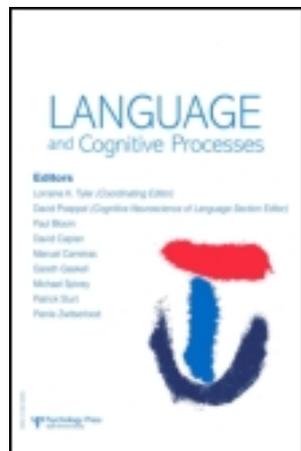


This article was downloaded by: [University of South Carolina ]

On: 03 June 2014, At: 00:10

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Language and Cognitive Processes

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/plcp20>

### Repeated names, overt pronouns, and null pronouns in Spanish

Carlos Gelormini-Lezama<sup>a</sup> & Amit Almor<sup>a b</sup>

<sup>a</sup> Linguistics Program , University of South Carolina , Columbia, SC, USA

<sup>b</sup> Psychology Department , University of South Carolina , Columbia, SC, USA

Published online: 23 Jul 2010.

To cite this article: Carlos Gelormini-Lezama & Amit Almor (2011) Repeated names, overt pronouns, and null pronouns in Spanish, *Language and Cognitive Processes*, 26:3, 437-454, DOI: [10.1080/01690965.2010.495234](https://doi.org/10.1080/01690965.2010.495234)

To link to this article: <http://dx.doi.org/10.1080/01690965.2010.495234>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

## Repeated names, overt pronouns, and null pronouns in Spanish

Carlos Gelormini-Lezama<sup>1</sup> and Amit Almor<sup>1,2</sup>

<sup>1</sup>Linguistics Program, University of South Carolina, Columbia, SC, USA

<sup>2</sup>Psychology Department, University of South Carolina, Columbia, SC, USA

In two self-paced, sentence-by-sentence reading experiments we examined the difference in the processing of Spanish discourses with repeated names, overt pronouns, and null pronouns in emphatic and nonemphatic contexts. In Experiment 1, repeated names and overt pronouns caused a processing delay when they referred to salient antecedents in nonemphatic contexts. In Experiment 2, both processing delays were eliminated when an emphatic cleft-structure was used. The processing delay caused by overt pronouns referring to salient antecedents in nonemphatic contexts in Spanish contrasts with previous findings in Chinese, where null and overt pronouns elicited similar reading times. We explain both our Spanish findings and the Chinese findings in a unified framework based on the notion of balance between processing cost and discourse function in line with the informational load hypothesis.

**Keywords:** Anaphora; Spanish; Repeated name penalty; Pronouns.

Coherent discourse often includes repeated reference to a small number of concepts. Such repetition can be achieved by means of anaphors, that is, referential expressions that co-refer with an antecedent found earlier in the same discourse. Anaphors can take many forms, including pronouns (e.g., *he*, *she*), definite expressions (e.g., *the boy*, *those girls*), and repeated names (e.g., *Mary*). Co-referentiality relations between anaphors and their antecedents serve an important role in establishing and maintaining coherence in

---

Correspondence should be addressed to Carlos Gelormini Lezama, Linguistics Program, University of South Carolina, 909 Welsh Humanities Building, Columbia, SC 29208, USA. E-mail: carlosgelormini@yahoo.com

This research was partially funded by NIH R21AG030445 and NSF BCS0822617 grants.

---

© 2010 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business

<http://www.psypress.com/lcp>

DOI: 10.1080/01690965.2010.495234

discourse. The fact that the same expression, for example a proper name, can be used anaphorically or nonanaphorically precludes a taxonomic account of anaphoric devices that is based strictly on the form of the referential expressions. Indeed, the literature suggests that a complete explanation of anaphoric processing and use should consider, in addition to the form of the anaphoric expression, other factors including the syntactic configuration in which the anaphor appears (Chambers & Smyth, 1998; Crawley, Stevenson, & Kleinman, 1990; Frederiksen, 1981), the discourse pragmatics (Almor, 1999; Ariel, 1990; Prince, 1978), and the related memory processes (Almor, 1999; Gernsbacher, 1989; Sanford & Garrod, 1981).

One theory that focuses on the role of memory in anaphoric processing is Gernsbacher's structure building framework (1989), which views anaphoric expressions as memory retrieval cues. The theory predicts that repeated names should be processed faster than pronouns because they provide more information that can more effectively activate appropriate memory representations. For example, once we have mentioned *Evita*, it may seem more efficient to refer to her again as *Evita* rather than *she* because the interpretation of the anaphoric pronoun *she* relies on a small number of properties such as gender, number, and animacy that only provide a weak retrieval cue. By the same explanation, using nonrepetitive expressions such as *the queen of the poor* or *the thief*, which add new information by revealing the speaker's attitude towards this political figure, should be less effective as memory retrieval cues than the repetition of the name. Thus, in Gernsbacher's "anaphor as a memory retrieval cue" theory, the repetition of the name appears as the best choice in most cases by providing a strong and unambiguous cue that does not introduce new information that can interfere with memory retrieval.

However, Gordon, Grosz, and Gilliom (1993) showed that Gernsbacher's view does not adequately describe referential processing in all cases. They conducted a series of experiments that demonstrated that names are harder to process than pronouns when they refer back to an entity that is the referent of an antecedent name in subject position, an effect which they dubbed the *repeated name penalty* (RNP). Importantly, the RNP does not occur when the antecedent is a direct object, indicating that the RNP is sensitive to the syntactic position of the antecedent and not just to the differences between pronouns and repeated names in word length or frequency.

