorded both to the speaker and to the listener, there is only one number which represents the utilities for both players.

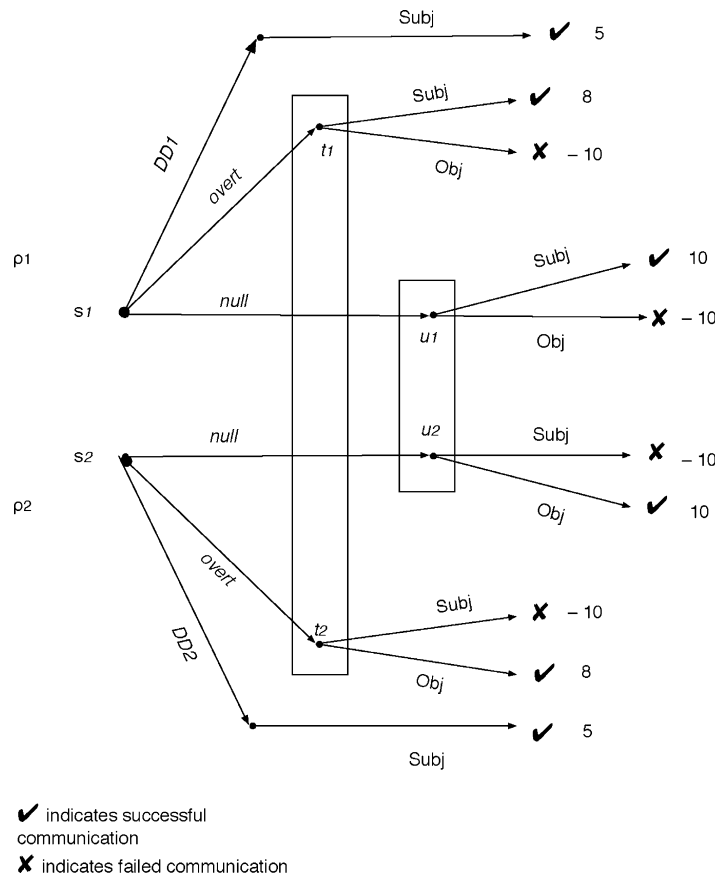


Fig. 2. Game for the Catalan pronouns.

This analysis differs from the one in [Clark and Parikh \(2007\)](#) in that the payoffs do not encode any asymmetry between referring to a subject antecedent and to an object antecedent. That is, the Context Factors (see 9c) and the hierarchy in 10 are not encoded in the payoffs; the payoffs for correctly interpreting a null pronoun are the same in s_1 and s_2 . However, the same result achieved in [Clark and Parikh \(2007\)](#) can be achieved by introducing this asymmetry in the probabilities of the information states, which is how [Clark and Parikh \(2007\)](#) modeled the various non-default antecedent assignments they dealt with.

In a nutshell, the proposal here is that $\rho_1 > \rho_2$, that is, the default option is that the referent of the subject of the current utterance U_i is the same as the referent of the previous utterance, U_{i-1} . This claim is empirically motivated and can be estimated from corpora. In order to do so, we coded three stories from a corpus of oral narrations ([Nocando, 2004](#)). In particular we coded those cases in which the subject of the utterance U_i appeared in the previous utterance U_{i-1} . Two-hundred seventeen clauses with a finite verb were coded: 76% ($n = 165$) referred to a previous subject and 24% ($n = 52$) to another constituent. Therefore, the two information states are not equally, or similarly, likely: the first one is much more likely than the second one. The idea that the default option is that the referent of the subject of the current utterance U_i is the same as the referent of the previous utterance U_{i-1} gets support from different sources other than corpus counts. First, the same claim is found in the literature from Centering Theory, in which it represents a Continue Transition, which is the preferred transition (see also [Walker et al. \(1998\)](#) on discourse continuity). Second, research about discourse structure has shown that a discourse normally sticks to the same topic, talking about the same objects and events (see [Jasinskaja and Zeevat \(2009\)](#)). Third, there is evidence from psycholinguistic studies that referents in subjects position are expected to be mentioned again. [Kaiser and Trueswell \(2008\)](#) carried out an eye-tracking experiment and they found that, after an SVO sentence, people were anticipating that a subject would be mentioned again. To sum up, subject continuity is not a property specific to a particular language or language family, but a cross-linguistic tendency and that is why it is best encoded in the probabilities. As [Jager \(2007\)](#) points out, probabilities in game theory should be used to represent cognitive and communicative tendencies, not particularities of a certain language. Subject continuity is

one of these communicative tendencies. This distribution of probabilities is common knowledge for the participants in a conversation and, thus, they can employ them to make the most efficient use of their resources.

