GELORMINI LEZAMA, Carlos. **A review of the repeated name penalty: implications for null subject languages.** Revista LinguíStica / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro. Volume 8, número 2, dezembro de 2012. ISSN 1808-835X 1. [http://www.letras.ufrj.br/poslinguistica/revistalinguistica]

A REVIEW OF THE REPEATED NAME PENALTY: IMPLICATIONS FOR NULL SUBJECT LANGUAGES

Carlos Gelormini Lezama (LPEN / INECO-CONICET)

ABSTRACT
This is a critical review of the anaphoric processing delay known as the *Repeated Name Penalty* (RNP: Gordon, Grosz, & Gilliom, 1993). In this paper I argue that the RNP should be understood as an interaction effect between the anaphor type and the discourse prominence of the referent, and not merely as a pairwise comparison between sentences with repeated names and corresponding sentences with pronouns. I further propose that in null subject languages, the relevant anaphor that should be contrasted with the repeated name is the null pronoun because this type of pronoun represents the least informative anaphor available.

KEY WORDS: Repeated Name Penalty; Informational Load Hypothesis; Null Subject Languages; Anaphoric Processing.

1. INTRODUCTION

Coreference relations play a major role in the creation of discourse coherence. The mechanism by which two linguistic expressions are understood to refer to the same entity has been and still is a central topic in the field of psycholinguistics. An anaphor is a linguistic expression that refers back to another linguistic expression, the antecedent, with which it corefers. The felicitous use of anaphoric expressions appears to depend on a number of factors including the syntactic function of the antecedent (Carminati, 2002, 2005; 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). Anaphoric expressions, such as repeated names, overt pronouns, and null pronouns, have been shown to contribute to discourse coherence in a differential manner.

This paper is a critical review of the Repeated Name Penalty (RNP: Gordon, Grosz & Gilliom, 1993), a phenomenon whose definition remains controversial but which, broadly speaking, refers to a processing delay caused by repeated name anaphors in certain contexts. The final goal of this paper is to propose a definition of the RNP for languages that allow for the omission of the grammatical subject. In addition, this review is an attempt to provide a possible answer to an ongoing debate

over the interpretation of the RNP in Brazilian Portuguese and Spanish. I consider the RNP as an interaction effect between the discourse prominence of the referent and the anaphor type. Specifically, I propose that in null subject languages, the critical comparison should be between repeated names and null pronouns.

2. ANTECEDENTS OF THE RNP

For more than thirty years, studies of language use have reported a preference for pronouns over definite descriptions or repeated names in certain contexts. Marslen-Wilson, Levy and Tyler (1982), for example, asked subjects to tell a simple comic-strip story involving two characters. They found that once a character had already been established in an action sequence, subjects preferred to reference this character by means of a pronoun. They also found that repeated names were reserved for re-establishing a certain referent into a central role. Hudson-D'Zamura (1988) showed that readers judged two-sentence passages to be more coherent when the second sentence contained a pronoun, rather than a repeated name, coreferring with the subject of the previous sentence. Similar results were obtained by Brennan (1995), Chafe (1976) and Fletcher (1984). Although these were studies of language use, rather than studies of anaphoric processing, they nevertheless count as antecedents for the preference for pronouns over repeated names both when the antecedent is syntactically salient or when the referent is prominent in the discourse.

Sanford, Moar, and Garrod (1988) showed that introducing a character in a discourse narrative by means of a name, rather than a definite description, enhances its accessibility for subsequent pronominal reference. They suggested that pronouns serve a different function from that of full definite noun phrases: pronouns have a reference maintenance function, whereas full noun phrases serve to introduce new entities in the discourse. Through a series of recognition tasks, category decision tasks, and lexical decision tasks, Cloitre and Bever (1988) found that anaphor resolution was faster for pronouns than for repeated nouns. They claimed that pronouns, relative to repeated nouns, provide a more direct access to the conceptual representation of the referent.
Thus, several studies have shown that coreference through pronouns, relative to coreference through repeated names, is much more effective in the creation of discourse coherence. This effect has been found in a variety of studies using different methodological paradigms, including response times to probe words (Cloitre & Bever, 1988), production of coherent speech (Brennan, 1995), grammaticality judgments (Gordon & Hendrick, 1997; Hudson, Tanenhaus & Dell, 1986), sentence by sentence self-paced reading (Gordon, Grosz & Gilliom, 1993), and eye-tracking during reading (Kennison & Gordon, 1997).