The fact that the syntactic function of the antecedent affects the resolution of anaphors can be regarded as one instance of a more general tendency: the form of the anaphor that is normally used in a real linguistic exchange depends on the discourse salience of the entity referred to by the antecedent, a phenomenon that appears to be reflected across all human languages (Ariel, 1990; Garrod & Sanford, 1984; Gernsbacher, 1989; Givon, 1987; Gordon et al., 1993; Grosz, Joshi, & Weinstein, 1983; Gundel, Hedberg, & Zacharski, 1993; van-Dijk & Kintsch, 1983). Unstressed

pronouns in non *pro drop* languages, for instance, are nearly always used to refer to the most salient entity in the discourse. On the other hand, fuller expressions such as proper nouns tend to refer to less salient discourse entities.

Ariel's (1990) accessibility theory explains the inverse relationship between the salience of the referent of the antecedent and the amount of information carried by the anaphor, in terms of the different processing costs associated with different forms of reference. Proper nouns and definite expressions are claimed to have a higher processing cost than pronouns. Because of this, establishing reference to a salient entity only requires a low-cost anaphor, while establishing reference to a nonsalient entity requires a higher-cost anaphor.

Almor's (1999) informational load hypothesis (ILH) provides a more elaborate account of the use of anaphoric expressions as a whole. In this view, cost is associated with the amount of semantic information that the referring expression activates. The processing cost of an anaphor needs to be justified in terms of the discourse function it serves within a particular context: either identifying the referent, adding new information, or both. In contrast to theories that focus only on the processing cost associated with the use of each kind of anaphor, the ILH explains anaphor processing as reflecting a balance between function and processing cost. The ILH attributes aspects of both anaphor function and anaphor processing cost to mechanisms of verbal working memory. One function of an anaphoric expression is to re-activate information that is maintained in working memory and, thus, to establish a coherent link with the previous discourse. However, if the only consideration in anaphor processing were the correct identification of the entity, then repeated names would always be preferred. Yet, because the capacity of our working memory is limited, the use of an unnecessarily repeated anaphor results in superfluous activation of information in working memory, which in this theory, is part of the anaphor's processing cost. According to the ILH, the RNP reflects an unnecessary activation of information that serves no discourse function.

Although these processing principles are universal and not specific to only one language, most of the evidence for the RNP is based on research done in English. Evidence from other languages that have different referential systems can therefore provide important evidence about the generality and universality of these processing principles. In particular, because English is a *non pro drop* language, it is important to ascertain whether the same principles can explain referential processing in *pro drop* languages, which differ from English in that in many contexts, overt reference can be omitted altogether. Most Romance languages, like Spanish and Italian, are *partially pro drop* or *null subject* in that the grammatical subject can be omitted in some contexts. Other languages, like Chinese, are *fully pro drop* in that they

allow both subjects and objects to be dropped. However, unlike most Romance languages, Chinese lacks strong subject–verb agreement features and displays an impoverished morphosyntactic system. Because the referent of the omitted pronoun cannot be linked to a verbal suffix, it is commonly assumed that Chinese speakers rely heavily upon contextual and pragmatic cues to retrieve the correct antecedent of the null pronoun. Yang, Gordon, Hendrick, and Wu (1999) examined co-reference in Chinese in a series of self-paced reading task experiments that looked at the influence of three kinds of anaphor (null pronouns, overt pronouns, and repeated names) on the processing of Chinese sentences. Their results showed that the RNP extends to Chinese and that syntactic factors play a similar role in the processing of anaphoric expressions in Chinese as in English. In both languages the RNP occurs when the repeated name co-refers with an antecedent in subject position but is eliminated when the antecedent is in object position.

An additional finding of Yang et al. (1999) was that the processing of both null pronouns and overt pronouns was affected similarly by the salience of the referent. Yang et al. therefore concluded that there is no need to posit special mechanisms for how null pronouns contribute to discourse coherence in Chinese. This finding raises the intriguing question of why both null and full pronouns are used in Chinese given that Yang et al. found that the two kinds of anaphors make an indistinguishable contribution to discourse coherence. Yang et al. noted that Chinese speakers use the null pronoun extensively and suggested two possible explanations for this tendency. The first explanation is that the null pronoun involves less articulatory effort. This explanation appears to rely on the Gricean maxim of quantity, which states that speakers will normally tend to choose the least complex linguistic expression that is sufficiently informative for their communicative purposes (Grice, 1975). However, the fact that Chinese speakers also use the overt pronoun makes it unclear how this Gricean principle could account in and of itself for the extensive use of the null pronoun in this language. In addition, it is worth noting that this Gricean explanation would have predicted processing differences between null and overt forms, contrary to Yang et al.'s findings. The second explanation that Yang et al. provided is that the null pronoun may have some stylistic value related to social stratification or to the expression of certain social attitudes. This explanation is interesting but is not convincing because there is no other evidence that the use of null pronouns in either pro drop or null subject languages can be accounted for by the notion of style.

