Language Processing, Acceptability, and Statistical Distribution:
A Study of Null and Overt Subjects in Brazilian Portuguese

Amit Almor^{1, 2}

Jefferson de Carvalho Maia²

Maria Luiza Cunha Lima³

Mirta Vernice⁴

Carlos Gelormini-Lezama⁵

¹Psychology Department, University of South Carolina, Columbia, SC, USA

²Linguistics Program, University of South Carolina, Columbia, SC, USA

³Universidade Federal de Minas Gerais, Brazil

⁴University of Milano-Bicocca, Italy

⁵Departamento de Ciencias Sociales, Universidad de San Andrés, Buenos Aires,
Argentina

Contact: Amit Almor
Department of Psychology
1512 Pendleton St.
Columbia, SC, USA 29208
Email: almor@sc.edu

Abstract

Two experiments and a corpus study tested whether Brazilian Portuguese (BP), which has been argued to be shifting from null subjects towards overt subjects indeed shows a comprehension preference for reduced over fuller anaphors for salient antecedents, and whether comprehension is better explained as an imbalance between processing cost and discourse function (pragmatics account), or simply the frequency of different constructions (usage-based account). Sentences with overt pronouns were read slower (Experiment 1) but rated more acceptable (Experiment 2) than sentences with null pronouns when the antecedents were salient. Sentences with repeated names showed weaker effects than overt pronouns. The corpus analysis confirmed that BP is changing, but that null subjects are not yet infrequent, especially in academic writings. We argue that these results reflect a "pronoun avoidance strategy" in BP related to its transitory state, and propose a new view that integrates elements from both the pragmatics and usage-based accounts.

Null pronouns, overt pronouns and repeated names in Brazilian Portuguese

Most discourses and texts span over multiple informational units such as sentences and utterances, and maintaining coherence between these units is necessary for effective linguistic communication (Halliday & Hasan, 1976). An important linguistic device for establishing and maintaining discourse coherence is repeated reference by means of anaphoric¹ expressions (Garrod & Sanford, 1982a; Grosz, Joshi, & Weinstein, 1983, 1995; Kintsch & van Dijk, 1978). Because this is true of all human languages, it is not surprising that they all display anaphors. More surprising is the observation that, despite differences in the inventory of referential forms in different languages, less informative expressions, such as the English unstressed pronoun, are commonly preferred when the antecedent is readily accessible to comprehenders (Gundel, Hedberg, & Zacharski, 1993).

This preference suggests the existence of some underlying constraints that are not language-specific. Indeed, many of the factors that are known to affect the processing of anaphors appear to be universal in nature and not specific to only some languages. For example, reference resolution is known to be affected by the relative syntactic configuration of the anaphor and its antecedent (Chambers & Smyth, 1998; Crawley, Stevenson, & Kleinman, 1990; Frederiksen, 1981; Gordon, Grosz, & Gilliom 1993), by discourse pragmatic principles (Almor, 1999; Ariel, 1990; Prince, 1978), and by memory constraints (Almor, 1999; Gernsbacher, 1989; Sanford & Garrod, 1981), all of which are quite likely involved in the processing of any natural language.

¹ In this paper we use the term “anaphor” to denote any expression that refers to an already established discourse entity (the “antecedent”). This is to be distinguished from the use of this term in certain work in linguistics, where it is used to denote only certain forms of referential expressions.

Despite the arguably universal nature of many of the factors that have been shown to affect anaphor processing, most of the related empirical research has been carried out in the framework of a single language, most often English. This could limit our understanding of the extent to which general principles can account for anaphor processing in different languages with different referential inventories. Moreover, research on reference processing in bilinguals, which has recently become a focus of considerable effort (e.g., Cho, 2010; Filiaci, Sorace, & Carreiras, 2013; Roberts, Gullberg, & Indefrey, 2008; Sorace & Filiaci, 2006), can clearly benefit from the better understanding of the differences and similarities in coreference processing in languages other than English.

The research we report here was undertaken to fill this gap in the literature by examining and comparing the processing of different types of anaphors (null pronouns, overt pronouns, and repeated names) in Brazilian Portuguese, a language that offers a window into the factors underlying reference choice and processing because it is currently undergoing a change in its referential preferences.