3. GORDON'S REPEATED NAME PENALTY

Gordon, Grosz, & Gilliom (1993) presented the first study where the RNP is isolated as a specific phenomenon. Within the framework of Centering Theory (Grosz, Joshi, & Weinstein, 1983; Grosz, & Sidner, 1986) non initial utterances are taken to contain a backward looking center which serves as a link to the previous discourse. Forward looking centers are all of the potential links for subsequent reference. The authors carried out five experiments examining the processing of discourses containing sentences with pronouns and repeated names. The first three experiments examined how the features of the backward looking center influence the processing of a discourse. Experiment 1 showed that

coreference was best achieved by means of pronouns, which yielded significantly shorter reading times than comparable sentences with repeated names. In the discussion of this experiment the authors introduce the RNP in the following terms: "we will refer to the elevated reading times observed when the backward looking center was realized as a name as the repeated-name penalty, a theoretically neutral description." Experiment 2 failed to elicit a RNP when the backward looking center was the first entity to be realized but was not a grammatical subject. This was taken to prove the importance of the grammatical subject in linking back to prior discourse and consequently, in the creation of discourse coherence. Experiment 3 gave further evidence that the RNP is elicited for grammatical subjects but not for surface initial non-subject constituents.

The last two experiments examined how the backward looking center is affected by factors in the previous sentence. Experiment 4 showed that the elicitation of a RNP depends crucially on the discourse prominence of the referent as established by the antecedent: a RNP was found when the entity was the most prominent member of the set of forward looking centers. In contrast, a RNP was not found when the antecedent was not a prominent forward looking center. Experiment 5 further showed that a RNP is also elicited when the antecedent of the repeated name is a surface initial non-subject, similarly to when it is a non-initial subject.

Gordon, Grosz, and Gilliom's (1993) results contrast with previous research by Gernsbacher and Hargreaves (1988) and Gernsbacher (1989, 1990) who had found a repeated name advantage in probe recognition tasks. Gernsbacher claimed that the more informative the anaphor was, the more effective it would be in enhancing the memory representation of its referent and would be thus processed faster. As repeated names create a full match between the anaphor and its antecedent, they would constitute an optimal retrieval cue. Gordon, Grosz and Gilliom (1993), Garrod, Freudenthal, and Boyle (1994), and Gordon, Hendrick, Ledoux, and Yang (1999) considered that this effect may only be a reflection of the relationship between the anaphoric expression and the probe word, but not the relationship between the anaphoric expression and its antecedent. Thus, the discrepancy between Gordon's and Gernsbacher's results is methodologically relevant because it suggests that self-paced reading experiments and probe word recognition tasks may be measuring different processes. Gordon, Hendrick, and Foster (2000) further argued that word recognition tasks do not provide evidence about the factors that affect anaphoric processing.

Gordon and Hendrick (1998) reviewed different proposals on how anaphoric processing is affected by structural factors. Parallel function predicts that a pronoun will be interpreted as coreferential with an antecedent that has the same grammatical function (Grober, Beardsley, & Caramazza, 1978; Sheldon, 1974). This contrasts with the subject assignment proposal, which predicts that a pronoun will be interpreted as coreferential with the antecedent occupying the subject position of the previous sentence (Carminati, 2002; Crawley, Stevenson, & Kleinman, 1990; Fredericksen, 1981). Grammatical matching makes the prediction that a pronoun will be linked to the antecedent that shares the most grammatical features (Smyth, 1994). Structural prominence predicts that a pronoun will be interpreted as coreferential with the most syntactically salient antecedent (Gordon, Grosz, & Gilliom, 1993; Gordon & Scearce, 1995).