If $\rho_1 > \rho_2$, there are four pure Nash equilibria (we will be assuming that $\rho_1 = 2/3$ and $\rho_2 = 1/3$, for the purposes of showing the calculations):

- (19) a. $\{(s_1, \text{null}), (s_2, \text{overt}), (\{u_1, u_2\}, \text{Subj}), (\{t_1, t_2\}, \text{Obj})\}$.
The expected payoff is: $\rho_1(10) + \rho_2(8) = 2/3(10) + 1/3(8) = 28/3$.
- b. $\{(s_1, \text{DD1}), (s_2, \text{null}), (\{u_1, u_2\}, \text{Obj}), (\{t_1, t_2\}, \text{Obj})\}$.
The expected payoff is: $\rho_1(5) + \rho_2(10) = 2/3(5) + 1/3(10) = 20/3$.
- c. $\{(s_1, \text{overt}), (s_2, \text{null}), (\{u_1, u_2\}, \text{Obj}), (\{t_1, t_2\}, \text{Subj})\}$.
The expected payoff is: $\rho_1(8) + \rho_2(10) = 2/3(8) + 1/3(10) = 26/3$.
- d. $\{(s_1, \text{null}), (s_2, \text{DD2}), (\{u_1, u_2\}, \text{Subj}), (\{t_1, t_2\}, \text{Subj})\}$.
The expected payoff is: $\rho_1(10) + \rho_2(5) = 2/3(10) + 1/3(5) = 25/3$.

No other strategy is a Nash equilibrium. For example, $\{(s_1, \text{null}), (s_2, \text{null}), (\{u_1, u_2\}, \text{Subj}), (\{t_1, t_2\}, \text{Obj})\}$ is not a Nash equilibrium. The speaker would always use a null pronoun, regardless of whether she wants to refer to the subject or the object, while the hearer would always understand a null pronoun as referring to the subject. This means there would always be miscommunication whenever the speaker refers to the object. The expected payoff for this strategy is: $\rho_1(10) + \rho_2(0) = 2/3(10) + 1/3(0) = 20/3$. Given the strategy the hearer is using, it is in the speaker's best interest to deviate from her own strategy and to use an overt pronoun when she wants to refer to the object: that is, she should use the strategy in 19a. This would increase her payoffs from 20/3 to 28/3.

There is a single Pareto–Nash equilibrium, which is the first equilibrium of the list above. According to this equilibrium, the speaker should use a null pronoun to refer to Subj and an overt pronoun to refer to Obj. The hearer should interpret a null pronoun as referring to Subj and an overt pronoun as referring to a Obj. It is easy to see that this strategy is equivalent to the predictions of the Position of Antecedent Hypothesis.

We capture the PAH in our model with two simple assumptions: (i) the payoffs for null pronouns are greater than the payoffs for overt pronouns and (ii) p_1 is greater than p_2 . However, the game-theoretical model is not a mere translation of the PAH, because it makes more specific predictions of when the biases predicted by the PAH will be obeyed and when they will be violated. In the default case, if $\rho_1 > \rho_2$, equilibrium 19a, which is equivalent to the PAH, is the Pareto–Nash equilibrium. If there is a change in the probabilities and $\rho_1 < \rho_2$ (if it becomes more likely that we are referring to the previous object), equilibrium 19b becomes the Pareto–Nash equilibrium. This equilibrium claims that the null pronoun should be used to refer to the previous object and, thus, encodes a violation of the biases emerging from the PAH.

This prediction of the game-theoretical model is borne out: if $\rho_1 < \rho_2$, a null pronoun is able to refer to a previous object as the following naturally occurring example from our corpus shows:

- (20) Altre cop Ø tira la granoteta fora i Ø cau a l'aigua.
“Again Ø_{big frog} pushes the little frog outside and Ø_{little frog} falls in the water.”