Co-referential processing

As already mentioned, the choice of an anaphoric expression is highly dependent on the discourse salience of the entity introduced by the antecedent (Almor & Nair, 2007; Ariel, 1990; Garrod & Sanford, 1982b; Gernsbacher, 1989; Givón, 1987; Gordon et al., 1993; Grosz, Joshi, & Weinstein, 1983; Gundel, Hedberg, & Zacharski, 1993; van-Dijk & Kintsch, 1983). Salient discourse entities are usually referred to by means of the least informative expression available in the language (e.g., null pronouns in a null-subject language), whereas non-salient entities are usually referred to by means of heavier anaphors such as definite expressions or proper names (Gordon & Searce, 1995; Grosz et al., 1983; van-Dijk & Kintsch, 1983). Thus, there appears to be an inverse relationship between the

salience of the antecedent and the amount of information contained in the anaphor: the more salient the antecedent is, the less informative the anaphor is likely to be.

Almor's Informational Load Hypothesis (ILH; Almor, 1999; Almor & Nair, 2007) is a pragmatics-based theory that explains this inverse relationship between the salience of the antecedent and the information load of the anaphor as the outcome of an optimization process reflecting a balance between the amount of information activated by the anaphor in working memory and its discourse function. When salient discourse entities are referred to by means of a highly informative anaphor, a larger amount of information is activated in working memory without serving any discourse function, causing a processing delay. However, when the same highly informative anaphor refers to a non-salient antecedent, the reduced salience of the latter results in less overall working memory activation and, at the same time, the information activated by the anaphor now serves the function of helping to correctly identify the antecedent and reactivate its representation. In this case, the cost associated with activating the information conveyed by the anaphor is balanced with its discourse function; therefore, no overall processing delay occurs.

The slower reading of repeated names relative to pronouns when the antecedent is salient is one notable example of imbalance between cost and function: Gordon, Grosz, and Gilliom (1993) showed that, in English, sentences containing a repeated name cause a processing delay relative to sentences with pronouns when the antecedent is salient, but not when the antecedent is not salient. This "repeated-name penalty" (RNP) has also been extended to Mandarin Chinese (Yang, Gordon, Hendrick & Wu, 1999) and Spanish (Gelormini-Lezama & Almor, 2011, 2013), suggesting that this might be a universal phenomenon.

Another instance of imbalance between cost and function is the “overt pronoun penalty” (OPP), which is the slower reading of overt pronouns referring to salient antecedents relative to null pronouns. The OPP was first reported in Spanish (Gelormini-Lezama & Almor, 2011) in which overt pronouns can serve the function of either retrieving a non-salient antecedent, or marking a contrast or emphasis. In non-contrastive contexts, an overt pronoun referring back to a salient antecedent represents a case where the richer semantic content of the overt pronoun in comparison to a null pronoun is unnecessary. According to Gelormini-Lezama and Almor, the use of a more complex anaphoric form than would be required to retrieve an already accessible antecedent is penalized with longer reading times.

While there has also been other research on null and overt pronouns in Spanish and Italian, which are typologically closer to BP, much of this research has focused on intra- rather than inter-sentential anaphora (e.g., Carminatti, 2002; Filiaci, 2011; Filiaci, Sorace, & Carreiras, 2013). For example, contrasting Italian and Iberian Spanish, Filiaci (2011) and Filiaci, Sorace, & Carreiras (2013) found a stronger preference for null over overt pronouns subject anaphors in Italian than in Spanish. The relevance of their results to the present discussion is somewhat limited however, because they only examined intra-sentential anaphora. As argued and shown by Carminatti (202), these two types of anaphora are likely subject to different types of constraints and should therefore be studied separately. Nevertheless, the body of studies mentioned above, together with other research showing regional variability in preferences for null pronouns in different regional variants of Spanish (Cameron, 1992; Flores-Ferrán, 2004), suggest that there may exist differences between languages in the likelihood of observing an OPP.

Yang et al. (1999) reported a study that did compare the processing of null and overt

pronouns as inter-sentential anaphors with subject antecedents in Chinese, which is a typologically different language from the Romance languages discussed thus far. This study found that although Chinese utilizes null pronouns, the OPP was not elicited in this language, in that no significant differences between reading times of sentences containing null and overt pronouns were found. One possible reason for this difference is that Chinese, unlike Spanish, has an impoverished verbal morphology that does not facilitate the process of reference resolution. Because the extra semantic features of the overt pronoun could serve this identification function in Chinese, the OPP is not generated (Rohrbacher, 1999). Indeed, Chinese is typologically distant from BP, and, unlike BP, which involves morphological agreement in anaphoric reference, Chinese relies on context-driven inferences to establish anaphor antecedent relations (Huang, 1994, 2000).