In summary, Yang et al. (1999) convincingly showed that the RNP also exists in Chinese, a pro drop language, thus providing empirical support to the argument that this phenomenon may be universal and extend to languages with different referential inventories. At the same time, Yang et al.

also left some important questions unanswered about the difference between null and overt pronouns. Here we extend this research to a partially pro drop or null subject language, in the hope of further testing the universality of the processes underlying the RNP, and getting a better understanding of the difference in processing cost and discourse function between null and overt pronouns.

To the best of our knowledge, the RNP has not yet been studied in Spanish, a language, which lies at an intermediate place between non pro drop languages like English and fully pro drop languages, like Chinese. The distribution of null vs. overt pronouns in Spanish is customarily explained in terms of the notion of emphasis (Luján, 1985; Luján, 1986). For example, in 1973, the Real Academia Española (RAE) declared that the pronominal subject is used to highlight the role of the subject or to underline its importance (Bosque & Demonte, 1999). Gili Gaya (1943) argued that overt pronouns are used when expressing the subject by means of a verbal suffix does not sufficiently highlight the involvement of the actor/subject in a certain action.

Although the RNP has not been studied in Spanish, the syntactic configuration of pronouns and their antecedents has been studied in both Spanish and Italian. Carminati (2002, 2005) showed that, in Italian, which is also a null subject language, overt and null pronouns are not in free variation, such that the null pronoun is linked to antecedents with salient referents more often than its overt counterpart. Carminati further argues that, within a sentence, salience is determined by syntactic position. Her position of the antecedent hypothesis (PAH) states that the null pronoun is preferably linked to an antecedent in the highest specifier of an inflectional phrase (Spec. IP), whereas the overt pronoun is linked to an antecedent in a lower syntactic position, such as the complement of the verb. This hypothesis predicts that reference resolution will exclusively be based on the syntactic configuration of the sentence.

Alonso-Ovalle, Fernández Solera, Frazier, and Clifton (2002) attempted to extend Carminati's results from Italian into Spanish by using an offline grammaticality judgment task. They found that in Spanish, raters also preferred to link the null pronoun with an antecedent in the Spec. IP more often than the overt pronoun. However, because they used an offline grammaticality judgment task, their results could reflect meta-linguistic processes that do not accurately reflect the online process of anaphora resolution in Spanish and may instead reflect prescriptive grammatical formulas.

In addition to the methodological concerns about the offline methodology employed in some of Carminati's (2002, 2005) experiments and in the Alonso-Ovalle et al. (2002) study, it is also not clear whether the syntactic interpretation favoured by these authors is necessary. In particular, the

results of both studies can be explained in a pragmatic framework that highlights the balance between processing cost and discourse function as the most important factor (e.g., Almor, 1999; Ariel, 1990). One example for this type of explanation is the ILH (Almor, 1999), which, as described above, states that anaphors are chosen and processed in accordance with balancing processing cost, which is related to the amount of semantic information they activate, and their discourse function. The overt pronoun carries more semantic information than the null form because it has gender and number features. However, when these features do not help in the identification of the antecedent, or add any relevant information, their cost serves no function. The ILH thus predicts that these overt forms will be more difficult to process unless their use has some other justification.

We focus here on two possible functions of the overt pronoun in Spanish. The first is to identify an antecedent that is not in the default subject position in the previous sentence. The second possible function of the overt pronoun is to signal emphasis or contrast. According to pragmatic explanations, such as the ILH, the use of the overt pronoun could be justified in such cases. In summary, such explanations predict that the overt pronoun will be more cost-effective in the following two situations: (1) when it picks an object rather than a subject antecedent; and (2) when it serves an emphatic pragmatic function. These two predictions are tested in Experiments 1 and 2, respectively.

## EXPERIMENT 1

This experiment examined the difference in the processing of Spanish discourses with repeated names, overt pronouns, and null pronouns with respect to the subject–object distinction. In order to study this issue in a discourse context, we used a self-paced reading paradigm with a sentence-by-sentence presentation. Previous research suggested that the RNP is a phenomenon concerning discourse integrative processes that do not necessarily occur immediately upon encountering an anaphor. In line with this notion, Nair and Almor (2010) showed that a word by word stimuli presentation is not well suited for studying the processes underlying the RNP but a sentence-by-sentence presentation is.

The main hypotheses of this experiment are: (a) null pronouns are the fastest to process when their antecedents are in subject position but they are the slowest to process when their antecedents are in object position; (b) overt pronouns are read slower than null pronouns with syntactically salient antecedents; (c) the RNP can be extended to Spanish. Hypotheses (a) and (b) follow Almor's (1999) view that the processing cost of an anaphor depends on the function that it serves within a specific linguistic context. Because the

default option for salient antecedents is the null pronoun, the use of an overt pronoun referring to subject antecedents would imply a much higher cost with no gain in adding new information or identifying the antecedent. If, on the other hand, null pronouns and overt pronouns are indeed in free variation in Spanish as Yang et al. (1999) argued they are in Chinese, then there should be no significant difference in the processing of these two kinds of anaphors. Hypothesis (c) reflects the fact that all extant explanations of the RNP would not predict differences between non pro drop, pro drop, and null subject languages, thus suggesting that the RNP can be elicited in Spanish too. We hypothesise that if the RNP occurs in Spanish, it will be sensitive to the syntactic status of the antecedent, just as it has been shown to be the case in English and Chinese. This would constitute further evidence of the RNP as a universal phenomenon. At the same time, hypothesis (b) predicts that in Spanish there is a penalty for the use of overt pronouns with salient antecedents, which would contrast with what has been found in Chinese (Yang et al., 1999), where reading times of sentences containing null and overt pronouns did not significantly differ.