The speaker is narrating a story about two frogs. In the first clause of example 20, a null subject pronoun refers to the big frog, while an NP in direct object position refers to the little one. In the second clause, a null pronoun can felicitously refer to the little frog, given that the semantic content of the sentences clearly biases the hearer in this direction: if x pushes y, y and not x, is the one likely to fall. With this extra information, the speaker can use the more economical form, the null pronoun, which the hearer can interpret correctly.

Note that Carminati presents the PAH not as a grammatical constraint, but as a pragmatic principle which expresses preferences that can be violated. However, she does not provide a mechanism to express when the biases derived from the PAH can be violated. By translating the PAH into games of partial information, it becomes obvious how to do so: the probabilities of each information state encode the shared knowledge about the likelihoods of these states and, depending on how agents assess them, the PAH biases will be obeyed or violated.

Consider again in light of this analysis the results from the reading-time experiment (Experiment 2), which are repeated below in Table 7.

Table 7
Results for Experiment 2.

Condition	Second sentence reading time	Difference (observed – expected)	% correct
Condition 1: subj + null	2464	–24.2	90.5
Condition 2: subj + pron	2928	298.7	90.0
Condition 3: obj + null	2587	45.4	91.1
Condition 4: obj + pron	2701	–1.4	91.2

The corrected reading times (in the Difference column) showed the asymmetry predicted by the PAH even in cases of semantic/pragmatic biasing (that is, even when the probabilities are being manipulated). However, the effect is very different in each one of the four conditions:

- Conditions 1 and 2 correspond to cases in which there was a bias towards the subject. While our proposal is that for unbiased situations $\rho_1 > \rho_2$, in these cases the difference between both probabilities is still larger. Thus, as expected, the null pronoun condition (Condition 1) is greatly favored, while the overt pronoun condition (Condition 2) receives a large penalty.
- Conditions 3 and 4 correspond to cases in which there was some bias towards the object. So is this bias towards the object capable of eliminating the initial bias towards the subject? Looking at the results, the answer seems to be no, although they certainly show the effect of the conflicting biases. In Condition 3, there is some cost that is reflected in the reading times: that is, the biasing does not render the null pronoun completely felicitous. However, note that the penalization is much smaller than in Condition 2, as we would expect (thus, the semantic bias is indeed doing some work). In Condition 4, there is some facilitatory effect: that is, in spite of the bias towards ρ_2 , the overt pronoun is still easing the processing of the sentence. However, the facilitatory effect is quite small, particularly if we compare it with the one in Condition 1.

We take these results to indicate that the initial difference between ρ_1 and ρ_2 is fairly big and, thus, it takes many extra signals to compensate for this initial difference and to reverse the probabilities. So, even in the case of some semantic bias, ρ_1 continues to be greater than ρ_2 , and thus the overt pronoun is needed to indicate reference to the object.

This proposal can be tested experimentally by constructing sentences with several degrees of biasing. Consider two conditions: (1) mild bias, with just some semantic bias in the sentences and (2) strong bias, with semantic bias reinforced by discourse connectives. The prediction of our approach would be that if we are referring to Obj with a null pronoun, the reading times of the conditions with mild bias would still show some penalty, while the ones of the conditions with strong bias would not. Experiment 3 shows that this prediction is fulfilled.

3.3. Experiment 3: self-paced reading experiment with different degrees of biasing

Experiment 3 is very similar to Experiment 2 and it also uses the self-paced reading methodology. The goal of this experiment is to test how context affects the processing and biasing preferences of the null pronoun.