One possible explanation for the cross-linguistic variability of the OPP is that the use of the overt pronoun with salient antecedents is not penalized in some languages because this configuration (i.e., overt pronoun - salient antecedent) is more frequent in these languages than it is in Argentinian Spanish. This interpretation is compatible with usage-based models of language processing, which predict difficulty during comprehension when the input requires an infrequent interpretation (e.g., MacDonald, 2013; MacDonald, Pearlmutter, & Seidenberg, 1994; Seidenberg & MacDonald, 1999; Trueswell & Tanenhaus, 1994).

One of the best articulated versions of this usage-based view, MacDonald's (2013) Production-Distribution-Comprehension (PDC) framework, argues that most processing constraints originate in production, whereas comprehension is mainly driven by the expectations related to the statistical patterns resulting from such production biases. The PDC is unique in that it argues for production-specific constraints that are related to

memory and attention limitations and not to discourse pragmatics as the primary cause of the statistical patterns. As such, while this framework is compatible with the existence of linguistic pragmatic constraints, it would likely predict that the RNP and the OPP, as comprehension phenomena, reflect the statistics of the language more than pragmatic constraints operating during comprehension. In this view, while the underlying production mechanisms could explain the distributional patterns, actual processing during comprehension is driven by expectations based on the distributional patterns themselves and not on the mechanisms that may have given rise to these patterns. Therefore, an alternative explanation for the RNP and the OPP is that instead of being the result of an imbalance between cost and function during comprehension as claimed by the ILH, these effects merely reflect statistical sensitivity to usage patterns.

According to usage-based explanations (MacDonald, 2013), languages in which referential preference has shifted away from the least informative forms available (overt pronouns in non-null-subject languages and null pronouns in null-subject languages) should show no penalty related to the use of fuller forms. Brazilian Portuguese, which we describe in more detail below, is an example of a formerly full null-subject language in which the use of null subjects has become significantly more restricted. This language can thus provide an informative context in which to contrast comprehension-pragmatics views with usage-based accounts.

It is important to note that while pragmatic frameworks of reference processing have not explicitly incorporated statistical sensitivity, there is no principled reason for this. This may reflect a general challenge facing statistical accounts in that they can too easily explain any possible finding as reflecting underlying statistics, without advancing the understanding of what creates the statistical distribution in the first place. As will become

clear after we report our results, it could be profitable to combine usage-based and pragmatic accounts both by appealing to pragmatic principles to explain underlying statistics, and by appealing to usage-based statistics to explain processing when pragmatic principles do not apply.

In sum, previous research on anaphoric processing, specifically on the RNP and the OPP, leaves some open questions. The first is whether the RNP is indeed a universal phenomenon, elicited both in non-null-subject and in null-subject languages. The second question is whether the OPP is unique to Spanish or whether it occurs in other null-subject languages as well. Finally, the third question is whether processing delays such as the RNP and the OPP largely reflect comprehenders' sensitivity to statistical usage patterns in their specific languages, pragmatic constraints, or a combination of statistical and pragmatic constraints.

To address these questions, we examine in this paper whether the RNP and the OPP extend to Brazilian Portuguese, a Romance language related to Spanish and similar to it in the use of null subjects but which, unlike Spanish, is undergoing change from being a full null-subject language to a non-null-subject language. We now turn to describe some relevant aspects of this change and the referential inventory of Brazilian Portuguese in more detail.

Brazilian Portuguese

Portuguese is very similar to Spanish in its verbal morphology and its pronominal system, including the use of both null and overt pronouns. The similarity in the pronominal system is most noticeable in European Portuguese, which remains a null subject language (Barbosa, 1995, 2000; Barbosa, Duarte, & Kato, 2005). In contrast, Brazilian Portuguese has undergone considerable linguistic changes over the past century, many of which related

to the classes of overt and null pronouns and to agreement markers on the verb phrase (Roberts & Kato, 1996). Many linguists have, as a result, argued for the necessity of positing different grammars for European and Brazilian Portuguese (henceforth, EP and BP respectively).