Within the framework of Centering Theory, Gordon and Hendrick (1998) proposed that the primary function of pronouns is to refer to entities that have already been mentioned in a discourse, whereas the primary function of names is to introduce entities in a discourse model. Therefore, pronouns, unlike names, are taken to be inherently anaphoric. The use of a name as an anaphor is thus expected

to produce a processing delay. Although this classification of pronouns and names may seem rather taxonomic, the authors point out that this RNP is greater when the antecedent is structurally prominent. Thus, this study placed a greater emphasis on the role that syntactic salience plays in the process of coreference. The authors review empirical evidence from acceptability judgments, reading paradigm studies, and corpus analyses which converge to show the importance of structural prominence on anaphor resolution. Gordon and Hendrick (1997) found coreference to be highly acceptable when a name precedes a pronoun but not when a name precedes a name. They further showed that when the antecedent is syntactically salient, acceptability judgments increase for name-pronoun sequences but decrease for name-name sequences. Reading time data also provide evidence that the syntactic salience of the antecedent favors name-pronoun coreference and disfavors pronoun-name coreference. Reading time experiments have further shown that the RNP is greater when the antecedent is the grammatical subject of the preceding sentence than when it is the grammatical object (Gordon, Grosz, & Gilliom, 1993; Gordon & Chan, 1995; Kennison & Gordon, 1997). Additionally, Gordon, Hendrick, Ledoux, and Yang (1999) showed that the RNP is eliminated when the antecedent is the possessor in a possessive expression (e.g., *Bill's aunt owns a lake house. Bill likes to go swimming there*) or one of the two components of a conjoined noun phrase (e.g., *John and Mary went to the store. John wanted to buy candy*). Lappin and Leass (1994) successfully designed and implemented an algorithm for pronominal anaphora resolution which includes measures of antecedent salience determined by syntactic structure. This approach is compatible with Centering Theory in that both end up giving a key role to the syntactic function of the antecedent of the anaphoric expression.

4. ALMOR'S INFORMATIONAL LOAD HYPOTHESIS

Almor's (1999) Informational Load Hypothesis (ILH) views the processing of anaphoric expressions as reflecting a general pragmatic principle in line with Grice's conversational maxims (Grice, 1975) and Relevance Theory (Sperber & Wilson, 1995). Grice's maxim of quantity states that speakers should use the least complex linguistic form that is sufficiently informative for their communicative purpose. Thus, the choice and processing cost of an anaphor will depend, among other factors, on how much information is required for the identification of the antecedent of the anaphoric expression. More generally, the ILH explains anaphoric processing as reflecting a balance between function and cost. The processing cost of an anaphor needs to be justified in terms of the discourse function that it serves in a particular context: either identifying the antecedent, adding new information, or both. The informational load of an anaphoric expression is defined with respect to its antecedent, and it refers to the semantic overlap between the representation of the anaphor and of the referent (Almor and Eimas, 2008).

In this view, the RNP would be the consequence of using an anaphor that has a high informational load with no functional justification. The maximal overlap between a repeated name and its antecedent, far from creating a processing advantage (as had been previously predicted by Gernsbacher, 1989, 1990), results in a processing delay because the high informational load of the repeated name serves no discourse function. It does not help identify the antecedent because it refers to the most prominent discourse entity, and it does not add new or relevant information. Crucially, if the repeated name refers to a non prominent discourse entity, the RNP is eliminated. This is because the high informational load of the repeated name now has the function of identifying a non default antecedent.

5. TWO RNPS IN CHINESE.

Yang, Gordon, Hendrick, and Wu (1999) examined the processing of repeated names, overt pronouns, and null pronouns in Mandarin Chinese with a self-paced reading paradigm. In Experiment 1, the first sentence introduced two characters that occupied the subject and object position. In the second and critical sentence, the entity that had been introduced in the subject position could be referred to by means of a repeated name, an overt pronoun, or a null pronoun, and the entity that had been introduced in the object position could be referred to by means of a repeated name or an overt pronoun. Results showed that sentences with repeated names were read slower than sentences with overt and null pronouns. This difference was eliminated for the second entity in the object position. Experiment 2 further showed that the RNP was elicited when the repeated name referred to a subject antecedent but not when it referred to an object antecedent. The authors concluded that for the passages studied, there was no need to posit new mechanisms for how null pronouns contribute to discourse coherence. In this view, null and overt pronouns are two types of reduced expressions which make an almost identical contribution to discourse coherence.