Materials. The materials consisted of 16 two-sentence discourses with eight conditions. In these discourses, the first sentence introduces two individuals by means of two proper names of the same grammatical gender, one in subject position, the other in either object or indirect object position. The second sentence contains either an overt or a null pronoun, it is semantically biased so that the pronoun refers either to the previous subject or previous object and this bias is either mild or strong. The degree of biasing is affected by means of discourse connectives. The idea behind this move is to increase discourse coherence by explicitly marking the rhetorical relation between sentences with connectives. The connectives used for the subject biasing condition were those marking narration, elaboration or explanation (*after*, *in addition*, *it turns out*), while for the object biasing condition, we used connectives that mark result and violated expectation (*that's why*, *however*). These two sets of relations tend to trigger subject and object interpretation of the pronoun respectively (Stevenson et al., 2000; Kehler, 2002; Hobbs, 1979). Note that the aim of placing a connective is not to change the interpretation of the pronoun, but to increase discourse coherence and make the intended interpretation more accessible in order to study how this affects the processing of pronouns.

The eight conditions of the experiment are:

- (21) a. Condition 1: Null pronoun + mild bias towards subject
El Joan va deixar en ridícul el Dani davant de tothom. Ø Es va excusar repetidament.
“John made fun of Dani in front of everyone. Ø Apologized many times.”
- b. Condition 2: Null pronoun + strong bias towards subject
El Joan va deixar en ridícul el Dani davant de tothom. Després, Ø es va excusar repetidament.
“John made fun of Dani in front of everyone. Afterwards, Ø apologized many times.”
- c. Condition 3: Overt pronoun + mild bias towards subject
El Joan va deixar en ridícul el Dani davant de tothom. Ell es va excusar repetidament.
“John made fun of Dani in front of everyone. He apologized many times.”
- d. Condition 4: Overt pronoun + strong bias towards subject
El Joan va deixar en ridícul el Dani davant de tothom. Després, ell es va excusar repetidament.
“John made fun of Dani in front of everyone. Afterwards, he apologized many times.”
- e. Condition 5: Null pronoun + mild bias towards object
El Joan va deixar en ridícul el Dani davant de tothom. Ø Es va ofendre molt.
“John made fun of Dani in front of everyone. Ø Was very offended.”
- f. Condition 6: Null pronoun + strong bias towards object
El Joan va deixar en ridícul el Dani davant de tothom. Per això, Ø es va ofendre molt.
“John made fun of Dani in front of everyone. That’s why Ø was very offended.”
- g. Condition 7: Overt pronoun + mild bias towards object
El Joan va deixar en ridícul el Dani davant de tothom. Ell es va ofendre molt.
“John made fun of Dani in front of everyone. He was very offended.”
- h. Condition 8: Overt pronoun + strong bias towards object
El Joan va deixar en ridícul el Dani davant de tothom. Per això, ell es va ofendre molt.
“John made fun of Dani in front of everyone. That’s why he was very offended.”

The conditions for each item set were counterbalanced and incorporated into a self-paced reading experiment together with 24 filler items and 5 practice items. Sixteen counterbalanced lists were constructed (the last eight lists with the items in reverse order), with a single randomization for all lists.

Procedure. The procedure is the same as explained for Experiment 2.

Participants. Thirty-two members from Universitat Pompeu Fabra community took part in this experiment. They had not participated in either Experiment 1 or Experiment 2.

Results. Table 8 contains the results for this experiment. The second column contains the raw reading time; the third column the difference between the observed and the expected reading time and the fourth column the percentage of correct answers.

The average reading times for the main clause were computed, after eliminating times that were longer than 7000 ms and shorter than 200 ms (about 3% of the total number of trials). The number in the “% correct” column refers to the percentage of answers in which subjects understood the pronoun as referring to the pragmatically biased antecedent.

Table 8
Results for Experiment 3.

Condition	Second sentence reading time	Difference (observed – expected)	% correct
Cond 1: subj + null + mild	2447	–176	92
Cond 2: subj + null + strong	2570	–423	92
Cond 3: subj + pron + mild	3077	288	83
Cond 4: subj + pron + strong	3342	275	85
Cond 5: obj + null + mild	2609	170	81
Cond 6: obj + null + strong	2757	–124	82
Cond 7: obj + pron + mild	2783	97	94
Cond 8: obj + pron + strong	3119	–104	77