One of the most striking recent changes in BP concerns the conditions under which null pronouns are used. Over the past thirty years, a large body of linguistic studies has demonstrated that BP, unlike EP and other null-subject languages such as Spanish and Italian, has been shifting towards a preference for overt subject positions in finite sentences (Cavalcante & Duarte, 2008; Duarte, 1995, 1996, 2000, 2003; Kato, 2000; Paredes Silva, 1988; Tarallo, 1983).

Some linguists noted that the increase in the use of overt subjects is closely related to a rise of new subject pronominal forms and to a rapid simplification of verbal morphology in BP (e.g., Duarte, 1995, 1996, 2000, 2003; Holmberg, Nayudu & Sheehan, 2009). In particular, there has been a wider use of the second-person pronoun “*você*” (“you”) than before, which, unlike the previously common second-person pronoun “*tu*,” requires a third-person morphological marker on the agreeing verb. Concurrently, BP has also seen a rise in the use of the first-person plural pronoun “*a gente*” (“we”; literally “the people,” including the speaker and the addressee/s), which also requires a third-person marker on the agreeing verb.

These changes have resulted in an impoverished verbal morphology in modern BP that does not distinguish between second-person singular, third-person singular, and first-person plural, thus possibly motivating the use of an overt pronoun with person and number features (Rohrbacher, 1999). Barbosa et al. (2005), based on Duarte (1996), claim that the use of overt pronouns as reflected in written texts has grown from 20% in 1845 to 74% in

1992. Thus, according to Duarte (2003, p. 126), even though the change towards filling the subject position proceeds slowly, overt subjects are already the unmarked option in BP in non-coordinate clauses.

Note that, despite these changes, the use of null subjects is still mandatory in BP, but only in specific contexts, for instance in expletive constructions such as weather predicates (e.g., “*Ø Está chovendo*” – *It is raining*) and generic reference sentences (e.g., “*Ø Dizem que ela foi demitida*” – *People say that she was fired*). In other contexts, the use of null pronouns is not mandatory but strongly preferred, for instance, in coordinate clauses with co-referential subjects (a context which favors null subjects even in non-null-subject languages; e.g., “*Ela foi pra casa e Ø estudou a tarde toda*” – *She went home and [she] studied all afternoon*) and in replies to yes/no questions (e.g., Q: *Você foi pra casa?* R: *Ø Fui* – Q: *Did you go home?* R: *I did*). However, when making reference to highly accessible antecedents in subject position, with no intervening linguistic material between the anaphoric null pronoun and the onset of the clause to which it belongs, the use of the null pronoun is increasingly uncommon (albeit still acceptable; Duarte, 2003). Finally, null pronouns have in fact become more common as grammatical objects (Corrêa, 1991; Cyrino, Duarte & Kato, 2000; Galves, 1996; Nunes, 1996). In summary, at present, BP can be regarded as a partial null-subject language, in which null subjects are still used, nonetheless under more restricted conditions than in consistent null-subject languages like EP, Spanish, and Italian (Holmberg, Nayudu & Sheehan, 2009; Kato, 2000).

To date, there has been very little empirical work on the processing of null and overt pronouns in BP. This is unfortunate because the recent changes in pronoun use preferences in BP can provide a valuable opportunity to examine the role of usage patterns and frequencies in the relation between anaphor form and referent accessibility and to

investigate how an ongoing language change impacts online processing and how language processing is related to the metalinguistic perceptions of speakers.

Available empirical investigations of processing pronouns in BP only examined repeated names and overt pronouns in a very restricted context and without manipulating the salience of the antecedent (Queiroz & Leitão, 2008; Leitão, 2005; Leitão & Simões, 2011; Leitão, Ribeiro, & Maia, 2012). Most relevant here is a dissertation by Maia (1997), which describes a psycholinguistic study of inter-sentential object anaphora that extends to BP the Overt Pronouns Constraint - a theoretical postulate made by Montalbetti (1984). Informally, this postulate states that overt pronouns that are in contrastive distribution with null pronouns have stricter structural constraints on their possible antecedent in comparison to null pronouns. Testing speakers of BP residing in the US, Maia verified that they indeed interpret overt pronouns in the object position more restrictively than null pronouns, leading him to propose that, in BP, null objects are interpreted as co-referring with sentence topics, whereas overt object pronouns prime are interpreted as co-referring with only grammatical subjects. While this study is interesting in and of itself, it does not speak to what happens when the pronouns are in the subject position and the antecedent is in a previous sentence.