## METHOD

### Participants

Forty-five graduate and undergraduate students from the University of Buenos Aires (UBA) and the Instituto de Enseñanza Superior en Lenguas Vivas J. R. Fernández (IESLV) participated in a single session lasting approximately 20 minutes. They were between 20 and 40 years of age and they were all native speakers of River Plate Spanish, the dialect spoken in Buenos Aires and neighbouring areas. The distinctive features of this dialect of Spanish are mainly lexical and phonological and thus play no role in this study.

### Materials

A set of 36 passages like the one in Table 1 was constructed. Each passage consisted of two sentences. Sentence (1) contained two referring expressions in the form of proper names such as *Juan* and *María*. These names could appear in two different syntactic positions: in subject position or as complement/object of the verb. Examples: *Juan met with María* vs. *María met with Juan*, where the proper name *Juan* appeared as grammatical subject or object, respectively. Sentence (2) made reference to the relevant proper name (*Juan*) in the previous sentence by means of one of the three different kinds of anaphors: (a) a repeated name: *Juan found her sad*; (b) an overt pronoun: *He found her sad*; or (c) a null pronoun: *(Null Pro) found her sad*;

TABLE 1  
Examples of all six conditions in Experiment 1

<i>Anaphor form</i>	<i>Antecedent salience</i>	
	<i>Salient (subject)</i>	<i>Nonsalient (object)</i>
Repeated name	Juan se encontró con María. Juan la vio triste. <i>Juan met with María.</i> <i>Juan found her sad.</i>	María se encontró con Juan. Juan la vio triste. <i>María met with Juan.</i> <i>Juan found her sad.</i>
Overt pronoun	Juan se encontró con María. Él la vio triste. <i>Juan met with María.</i> <i>He found her sad.</i>	María se encontró con Juan. Él la vio triste. <i>María met with Juan.</i> <i>He found her sad.</i>
Null pronoun	Juan se encontró con María. La vio triste. <i>Juan met with María.</i> <i>NULL found her sad.</i>	María se encontró con Juan. La vio triste. <i>María met with Juan.</i> <i>NULL found her sad.</i>

yielding six different conditions as shown in Table 1. Most items used simple action verbs as the main predicate, but some items used psychological verbs. Importantly, the same verbs were used in all the conditions for each item, making it unlikely that verb type would have affected our results. Each two-sentence passage was followed by a yes/no comprehension question in order to ensure that participants were processing the sentences as they read them and not merely looking at them. The names of the two characters in the experimental items were of different genders so that there was no ambiguity in the matching of the overt pronoun with its proper antecedent. A set of 36 filler items was constructed to reduce the predictability of the experimental items and mask the purpose of the experiment. These fillers were also two-sentence coherent passages including proper names but they contained syntactic structures that were not relevant to the experimental manipulation in question. Filler items were also followed by a comprehension question.

## Design

Each experimental passage was presented to each participant in only one condition, but every passage occurred in all six conditions across participants. Thus, each participant read 36 experimental items (six per condition) together with 36 filler items. The order of trials was randomised. There was a short practice session consisting of five filler passages to familiarise participants with the reading task. This session also included yes/no questions for the participants to know that they were expected to understand the sentences as they read them.

## Procedure

Participants read the instructions on the screen and began the practice session to become familiar with the self-paced reading task. At the start of each trial the sentence *Presione la barra espaciadora* (“press the space bar”) was presented on the screen. Once the participants had pressed the space bar, they were presented with an entire sentence on the screen. Participants were instructed to press the space bar as soon as they were ready to advance to the next sentence. After both experimental and filler passages, participants were presented with a comprehension question. Participants were required to press the space bar to answer “yes” and to press the “shift” key to answer “no”. Half of the questions had “yes” answers and half of them had “no” answers. The experiment was run on a Windows-based personal computer running the E-prime software. The time lapse from the presentation of Sentence 2 to the participant’s pressing of the space bar was recorded and was the dependent measure.

## RESULTS

Table 2 shows the mean reading time of the critical second sentence for all six conditions. A  $3 \times 2$  repeated measures analysis of variance (ANOVA) with factors Anaphor form (Repeated name, Overt pronoun, Null pronoun) and Antecedent salience (Subject, Object) was conducted with both participants ( $F_1$ ) and items ( $F_2$ ) as a random factor. These analyses revealed a main effect of Antecedent salience such that anaphoric expressions were read slower when they were associated with an object antecedent rather than a subject antecedent,  $F_1(1, 44) = 9.50$ ,  $MSE = 172919$ ,  $p < .005$ ,  $F_2(1, 35) = 6.11$ ,  $MSE = 217689$ ,  $p < .05$ , but no main effect of Anaphor form,  $F_1(2, 88) = 1.96$ ,  $MSE = 244870$ ,  $ns$ ,  $F_2(2, 70) = 1.65$ ,  $MSE = 205167$ ,  $ns$ . However, there was a significant interaction between the two factors,  $F_1(2, 88) = 11.97$ ,  $MSE = 280356$ ,  $p < .001$ ,  $F_2(2, 70) = 15.76$ ,  $MSE = 152256$ ,  $p < .001$ .