Deviations from regressions were computed to account for length differences, following the same method explained for Experiment 2 and these results will be the ones to be discussed here. Negative numbers indicate that reading times were shorter than expected (i.e., that condition was read faster than expected) and positive numbers that they were larger than expected (i.e., that condition was read slower than expected). In the first four conditions, the ones with bias to the subject, we observe the pattern we have reported so far: conditions with null pronouns are read faster than expected, while conditions with overt pronouns are read slower than expected. The different level of biasing increases the ease of processing in the conditions with null pronouns (Condition 1 vs. Condition 2), while it does not have any significant effect in the conditions with overt pronouns (Condition 3 vs. Condition 4). The most interesting results for our purposes is the contrast observed for Conditions 5 and 6: the ones with bias to the object and null pronoun. If the bias is mild (Condition 5), we see some difficulty in processing (the sentences are read slower than expected). However, if the bias is strong (Condition 6), this difficulty disappears and the sentence is read faster than expected. This is exactly what our model predicts: if the bias is strong enough so that the probabilities are switched, the speaker can use the more economical form and can expect the hearer to process the sentence without problems. Finally, we see a parallel pattern with the overt pronoun in Conditions 7 and 8, although both the ease and the difficulty of processing are less extreme. Note that we predict that a speaker should not produce an overt pronoun with a strongly biased sentences and, therefore, the sentences in Condition 8 should be unnatural. This may be somewhat reflected in the percentage of correct answers, which is the lowest for all conditions (77%). However, it seems that the strong bias overrides the conflicting linguistic cue and hearers can anyway process the sentence with ease (although less than in Condition 6, with the null pronoun).

The data regarding the difference between observed and expected reading times was submitted to an ANOVA analysis. The effect of type of pronoun was significant ($F(1, 31) = 7.55, p < 0.01$; $F(1, 15) = 26.67, p < 0.001$), as well as the effect of type of biasing, although only marginally by items ($F(1, 31) = 7.74, p < 0.01$; $F(1, 15) = 3.41, p = 0.083$). In addition, there was also a significant bias by pronoun interaction ($F(1, 31) = 11.85, p < 0.001$; $F(1, 15) = 12.94, p < 0.001$).

These results clearly show that contextual information is crucial in assigning antecedents to different referential forms and in processing them. Contextual information is explicitly captured in the different information states of games of partial information. These games provide a way of modeling the context and placing probabilities to different contextual state of affairs and this is one of the reasons why they are highly suitable to analyze phenomena such as the one studied in this paper.

4. Conclusion

In this paper, we have presented an analysis of the distribution of null and overt pronouns in Catalan. We have presented psycholinguistic evidence that the Position of Antecedent Hypothesis holds for Catalan: overt and null pronouns have different antecedent biases. This asymmetry can be naturally captured using the tools from Game Theory: speakers and listeners are rational agents using the available linguistic resources in the most efficient way in the communication game. Furthermore, by redefining a pragmatic principle like the PAH into a game of partial information, the analysis predicts not only when the biases derived from the PAH will be obeyed but also when they will be violated.

There remain several issues to be addressed in future work. In the first place, in this paper we provide evidence that the PAH holds for Catalan across-sentences and that null and overt pronouns tend to refer to syntactically different antecedents. However, it could be that what really matters is not the grammatical function, but some aspect of the information structure of the sentence. For example, a possible version of the PAH recast in pragmatic terms would be the following: null pronouns prefer to refer to antecedents in topic positions and overt pronouns to antecedents in non-topic positions (see Frana (2007) and Vallduví (1992) for proposals along this line). Since, in Catalan preverbal subjects are topics, our experiments cannot distinguish between the original PAH and the modified PAH. In contrast, experimental material in which the topic occupies a different syntactic position (by means of a left-dislocated object, for example) would allow us to distinguish both hypotheses. If the pragmatic hypothesis turned out to be correct, this may explain why our results were somewhat weaker than Carminati's results for intrasentential anaphora: the context of our experimental items may be too limited to construct the pragmatic structure of the sentence. We leave the investigation of this hypothesis for future work.

Also, we believe that the game-theoretic machinery used here can be used to model other uses of the overt pronouns such as focal and contrastive uses. In particular, the different contrastive uses seem naturally amenable to a

game-theoretic analysis. Rigau (1989) notes that stressed and unstressed pronouns convey different kinds of contrast. As 22 shows, the stressed overt pronoun conveys exhaustivity, what is predicated of the subject is not true of anyone else, and thus is incompatible with the second clause in 22b. In contrast, the unstressed overt pronoun does not convey exhaustivity, but a weaker meaning and thus is compatible with the second sentence in 22c.