Overall, it is unclear what kind of RNP the authors were hoping to find in Chinese. The fact that both overt and null pronouns were read faster than repeated names allowed them to claim that there is a RNP without having to specify if the RNP is about the comparison between repeated names and overt pronouns or between repeated names and null pronouns. It should be noted, however, that the definition of the RNP that the authors give here is: "under specific circumstances, sentences with repeated names are read more slowly than sentences with pronouns, an effect that has been called the repeated-name penalty." The emphasis here is on the fact that sentences with repeated names are read more slowly only under certain circumstances. If the RNP were a generalized phenomenon, it would not reveal that much about anaphoric processing. Therefore, claiming that the RNP is the slower reading of sentences containing repeated names relative to sentences containing pronouns without making reference to, at least, antecedent salience and interaction effects, is quite a simplification.

6. ONE RNP IN SPANISH

In two self-paced sentence-by-sentence reading experiments, Gelormini-Lezama and Almor (2011) examined the difference in the processing of Spanish sentences with repeated names, overt pronouns, and null pronouns referring to salient and non salient antecedents. In this study, each item consisted of two sentences. Sentence (1) contained two referring expressions in the form of proper names such as *Juan* and *María*. Examples: *Juan met with María/ María met with Juan*, where the name *Juan* appeared in subject or object position. Sentence (2) made reference to the relevant proper name *Juan* by means of (a) a repeated name: *Juan found her sad*; (b) an overt pronoun: *He found her sad*; or (c) a null pronoun: *(Null) found her sad*. Their results showed that sentences with repeated names referring to subject antecedents were read significantly slower than corresponding sentences with null pronouns. This repeated name disadvantage was eliminated when the antecedent was in the object position.

In addition, the authors found an Overt Pronoun Penalty (OPP), such that sentences with overt pronouns referring to subject antecedents were read significantly slower than corresponding sentences with null pronouns. Importantly, this effect was eliminated when the antecedent of the anaphoric expression was in the object position. These results from Spanish sharply contrast with the findings from Chinese (Yang, Gordon, Hendrick, & Wu, 1999), where null pronouns were found to have no processing advantage relative to overt pronouns. Gelormini-Lezama and Almor (2011) argue that the

discrepancy between these two languages may be due to the fact that Spanish, unlike Chinese, has a morphologically rich verbal system, which makes the extra semantic features of the overt pronoun less important for the correct identification of the antecedent.

In sum, the Spanish results show that both repeated names and overt pronouns cause a processing delay relative to null pronouns when the antecedent is in the subject position. Both effects are eliminated when the antecedent occupies the object position or when pronouns are embedded in an emphatic structure. Interestingly, sentences with repeated names referring to subject antecedents were not read slower than corresponding sentences with overt pronouns. In fact, sentences with overt pronouns were read slower than sentences with repeated names although this difference fell short of significance.

7. TWO RNPS IN BRAZILIAN PORTUGUESE?

Results in Brazilian Portuguese have been mixed. Leitao (2005) found slower reading times for repeated names relative to overt pronouns in the object position referring to object antecedents. Queiroz and Leitao (2008) found slower reading times for repeated names relative to overt pronouns with antecedents in the subject position. The authors take their results as evidence for the RNP in Brazilian Portuguese. It should be noted, however, that their manipulation is different than the one in Gelormini-Lezama and Almor (2011), where anaphors always occupied the subject position and the manipulation involved the position of the antecedent and the type of anaphor.

Gelormini-Lezama et al. (2012) replicated the Spanish experiment in Brazilian Portuguese and found similar results in both languages, namely, that sentences with repeated names were read slower than corresponding sentences with null pronouns when the antecedent was in the subject position and that this difference was eliminated when the antecedent was in the object position. Importantly, the authors also found an OPP in Brazilian Portuguese, such that sentences with overt pronouns referring to subject antecedents generated a processing delay relative to corresponding sentences with null pronouns, an effect that was eliminated when the antecedent of the anaphoric expression was in the object position.