In sum, to our knowledge, there are no published studies assessing how null pronouns are processed as co-referring expressions by BP speakers in the linguistic contexts commonly used in psycholinguistic experiments of discourse reference processing (e.g., Gordon, Grosz, & Gilliom, 1993). The work we report here aims to fill the gap in the literature by examining the processing of referential expressions in BP. Specifically, we ask whether the RNP and the OPP occur in BP. The lack of an OPP in BP could suggest a quick adaptation of BP speakers to usage patterns and preferences, whereas evidence of an

OPP in BP despite the shifting preference for overt over null-subject pronouns could be indicative of other processing principles beyond usage-based expectations only.

Experiment 1

Our first experiment examined reference processing in BP, which historically exhibited a preference for null subjects similar to Spanish, but in which, at present, null subjects are grammatical but infrequently used in comparison to overt subjects (Cavalcante & Duarte, 2008; Duarte, 1995, 1996, 2000, 2003; Kato, 2000; Paredes Silva, 1988; Tarallo, 1983). Therefore, if referential processing in comprehension closely reflects usage patterns, then BP readers should experience an RNP but not an OPP. This experiment closely followed the design and procedure used by Gelormini-Lezama & Almor (2011) in their Spanish study using, instead, BP items.

Methods

Participants. Forty-seven graduate and undergraduate students from Universidade Federal de Minas Gerais (14 males, 33 females) participated in a single self-paced reading session lasting approximately 20 minutes. All participants were native speakers of the Mineiro variety of BP (spoken in the Minas Gerais state region). Their age ranged from 18 to 43 years old (mean 24.5 years old). Both the gender and age range of this sample are representative of the undergraduate population from which the sample was drawn.

Insert Table 1 about here

Materials. Table 1 shows a sample item in all conditions. Each of our 36 items consisted of two sentences. Sentence 1 introduced two gender-distinct referents with proper names, the first in subject position and the second as the object of the sentence. Sentence 2, our critical sentence, started with a reference to either the subject or the object of the first

sentence by means of a repeated name, an overt pronoun, or a null pronoun in subject position. This resulted in a 2x3 design with six different conditions.

We constructed an item set in which we carefully controlled for item length, word frequency, and verb argument structure across conditions. We only used two-syllable proper names that are common in BP. To make sure variability in item length would not affect our results and that the items were not overly short, all the first sentences were comprised of 16-20 syllables. The second sentences were comprised of 15 syllables in the repeated name and overt pronoun conditions and of 13 syllables in the null conditions. Finally, we made sure that our items were semantically coherent pieces of discourse that described two-participant events, such that the choice of referent in the second sentence would not alter the overall semantics of the event; this ensured that the second sentence would provide a natural continuation for the first sentence in both subject and object conditions without being ambiguous.

We created additional 36 filler items of similar (albeit more variable) structure and length compared to our experimental items, and with different and varied referential schemes. Each two-sentence passage (items and fillers alike) was followed by a yes/no comprehension question.

Design. Each participant saw each experimental passage in only one condition, but each passage was seen in all 6 conditions across participants. The order of the 36 experimental items (6 per condition) and 36 filler items was randomized per participant. The experiment started with a short practice session of 5 trials, which also included yes/no questions.

Procedure. Participants read the instructions on the computer screen and began a practice session to familiarize themselves with the self-paced reading task. After pressing

the space bar to begin each practice and experimental trial, participants were presented with the first sentence of the bi-sentential passage, shown in its entirety at the center of the screen. Participants were instructed to press the space bar as soon as they finished reading the initial sentence, after which the entire second sentence replaced the first at the same position on the screen. After reading the second and final sentence of the passage, participants pressed the space bar again and were presented with a comprehension question. They were instructed to press the space bar to answer “yes” and the “shift” key to answer “no.” Half of the questions had “yes” answers and half had “no” answers.

The experiment was run on a Windows-based desktop computer running the E-Prime 1.0 software (Psychology Software Tools, Pittsburgh, PA). The time lapse between the presentation onset of the second sentence and the participants’ pressing of the space bar to proceed to the comprehension question was recorded and was the dependent measure of the experiment.