- (22) a. Qui vol venir, tu o en Joan?
 “Who wants to come, you or John?”
 b. *JO vull venir.. en Joan, no ho sé.
 c. Jo vull venir . . . en Joan, no ho sé.
 “I want to come . . . John, I don’t know.”

Note that the more complex form (the stressed overt pronoun) conveys a stronger, more complex meaning than the simpler form (the unstressed overt pronouns) and, therefore, in a way, mirrors the asymmetry found in the referring preferences of null and overt pronouns. It is another division of labor between linguistic forms and meanings which Game Theory could naturally account for.

In general, we hypothesize that many pragmatic notions, like focus, theme/rheme, and so on, are subject to a game-theoretic analysis. In these cases, we would view utterances as moves in a “super game” that covers discourse segments larger than a single utterance. The issue is of interest for us, since there is a clear interaction between syntactic position, information structure and the interpretation of discourse anaphors.

Finally, expected utility plays a central role in the analysis of linguistic choice. We would expect that expected utility should be fundamental to explaining linguistic variation of the type we discussed above. Recall that we derived the asymmetries between subject and other positions by increasing the probability mass associated with subjects, thus raising their expected utility. Another possibility, though, would be to try to derive the variation by manipulating the utilities and then deriving the probabilities via mixed strategies (see Myerson, 1991). The use of mixed strategies to model linguistic variation raises a number of conceptual issues that merit investigation. We leave these questions for future research.

Acknowledgements

We wish to thank the audiences at the University of Pennsylvania, Universitat Pompeu Fabra and the Ambiguity in Anaphora Workshop (ESSLI 06), where previous versions of this work were presented, for their useful comments and feedback. We also thank the two anonymous reviewers for their very detailed comments and feedback. The second author wishes to acknowledge the support of a grant from the NIH, grant NS44266.

References

- Alonso-Ovalle, Luis, Fernández Solera, Susana, Frazier, Lynn, Clifton, Charles, 2002. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Journal of Italian Linguistics* 14 (2), 151–169.
- Ariel, Mira, 2001. Accessibility theory: an overview. In: Sanders, T., Schilperoord, J., Spooren, W. (Eds.), *Text Representation: Linguistic and Psycholinguistic Aspects*. John Benjamins, Amsterdam.
- Arnold, Jennifer E., Wasow, Thomas, Losongco, Anthony, Ginstrom, Ryan, 2000. Heaviness vs. newness: the effects of structural complexity and discourse status on constituent ordering. *Language* 76, 28–55.
- Cameron, Richard, 1992. Pronominal and null subject variation in Spanish: constraints, dialects, and functional compensation. Ph.D. Thesis. University of Pennsylvania.
- Caramazza, Alfonso, Grober, Ellen, Garvey, Catherine, Yates, Jack, 1977. Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior* 16, 601–609.
- Carminati, Maria Nella, 2002. The processing of Italian subject pronouns. Ph.D. Thesis. University of Massachusetts.
- Clark, Robin, Parikh, Prashant, 2007. Game theory and discourse anaphora. *Journal of Logic, Language and Information* 16 (3), 265–282.
- DiEugenio, Barbara, 1998. Centering in Italian. In: Joshi, A.K., Walker, M., Prince, E. (Eds.), *Centering Theory in Discourse*. Oxford University Press, Oxford, pp. 114–137.
- Dimitriadis, Alexis, 1996. When pro-drop languages don’t: overt pronominal subjects and pragmatic inference. In: *Proceedings of Chicago Linguistics Society*, vol. 32. pp. 33–47.
- Ferreira, Fernanda, Clifton, Charles, 1986. The independence of syntactic processing. *Journal of Memory and Language* 25, 555–568.
- Ferreira, Victor S., Dell, Gary S., 2000. Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology* 40 (4), 296–340.