Therefore, it remains unclear what Brazilian Portuguese data indicate. It may be the case that repeated names are indeed read slower than both overt and null pronouns with subject antecedents, and in this case, Brazilian Portuguese would replicate the Chinese results. Leitao's (2005) and Queiroz and Leitao's (2008) data appear to support this hypothesis. In contrast, it may be the case that sentences with repeated names are read slower than sentences with null pronouns, but they are not read slower than sentences with overt pronouns, in which case, Brazilian Portuguese would replicate the Spanish results. Gelormini-Lezama et al.'s (2012) data appear to support this other hypothesis. Future research will have to address this unresolved issue.

8. DISCUSSION

Before proposing a definition of the RNP, I would like to draw the reader's attention to the fact that in the field of anaphoric processing, the notions of *antecedent salience* and *referent prominence* are often wrongly used interchangeably. On the one hand, an *antecedent* is a linguistic expression, and, as such, we can speak of its surface position, its syntactic function, its morphological complexity,

etc. On the other hand, a *referent* can be described as a discourse *entity* which can be prominent or not in the discourse representation. Much of the research reviewed in this paper shows that there is a high degree of correlation between the salience of the antecedent and the discourse prominence of the referent. In addition, a referent can be made prominent by other factors, such as the visual context, the communication situation, or by other non linguistic cues. Unsurprisingly, Centering Theory describes backward and forward looking centers as *semantic entities* and not as *linguistic expressions* (Gordon, Grosz, & Gilliom, 1993). This as an indication that in this seminal paper, the effect of antecedent salience in the elicitation of the RNP is the result of the impact that antecedent salience has on the discourse status of the semantic entity involved. Therefore, the RNP should be regarded as dependent on the discourse prominence of the referent rather than merely on the salience of the antecedent, which is only one of the ways in which discourse entities can be made prominent.

In addition, a persistent question remains concerning the exact interpretation of the RNP in null subject languages: should the RNP be understood as a comparison between repeated names and overt pronouns or as a comparison between repeated names and null pronouns? There is no doubt that in null subject languages both comparisons can be informative about the mechanisms underlying anaphoric processing. It may be the case, for example, that one of these two comparisons is a generalized phenomenon across null subject languages whereas the other one is not. Depending on the research question, either comparison could be illuminating. Yet, the discussion over the exact interpretation of the RNP in null subject languages should be based on *a priori* theoretical and methodological considerations independent of the data.

Strictly speaking, Gordon, Grosz, and Gilliom (1993) defined the RNP as a contrast between repeated names and pronouns. They never specified whether by *pronouns* they were referring to *overt pronouns* or *null pronouns*. This clarification was certainly unnecessary in a context where the language that was being studied was English, which lacks null pronouns. This may seem to show that for Gordon, Grosz, and Gilliom (1993), the RNP is indeed the comparison between sentences with repeated names and sentences with overt pronouns. It could also be argued that the authors were talking about default pronouns, which, in English, happen to be overt, and in null subject languages, happen to be null. It is also quite obvious that without null pronouns, the category of overt pronouns means something different than what it means in null subject languages. In fact, it has often been claimed that English overt pronouns behave much more like Spanish null pronouns whereas English stressed pronouns behave much more like Spanish overt pronouns (Luján, 1985, 1986; Bosque & Demonte, 1999).

Unfortunately, the results from Chinese (Yang, Gordon, Hendrick, & Wu, 1999) did not make it necessary for the authors to clarify which of the two possible comparisons the RNP was referring to, because, in fact, both possible RNPs were elicited: sentences with repeated names were read slower than both sentences with overt pronouns and sentences with null pronouns. In fact, the authors explicitly stated that the RNP in Chinese occurs for pronouns as a class and that there are no differences between reading times of sentences with overt pronouns and sentences with null pronouns.