TABLE 2  
Mean reading times (in ms) and standard error (in parentheses)  
for Sentence 2 in Experiment 1

<i>Anaphor form</i>	<i>Antecedent salience</i>	
	<i>Salient (subject)</i>	<i>Nonsalient (object)</i>
Repeated name	2,080 (108)	2,055 (96)
Overt pronoun	2,264 (157)	2,157 (100)
Null pronoun	1,812 (103)	2,412 (172)

Four planned contrasts aimed at testing our specific predictions revealed significant differences between: (a) null pronouns and repeated names in the subject condition:  $F_1(1, 44) = 8.05$ ,  $MSE = 200106$ ,  $p < .01$ ,  $F_2(1, 35) = 8.11$ ,  $MSE = 152908$ ,  $p < .01$ ; (b) null pronouns and repeated names in the object condition:  $F_1(1, 44) = 7.43$ ,  $MSE = 385011$ ,  $p < .01$ ,  $F_2(1, 35) = 10.24$ ,  $MSE = 187926$ ,  $p < .005$ ; (c) null pronouns and overt pronouns in the subject condition:  $F_1(1, 44) = 14.66$ ,  $MSE = 312440$ ,  $p < .001$ ,  $F_2(1, 35) = 19.83$ ,  $MSE = 170332$ ,  $p < .001$ ; (d) null pronouns and overt pronouns in the object condition:  $F_1(1, 44) = 5.57$ ,  $MSE = 262208$ ,  $p < .05$ ,  $F_2(1, 35) = 4.69$ ,  $MSE = 211809$ ,  $p < .05$ . Two posthoc contrasts were also performed to compare overt pronouns and repeated names. No significant differences were found in the subject condition:  $F_1(1, 44) = 2.61$ ,  $MSE = 290210$ , *ns*,  $F_2(1, 35) = 2.48$ ,  $MSE = 211677$ , *ns*; or in the object condition:  $F_1(1, 44) = 1.86$ ,  $MSE = 125701$ , *ns*,  $F_2(1, 35) = 1.11$ ,  $MSE = 137614$ , *ns*.

## DISCUSSION

In the subject antecedent conditions, the null pronoun led to the shortest reading times. The finding that the repeated name elicited longer reading times than the null pronoun indicates that the RNP extends to Spanish and that this phenomenon may well be a reflection of a linguistic universal. The overt pronoun anaphor was also severely penalised when it referred to an antecedent in subject position. It should be noted that this experiment made use of a nonemphatic context, a fact that may explain the long reading times of the overt pronoun with subject antecedents. The experimental items did not provide a context in which the overt pronoun served an emphatic function, and thus, in the subject antecedent conditions, this anaphor did not have an advantage over its null counterpart in either the identification of the antecedent or in the addition of new information. Following Gricean principles, pragmatic explanations like the ILH state that speakers should use the least complex linguistic form that is sufficiently informative for their communicative purpose. The absence of a contrastive context makes the overt pronoun in the subject antecedent conditions both unnecessary and inefficient, and this is reflected in longer reading times. We call this processing delay the overt pronoun penalty (OPP).

In the object antecedent conditions, the null pronoun led to the longest reading times of all six conditions. It thus appears that the default antecedent of the null pronoun is the subject antecedent and forcing the reader to link a null pronoun with an object antecedent requires a re-analysis. Both facts, that the null pronoun was the most efficient option for subject antecedents, and, at the same time, the least efficient option for object antecedents, reflect the fact that the null pronoun is the default option for subject antecedents in

nonemphatic contexts. Another relevant finding is that reading times of repeated names and overt pronouns did not significantly differ in any condition. Overall, the results of this experiment indicate that the null pronoun is interpreted by default as referring to the subject, making more informative forms unnecessary. When reference is made to the object of the previous utterance, the null pronoun is ineffective and the additional information carried by both repeated names and overt pronouns serves some discourse function (identifying the correct, “nondefault” antecedent).

## EXPERIMENT 2

Experiment 1 showed that the RNP occurs in Spanish similarly to English and Chinese: repeated names are read slower than the default pronoun in each language when the antecedent is in subject position. An additional finding was that the overt pronoun in Spanish, unlike in Chinese, elicited longer reading times than the null pronoun when the antecedent was in subject position. We hypothesised that this OPP reflects a violation of the balance between cost and function. In the context we used in Experiment 1, overt pronouns did not serve their expected pragmatic function in Spanish, which is to express emphasis or contrast and were thus useful only as means of identifying a nondefault antecedent. If this interpretation is correct, embedding the same anaphoric expressions in a structure in which the overt pronoun can serve a more justified discourse function should eliminate the OPP. For Experiment 2 we therefore modified the items from Experiment 1 to include a contrastive structure similar to English cleft sentences. Spanish clefts are characterised by the absence of a cleft pronoun and by the fact that the clefted constituent is the subject of the copulative verb. Although Spanish clefts are structurally simpler, they serve a contrastive function similar to their English counterparts.