- Frana, Ilaria, 2007. The role of discourse prominence in the resolution of referential ambiguities. Evidence from co-reference in Italian. *University of Massachusetts Occasional Papers* 37.
- Grosz, Barbara, Joshi, Aravind, Weinstein, Scott, 1995. Centering: a framework for modelling the local coherence of discourse. *Computational Linguistics* 21 (2), 203–225.
- Hobbs, Jerry R., 1979. Coherence and coreference. *Cognitive Science* 3, 67–90.
- Jäger, Gerhard, 2007. Evolutionary game theory and typology: a case study. *Language* 83, 74–109.
- Jasinskaja, Katja, Zeevat, Henk, 2009. In Rieber, A., Solstad, T. (Eds.), *Contrast in Russian and English*. *Proceedings of Sinn und Bedeutung*, 13.
- Kaiser, Elsi, Trueswell, John, 2008. Interpreting pronouns and demonstratives in Finnish: evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes* 23 (5), 709–748.
- Kameyama, Megumi, 1985. Zero anaphora: the case of Japanese. Ph.D. Thesis. Stanford University.
- Kehler, Andrew, 2002. *Coherence, Reference and the Theory of Grammar*. CSLI Publications, Stanford, CA.
- Koornneef, Arnout W., Van Berkum, Jos J.A., 2006. On the use of verb-based implicit causality in sentence comprehension: evidence from self-paced reading and eye tracking. *Journal of Memory and Language* 54 (4), 445–465.
- Luján, Marta, 1985. Binding properties of overt pronouns in null pronominal languages. In: Kroeber, P., Eilforth, W., Peterson, K. (Eds.), *Proceedings of the Chicago Linguistics Society*, vol. 21. pp. 424–438.
- Myerson, Roger, 1991. *Game Theory: Analysis of Conflict*. Harvard Univ. Press, Cambridge, MA.
- Nocando, 2004. Non-canonical constructions in oral speech: a cross-linguistic study. Technical Report I+D HUM2004—04463. Universitat Pompeu Fabra, Barcelona.
- Osborne, Martin J., Rubinstein, Ariel, 1994. *A Course in Game Theory*. The MIT Press, Cambridge, MA.
- Parikh, Prashant, 2001. *The Use of Language*. Center for the Study of Language and Information, Stanford.
- Rigau, Gemma, 1986. Some remarks on the nature of strong pronouns in null-subject languages. In: Bordelais, I., Contreras, H., Zagana, K. (Eds.), *Generative Studies in Spanish Syntax*. Foris, Dordrecht.
- Rigau, Gemma, 1989. Connexity established by emphatic pronouns. In: Conte, M., Petöfi, J., Sözer, E. (Eds.), *Text and Discourse Connectedness*. John Benjamins Publishing, Amsterdam.
- Silva-Corvalán, Carmen, 1977. A discourse study of word order in the Spanish spoken by Mexican-Americans in West Los Angeles. Master's Thesis. University of California.
- Stevenson, Rosemary, Knott, Alistair, Oberlander, Jon, McDonald, Sharon, 2000. Interpreting pronouns and connectives: interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes* 15 (3), 225–262.
- Trueswell, John C., Tanenhaus, Michael K., Garnsey, Susan M., 1994. Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33 (3), 285–318.
- Turan, Umit, 1995. Null vs. overt subjects in Turkish discourse: a centering analysis. Ph.D. Thesis. University of Pennsylvania.
- Vallduví, Enric, 1992. *The Informational Component*. Garland, New York.
- van Rooij, Robert, 2006. Game theory for linguists. In: Benz, A., Jäger, G., Rooij, R.V. (Eds.), *Game Theory and Pragmatics*. Palgrave Macmillan, Basingstoke, NY.
- Walker, Marilyn, Joshi, Aravind, Prince, Ellen, 1998. *Centering Theory in Discourse*. Oxford University Press, New York.

Laila Mayol is a graduate student in the Linguistics Department of the University of Pennsylvania. Her areas of interest are anaphora in null-subject languages, game-theoretical approaches to meaning and the semantics and pragmatics of non-truth conditional meaning.

Robin Clark is an associate professor in the Linguistics Department of the University of Pennsylvania. His interests include game theory, agent-based modeling, and decision theoretic aspects of the neuroscience of language.