Almor's (1999) view of pronouns and names not as separate and unrelated categories but as different parts of a long continuum can be useful to provide a reinterpretation of the RNP. Pronouns and names can be regarded as extreme points in this continuum which spreads from the least to the most informationally loaded anaphor, an idea which is in line with theories which associate the choice of anaphoric expression to the degree of accessibility of the referent (e.g., Ariel, 1990; Chafe, 1994; Gundel, Hedberg, & Zacharski, 1993). Regarding the RNP as the comparison between the two

extremes of such a continuum is crucial for the redefinition of the RNP that I propose in this paper. The RNP in English was tested by contrasting the most and the least informative anaphors available: repeated names, which have a complete semantic overlap with the antecedent, versus the lightest pronouns available in English.

Following this idea of an anaphoric continuum, it is quite obvious that null and overt pronouns cannot possibly be considered to be on the same point along that continuum. Overt pronouns contain morphological features of gender and number which null pronouns lack. Thus, overt pronouns appear to be somewhat in the middle of a continuum that goes from repeated names to phonologically unrealized pronouns. Processing differences are not necessarily expected to occur between any two adjacent anaphor types along that continuum. In fact, Yang, Gordon, Hendrick, and Wu (1999) did not find any difference in processing times of Chinese sentences containing overt or null pronouns, whereas Gelormini-Lezama and Almor (2011) found a systematic OPP in Spanish. The fact that the OPP was not elicited in Chinese is an indication that this processing delay does not merely depend on the availability of null pronouns. The OPP seems to be determined by the interplay of a variety of factors including the syntactic salience of the antecedent, the discourse prominence of the referent, the degree of richness of the verbal morphology, and the pragmatic value of the overt pronoun in each language. Because the existence of null pronouns does not guarantee that the OPP will be elicited, it is quite clear that this processing delay does not exclusively depend on the anaphoric inventory of the language.

In contrast, a RNP has been found in every language tested thus far, provided that we understand the RNP as referring to the comparison of the two end points of such a continuum. Overall, the studies reviewed in this paper show that, contrary to Gordon's original proposal, the RNP has never been a theoretically neutral description. The RNP is a theoretically loaded concept and, as such, the debate over its definition is not merely a matter of linguistic gymnastics. Rather, I hope to show that its definition is a reflection of the theoretical perspective from which we consider this phenomenon. Gordon, Grosz, and Gilliom (1993) claimed that the RNP is supportive of the Centering Theory because it shows that using a repeated name where a pronoun is expected generates a processing delay. Specifically, one of the theses of Centering Theory is that in a locally coherent discourse fragment, the backward looking center should be preferentially realized as a pronoun rather than as a repeated definite description or a repeated name. The RNP is a violation of this guiding rule and thus, a penalty occurs. In contrast, Almor (1999) explained the RNP as giving support to the ILH: repeated names are higher cost anaphors that make no significant contribution in the identification of the antecedent or in the addition of relevant information when the antecedent is salient.

The fact that the same phenomenon can be explained by more than one theory is common in all fields of science. Ideally, at one point such theories will make different predictions. If we look deeper into the explanations of the RNP it is quite apparent that Almor (1999) and Gordon, Grosz, and Gilliom (1993) make different predictions. The ILH explains the RNP in English as an imbalance between processing cost and discourse function. The informational load of repeated names is high because there is complete semantic overlap between anaphor and antecedent. However, this excess of information is not enough to explain the RNP. What is crucial in this theory is that this informationally heavy anaphor does not help identify the antecedent, add any new information, or reactivate the referent in working memory. The RNP does not occur when the antecedent is in the object position or in any other non salient position. If the RNP were the consequence of the heavy informational load of the anaphor, then the syntactic function of the antecedent would be irrelevant. But this is obviously not true. The syntactic function of the antecedent does matter and, accordingly, the processing cost of an

anaphor will be determined, among other factors, by semantic overlap and antecedent salience. Almor (1999) also showed that salience can be affected by focus structures such as cleft sentences (e.g., *it was the robin that ate the apple*) where the subject is the focus and wh- pseudo cleft sentences (e.g., *what the robin ate was the apple*), where the object is the focus. In sum, in this view, the RNP is the consequence of using an anaphor with a high informational load to corefer with an antecedent whose referent is already prominent in the current discourse. The repetition of a name in such circumstances does not help reactivate the referent, because the referent is already prominent in the discourse and it does not add relevant information about the entity involved. When the antecedent is unfocused, the RNP is eliminated.