Our main hypothesis for this experiment was that the reading times of emphatic sentences with overt pronouns or repeated names referring to syntactically salient antecedents would be shorter than the reading times of comparable sentences containing null pronouns. Because the contrastive constructions serve to identify one entity as opposed to the other as the topic of the contrastive sentence, the relative salience of the two antecedents matters less. In this situation, the discourse no longer offers a default antecedent, and thus the additional information carried by the semantic features of an overt anaphoric expression are more functional in helping to identify the correct antecedent. At the same time, the use of the null pronoun will become more problematic because it depends on the availability of a default salient antecedent. In sum, we predicted that embedding the items in a contrastive structure would dissolve both penalties observed in

Experiment 1 because the additional semantic features provided by these higher-cost anaphors, will now have a discourse justification.

## METHOD

### Participants

Forty-two graduate and undergraduate students from the UBA and the IESLV participated in a single session lasting approximately 20 minutes. They were all native speakers of River Plate Spanish, and they were between 20 and 40 years of age.

### Materials

A set of 36 passages like the one in Table 3 was constructed. We used modified versions of the items from Experiment 1 in which the second sentence was embedded in an emphatic structure equivalent to cleft sentences in English. As in Experiment 1, Sentence (1) contained two referring expressions in the form of proper names such as *Juan* and *María*, which could appear in subject or object position. Examples: *Juan met with María* vs. *María met with Juan*, where the proper name *Juan* appeared as grammatical subject or object, respectively. Sentence (2) made reference to the relevant proper name (*Juan*) in the previous sentence by means of a contrastive structure containing one of the three kinds of anaphors: (a) a repeated name; (b) an overt pronoun; or (c) a null pronoun. This second

TABLE 3  
Examples of all six conditions in Experiment 2

<i>Anaphor form</i>	<i>Antecedent salience</i>	
	<i>Salient (subject)</i>	<i>Nonsalient (object)</i>
Repeated name	Juan se encontró con María. Fue Juan quien la vio triste. <i>Juan met with María.</i> <i>It was Juan who found her sad.</i>	María se encontró con Juan. Fue Juan quien la vio triste. <i>María met with Juan.</i> <i>It was Juan who found her sad.</i>
Overt pronoun	Juan se encontró con María. Fue él quien la vio triste. <i>Juan met with María.</i> <i>It was he who found her sad.</i>	María se encontró con Juan. Fue él quien la vio triste. <i>María met with Juan.</i> <i>It was he who found her sad.</i>
Null pronoun	Juan se encontró con María. Fue quien la vio triste. <i>Juan met with María.</i> <i>It was NULL who found her sad.</i>	María se encontró con Juan. Fue quien la vio triste. <i>María met with Juan.</i> <i>It was NULL who found her sad.</i>

sentence was introduced by the structure *Fue...quien* (“It was...who”); yielding six different conditions as shown in Table 3. Each two-sentence passage was followed by a yes/no comprehension question in order to ensure that participants were processing the sentences as they read them. The names of the two characters in the experimental items were of different genders so that there was no ambiguity in the matching of the overt pronoun with its proper antecedent. A set of 36 filler items was constructed to reduce the predictability of the experimental items and mask the purpose of the experiment. These fillers were also two-sentence coherent passages including proper names but they contained syntactic structures that were not relevant to the experimental manipulation in question. Filler items were also followed by a comprehension question.

## Design

Each experimental passage was presented to each participant in only one condition, but every passage occurred in all six conditions across participants. Thus, each participant read 36 experimental items (six per condition) together with 36 filler items. The order of trials was randomised. There was a short practice session consisting of five filler passages to familiarise participants with the reading task. This session also included yes/no questions for the participants to know that they were expected to understand the sentences as they read them.

## Procedure

The procedure of this experiment was identical to Experiment 1.

## RESULTS

Table 4 shows the mean reading time of the critical second sentence for all six conditions. We found a main effect of Antecedent salience such that anaphoric expressions were read slower when they were associated with an object antecedent rather than a subject antecedent,  $F_1(1, 41) = 7.30$ ,  $MSE = 208490$ ,  $p < .05$ ,  $F_2(1, 35) = 4.43$ ,  $MSE = 292365$ ,  $p < .05$ . We also found a main effect of Anaphor form,  $F_1(2, 82) = 21.01$ ,  $MSE = 294808$ ,  $p < .001$ ,  $F_2(2, 70) = 21.68$ ,  $MSE = 238708$ ,  $p < .001$ . However, there was no significant interaction between the two factors,  $F_s < 1$ .

To further explore the main effect of Anaphor form, we also performed posthoc pair-wise tests comparing reading times of sentences containing the different anaphor types. These tests revealed that sentences with repeated names were read faster than sentences with overt pronouns,  $F_1(1, 41) = 12.00$ ,  $MSE = 58655$ ,  $p < .005$ ,  $F_2(1, 35) = 5.20$ ,  $MSE = 116788$ ,  $p < .05$ , which were

TABLE 4  
 Mean reading times (in ms) and standard error (in parentheses)  
 for Sentence 2 in Experiment 2

<i>Anaphor form</i>	<i>Antecedent salience</i>	
	<i>Salient (subject)</i>	<i>Nonsalient (object)</i>
Repeated name	2,571 (96)	2,649 (120)
Overt pronoun	2,668 (100)	2,915 (108)
Null pronoun	3,072 (138)	3,215 (141)

in turn read faster than sentences with null pronouns,  $F_1(1, 41) = 14.43$ ,  $MSE = 182322$ ,  $p < .001$ ,  $F_2(1, 35) = 23.60$ ,  $MSE = 93251$ ,  $p < .001$ .

## DISCUSSION

Our hypotheses that the OPP and the RNP would be eliminated if the critical sentence contained a contrastive structure were confirmed. Sentences containing either overt pronouns or repeated names referring to subjects antecedents were no longer penalised in relation to sentences with null pronouns. Both the RNP and the OPP observed in Experiment 1 disappeared. Indeed, the null pronoun, which was the most efficient anaphor for subject antecedents in Experiment 1, elicited the longest reading times in this experiment for both subject and object antecedents.

The results of Experiment 2 support the ILH's cost-balance function-based explanation of the OPP observed in Experiment 1. In Experiment 1, this effect reflected the imbalance between the higher cost of a full pronominal form relative to null reference, in circumstances where the information conveyed by the overt pronoun made little contribution for either identifying the antecedent or for expressing contrast. In Experiment 2, on the other hand, the contrastive structure weakened the effect of salience, and thus, the semantic features carried by the overt pronoun made a discourse contribution by helping to identify the antecedent. Encountering this cleft-structure at the beginning of Sentence 2 indicates to readers the beginning of a contrastive construction in which either of the two entities that were introduced in Sentence 1 could be mentioned, thus weakening the salience difference between them (Birch & Rayner, 1997). This makes it harder to assign the null pronoun to a default salient antecedent, and makes the additional information in both repeated names and overt pronouns serve some discourse function, leading to the elimination of both the RNP and the OPP.

One possible objection might be that the use of cleft sentences in this experiment was somewhat infelicitous because the context did not offer a set

of alternative candidates. It might be argued then that the reader was forced to accommodate a presupposition for which the context offered no explicit evidence. However, Almor (1999) found that, in English, reading times of cleft sentences with anaphoric expressions were not differentially affected by the satisfaction or violation of the presupposition component. In addition, the cleft-structure was common to all conditions in our experiment. Thus, although we acknowledge that the infelicity of the Spanish clefts we used might have affected our results, we consider it unlikely that this can account for both the lack of interaction between Antecedent salience and Anaphor form, and the differences we found between the three referential forms.

## GENERAL DISCUSSION

The goal of this research was to examine the processing of discourses with repeated names, overt pronouns, and null pronouns. Two findings are particularly noteworthy: (1) for syntactically salient antecedents, both repeated names and overt pronouns elicit a processing penalty relative to null pronouns; (2) both processing delays are eliminated in emphatic contexts. The first finding shows that the RNP exists in Spanish in a similar way to English and Chinese: repeated names are read slower than null pronouns when their antecedent is syntactically salient. Therefore, as the overt pronoun in English, the null pronoun in Spanish appears to be the default pronoun, in the sense that the default pronoun in both languages is preferred over a repeated name for syntactically salient antecedents. The fact that this phenomenon has been demonstrated in English, Chinese, and now Spanish constitutes empirical evidence that helps substantiate the argument that the processes underlying the RNP occur universally, including in languages that have referential and morphosyntactic systems of considerably different complexity.

In addition, the two experiments we report here show that contrary to Carminati's PAH, the use and processing of Spanish anaphors is not exclusively determined by the syntactic configuration in which their antecedents appear. According to Carminati's hypothesis, our two experiments should have yielded similar results because they did not differ in the grammatical position of the antecedents.

The processing penalty observed here in Spanish for overt pronouns referring to salient antecedents in noncontrastive contexts sharply contrasts with Chinese, where the null pronoun was found to have no substantial processing advantage over the overt pronoun (Yang et al., 1999). Based on the pragmatic principles of cost and function (e.g., Almor, 1999: ILH; Ariel, 1990: accessibility theory), we attribute this discrepancy between Spanish and Chinese to the differences in the balance between processing cost and

discourse function in these two languages. Because Chinese lacks a rich verbal morphology, the gender and number information carried by the overt pronoun in Chinese makes this full form more useful than its Spanish counterpart. In contrast, the richer morphology of the Spanish language makes the verbal suffix crucial to the interpretation of the anaphor, thus making the semantic features of the overt pronoun less important for correctly identifying the antecedent. For subject antecedents in noncontrastive Spanish contexts, overt pronouns offer no additional functional advantage over null pronouns. Thus, the slower reading of overt pronouns in such contexts reflects an imbalance between cost and function. In this sense, the overt pronoun in Spanish does not serve the function of helping identify the antecedent in the same way as it does in Chinese. We also speculate that, historically, the lack of functionality of the overt pronoun in Spanish as means of referent identification, made the overt pronoun acquire a more specific function in this language, which is to indicate contrast. Experiment 2 showed that when the overt pronoun serves this function it is processed more easily than the null pronoun. This also suggests that in noncontrastive Spanish discourses, the overt pronoun represents a worse balance between cost and function than a repeated name despite having lower initial processing cost in terms of activated semantic information. This is compatible with the trend of slower reading of overt pronouns than repeated names in Experiment 1.

The disparity between Spanish and Chinese can thus be seen as the consequence of the difference between the rich vs. impoverished verbal morphology of these two languages, respectively. This leads to a universal prediction that the use of an overt pronoun in other pro drop languages in contexts in which it does not help to identify the antecedent or add useful information will result in the OPP. Because of the poor verbal morphology of Chinese, the overt pronoun in this language *does* help to identify the antecedent, and thus, it is not surprising that a processing delay has not been observed.