Almor's (1999) view contrasts with Gordon, Grosz, and Gilliom's (1993) interpretation of the RNP as the consequence of the violation of the first rule of Centering Theory, namely, that reference to the discourse focus should be realized as a pronoun (Grosz, Joshi, & Weinstein, 1983). Indeed, if the RNP comes from the violation of this pronoun constraint, then any non pronominal noun phrase should elicit a similar processing delay, which Almor (1999) proves not to be the case. Therefore, the RNP cannot be explained as a mere violation of the pronoun constraint because this constraint would be equally violated by any non pronominal noun phrase anaphor and not only by repeated names.

To conclude, the RNP can no longer be considered as the longer reading times of sentences containing repeated names relative to corresponding sentences with overt pronouns. Instead, what the RNP reflects is an interaction effect. This interaction effect is perfectly compatible with the felicitous inverse relationship that has been shown to exist between the prominence of the referent and the informativeness of the anaphor (Sperber & Wilson, 1986, 1995; Ariel, 1990; Gordon & Scearce, 1995; Grosz, Joshi, & Weinstein, 1983; van-Dijk & Kintsch, 1983; Grosz, 1981, Arnold, 1998, Gundel, Hedberg, and Zacharski, 1993). Within the framework of the ILH, I would like to propose that the RNP is an interaction effect between the anaphor form and the discourse prominence of the referent. In each language the relevant comparison should be between repeated names and the least informative anaphoric form available. Thus, in null subject languages the RNP occurs if sentences containing repeated names referring to prominent referents are read slower than corresponding sentences with null pronouns, and, importantly, if this effect is eliminated when the referent in question is made non prominent. This redefinition of the RNP as a comparison between the two extremes of an anaphoric continuum enables a fair comparison of reference processing across different language types and avoids a taxonomic account of anaphors that is only based on the form of the referential expression.

UMA REVISÃO DA PENALIDADE DO NOME REPETIDO: IMPLICAÇÕES PARA LÍNGUAS DE SUJEITO NULO

RESUMO
Esta é uma revisão crítica do atraso de processamento conhecido como Penalidade do Nome Repetido (PNR: Gordon, Grosz e Gilliom, 1993). Neste artigo, defendo que a PNR deve ser entendida como um efeito da interação entre o tipo de anáfora e a saliência do referente discursivo, e não apenas como uma comparação pareada entre sentenças com nomes repetidos e sentenças correspondentes com pronomes. Proponho também que, em línguas com sujeito nulo, a anáfora relevante que deve ser contrastada com o nome repetido é o pronome nulo, porque esse tipo de pronome representa a anáfora menos informativa disponível.

PALAVRAS-CHAVE: Penalidade do Nome Repetido; Hipótese da Carga Informacional; Línguas de Sujeito Nulo; Processamento Anafórico.

REFERENCES

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, 106, 748-765.

Almor, A. & Eimas, P. D. (2008). Focus and Noun Phrase Anaphors in Spoken Language Comprehension. *Language and Cognitive Processes* 23 (2), 201-225.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London and New York: Routledge.

Arnold, J. (1998). Reference form and discourse patterns. Ph.D. dissertation, Stanford University, Stanford, CA.

Bosque, I., & Demonte, V. (Eds.). (1999). *Gramatica descriptiva de la lengua española*. Madrid: Espasa Calpe.

Brennan, S. (1995). Centering attention in discourse. *Language and Cognitive Processes*, *10*, 137–167.

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.

Chafe, W. L. (1976). Giveness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25-55). New York: Academic Press.

Chafe, W. L. (1994). *Discourse, consciousness, and time.* Chicago: University of Chicago Press.

Chambers, C., & Smyth, R. (1998). Structural parallelism and discourse coherence. *Journal of Memory and Language*, 39, 593_608.

Cloitre, M., & Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes, 3,* 293-322.

Crawley, R., Stevenson, R., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, 4, 245_264.

Fletcher, C.R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior*, 23, 487–493.