One relevant example is the difference between Brazilian and European Portuguese. Brazilian Portuguese is undergoing a process of impoverishment in the inflectional paradigm of the verb, accompanied by a reduction in the use of the null pronoun (Barbosa, Duarte, & Kato, 2005; Duarte, 1993). Duarte (1995) also observed that spoken Brazilian Portuguese is gradually showing an increase in the use of overt pronouns. We argue that these ongoing linguistic changes reflect, in themselves, a growing functional need for the overt pronoun's richer features. More generally, we claim that in pro drop languages, the more generalised use of the overt pronominal in noncontrastive contexts is associated with the impoverishment of the verbal morphology. According to this view, and because these changes are restricted to the Brazilian dialect, we predict that the OPP will be significantly more

robust in European than in Brazilian Portuguese. We are currently collaborating with investigators in these countries to test this prediction.

To conclude, this study reported two effects, the RNP and the OPP, which show that the use of null and full anaphoric expressions in Spanish is subject to several constraints that affect both cost of processing and discourse function. Although both the RNP and the OPP can be conceptualised as penalties for overspecificity, there is a typological distinction between them. The RNP appears as a universal phenomenon, independent of the anaphoric inventory of each language, whereas the OPP is, presumably, confined to a subset of pro drop languages. We propose that the OPP reveals the narrow pragmatic function of the overt pronoun in pro drop or null subject languages where the verbal morphology of the verb encodes the grammatical features that are sufficient for the identification of the antecedent of an anaphoric expression.

Manuscript received 19 October 2009

Revised manuscript received 12 May 2010

First published online 23 July 2010

## REFERENCES

- Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, 106, 748–765.
- Alonso-Ovalle, L., Fernández Solera, S., Frazier, L., & Clifton, C. (2002). Null vs. overt pronouns and the topic-focus articulation in Spanish. *Journal of Italian Linguistics*, 14(2), 151–169.
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London and New York: Routledge.
- Barbosa, P., Duarte, M. E. L., & Kato, M. A. (2005). Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4, 11–52.
- Birch, S., & Rayner, K. (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition*, 25, 653–660.
- Bosque, I., & Demonte, V. (Eds.). (1999). *Gramática descriptiva de la lengua española [Descriptive Grammar of the Spanish Language]*. Madrid: Espasa Calpe.
- Carminati, M. N. (2002). *The processing of Italian subject pronouns* (Doctoral dissertation, University of Massachusetts at Amherst).
- Carminati, M. N. (2005). Processing reflexes of the feature hierarchy (Person > Number > Gender) and implications for linguistic theory. *Lingua*, 115, 259–285.
- Chambers, C., & Smyth, R. (1998). Structural parallelism and discourse coherence. *Journal of Memory and Language*, 39, 593–608.
- Crawley, R., Stevenson, R., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, 4, 245–264.
- Duarte, M. E. L. (1993). Do pronomes nulo ao pronome pleno: a trajetória do sujeito no português do Brasil [From null pronouns to full pronouns: The trajectory of the subject in Brazilian Portuguese]. In I. Roberts & M. A. Kato (Eds.), *Português Brasileiro: Uma viagem diacrônica [Brazilian Portuguese: A diachronic journey]* (pp. 107–128). Campinas, Brazil: Editora da UNICAMP.

- Duarte, M. E. L. (1995). *A Perda do Princípio "Evite pronome" no Português Brasileiro* [The loss of the avoid pronoun principle in Brazilian Portuguese] (Doctoral dissertation, UNICAMP, Campinas, Brazil).
- Frederiksen, J. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, 4, 323–347.
- Garrod, S., & Sanford, A. J. (1984). The mental representation of discourse in a focused memory system: Implications for the interpretation of anaphoric noun phrases. *Journal of Semantics*, 1, 21–41.
- Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, 32, 99–156.
- Gili Gaya, S. (1943). *Curso superior de sintaxis española* [Advanced course in Spanish syntax]. México: Minerva.
- Givón, T. (1987). *On understanding grammar*. New York: Academic Press.
- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17, 311–347.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics III: Speech acts*. New York: Academic Press.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 21st annual meeting of the Association for Computational Linguistics* (pp. 44–50). New Brunswick, NJ: Association for Computational Linguistics.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274–307.
- Luján, M. (1985). Overt pronouns in Spanish. *Chicago Linguistic Society Proceedings*, 21, 424–438.
- Luján, M. (1986). Stress and binding of pronouns. In A. Farley, P. Farley, & K-E. McCullough (Eds.), *Papers from the parasession on pragmatics and grammatical theory* (pp. 248–262). Chicago, IL: Chicago Linguistics Society, University of Chicago.
- Nair, V. A., & Almor, A. (2010). *On the measurement of the repeated name penalty with reading times and ERP*. Unpublished manuscript.
- Prince, E. F. (1978). A comparison of wh-clefts and it-clefts in discourse. *Language*, 34, 883–906.
- Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language*. Chichester, UK: Wiley.
- van-Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Yang, C. L., Gordon, P. C., Hendrick, R., & Wu, J. T. (1999). Comprehension of referring expressions in Chinese. *Language and Cognitive Processes*, 14, 715–743.