Frederiksen, J. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, 4, 323_347.

Garrod, S., Freudenthal, D., & Boyle, E.A. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, 33, 39–68.

Gelormini-Lezama, C., Maia, J. C., Vernice, M., Cunha Lima, M. L., Almor, A. (2012). *Repeated names, pronouns and null pronouns in Brazilian Portuguese and Italian*. Poster session presented at the 25th annual CUNY conference on human sentence processing. City University of New York. New York.

Gelormini-Lezama, C. & Almor, A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Language and Cognitive Processes,* 26(3), 437-454.

Gernsbacher, M.A. (1989). Mechanisms that improve referential access. *Cognition*, *32*, 99–156.

Gernsbacher, M.A. (1990). *Language comprehension as structure building.* Hillsdale, NJ.Erlbaum.

Gernsbacher, M.A., & Hargreaves, D. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, *27*, 699–717.

Gordon, P.C., & Scearce, K.A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition*, *23*, 313–323.

Gordon, P.C., & Chan, D. (1995). Pronouns, passives and discourse coherence. *Journal of Memory and Language*, *34*, 216–231.

Gordon, P.C., & Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition*, *62*, 325–370.

Gordon, P.C., & R. Hendrick (1998). The representation and processing of coreference in discourse. *Cognitive Science*, *22*, 389–424.

Gordon, P.C., Grosz, B.J., & Gilliom, L.A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*, 311–347.

Gordon, P.C., Hendrick, R., & Foster, K. (2000). Language comprehension and probe-list memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 766-775.

Gordon, P.C., Hendrick, R., Ledoux, K., & Yang, C.L.. (1999). Processing of reference and the structure of language: An analysis of complex noun phrases. *Language and Cognitive Processes*, 14, 353-379.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics III: Speech acts*. New York: Academic Press.

Grober, E. H., & Beardsley, Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition,6, 117-133.*

Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics,* 12, 175-204.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In *Proceedings of the 2lst Annual Meeting of the Association for Computational Linguistics,* Cambridge, MA.

Grosz, B. J. (1981). Focusing and description in natural language dialogues. In A. K. Joshi & B. L. Webber (Eds.), *Elements of discourse understanding* (pp. 84-105). Cambridge, England: Cambridge University Press.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69,* 274-307.

Hudson-D'Zmura, S. B. (1988). *The structure of discourse and anaphor resolution: The discourse center and the roles of nouns and pronouns.* Unpublished doctoral dissertation, University of Rochester.

Hudson, S.B., Tanenhaus, M.K., & O'Dell,G.S. (1986). The effect of the discourse center on the local coherence of a discourse. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (pp.96–101). Hillsdale, NJ: Erlbaum.

Kennison, S. M., & Gordon, P. C. (1997). Comprehending referential expressions during reading: Evidence from eye tracking. *Discourse Processes, 24, 229-252.*

Lappin, S. & Leass, H. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20, 535-561.

Leitão, M.M. (2005). *O processamento do objeto direto anafórico no português brasileiro*. Rio de Janeiro: UFRJ/ Faculdade de Letras. Tese de Doutorado em Lingüística.

Luján, M. (1985). Overt pronouns in Spanish. *Chicago Linguistic Society Proceedings, 21,* 424438.

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.

Marslen-Wilson,W., Levy, W., & Tyler, L.K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. Jarvella and W. Klein (Eds.), *Speech,place, and action* (pp.339–378). New York: Wiley.

Prince, E. F. (1978). A comparison of wh-clefts and it-clefts in discourse. Language, 34, 883_906.
Queiroz, K.L.; Leitão, M.M. Processamento do sujeito anafórico em português brasileiro. *Veredas on-line psicolinguística*, juiz de fora, 2008.

Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language.* Chichester, England: Wiley.

Sanford, A.J., Moar, K., and Garrod, S.C. (1988) Proper names as controllers of discourse focus. *Language and Speech*, 31 (1). pp. 43-56.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior, 13, 272-28* 1.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research, 23, 197-229.*

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.*Cambridge, MA: Harvard University Press.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford, UK: Blackwell.

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.