Insert Figure 1 about here

Results

The data from two participants were discarded because their accuracy was below 80%. The mean accuracy of the remaining participants was 87%. Figure 1 shows the mean reading time of the critical second sentence for all six conditions for the remaining 45 participants. Outliers were then identified by an adjusted boxplot rule, based on the upper and lower quartiles, along with a robust skewness estimator, separately for each subject (Hubert & Vandervieren 2008). Any remaining responses longer than 8000 ms were also

removed. Overall, outlier removal affected 5.56% of the data. The reading times of each type of anaphor in the remaining data are shown in Figure 1.

We analyzed log-transformed reaction times in R version 3.2.0 using mixed-effects models with R package lme4 version 1.1-7. We used model comparison to estimate the significance of each term starting with a maximal model containing the antecedent and anaphor and their interaction as fixed effects. We first tried to eliminate the interaction term and if this elimination did not result in a significant loss of model fit, we then attempted to remove each of the individual factors (Baayen, 2008). As recommended by Barr, Levy, Scheepers and Tily (2013), we included the maximal structure of by-participant and by-item random intercepts and slopes that allowed both compared models to converge. We report the coefficients in the final models thus identified, with p values approximated by the normal distribution (Barr et al., 2013).

An analysis with Antecedent (Subject vs. Object) and Anaphor (Null, Overt Pronoun, Repeated Name) as fixed factors and items and subjects as random factors showed that the interaction of the two fixed factors made a significant contribution to model fit, $\chi^2(2) = 12.736$, $p = .002$. The parameter estimates of the chosen model are shown in Table 2. To facilitate the examination of the differences between the three reference forms in the subject antecedent conditions we used treatment coding with the reference levels Subject and Null for the Antecedent and Anaphor factors respectively. The coefficients in Table 2 thus represent differences from this reference level.

As can be seen in the table, when the antecedent was the subject, sentences with overt pronouns were read significantly more slowly than ones with null pronouns. Sentences with repeated names were also read slower than ones with null pronouns but the effect was not quite significant. Importantly, following Gelormini-Lezama and Almor

(2011), we looked for the interactions between anaphor form (full form vs. null pronouns) and antecedent (subject vs. object), and not for simple pairwise differences between anaphor forms in just the subject antecedent conditions, as reflections of the RNP and the OPP. As can be seen in the table, the two interaction coefficients reflecting the RNP and the OPP, respectively, were both significant.

Insert Table 2 about here

Discussion

Despite the well-documented change of a preference for overt subjects than null pronouns in present-day BP, our results showed an OPP in BP as in Gelormini-Lezama and Almor's (2011) Spanish study. Unlike in Spanish, however, the RNP was weaker as it was reflected only in the interaction between Antecedent and Anaphor; repeated name anaphors were not read significantly slower than null pronouns when referring back to a subject antecedent. The OPP was reflected in both a significant interaction and in a significantly slower reading time of sentences with overt pronouns than sentences with repeated names.

Experiment 2

Experiment 1 demonstrated a consistent OPP in BP, a language in which the use of overt subjects has become prevalent. Although speakers of BP share strong intuitions about this shift in referential preferences, and linguists studying BP have gone as far as to argue that BP is no longer a consistent null-subject language (Holmberg, Nayudu, & Sheehan, 2009; Roberts & Kato, 1996; Tarallo, 1983), it may be that the items we used in Experiment 1 somehow make null subjects preferred over lexically expressed ones when referring to the subject of the immediately preceding sentence.

To address this possibility, Experiment 2 explicitly tested the judgments of BP speakers from the same population as that of Experiment 1 about the items we used in that experiment. Critically, Experiment 2 aimed to ascertain whether BP speakers would find our items with null pronouns more acceptable than our items with overt pronouns or repeated names when the antecedent was the subject of the previous sentence. Finding such a preference would undermine our assumption that BP speakers prefer overt subjects in the items we used in Experiment 1.

Methods

Participants. Seventy-two different undergraduate students from Universidade Federal de Minas Gerais (14 males, 58 females) participated in two acceptability judgment sessions administered at the beginning of two different classes, and lasting approximately 15 minutes each. All participants were native speakers of the Mineiro variety of BP (spoken in the Minas Gerais state region). Their age ranged from 17 to 51 years old (mean 21.6 years old).

Materials. The same set of experimental and filler passages from Experiment 1 was used in this experiment.

Design. 6 lists were created in order to counterbalance the 6 experimental conditions. The order of the 36 experimental items (6 per condition) and 36 filler items was randomized for each list. Each participant rated each experimental passage in only one condition, but across participants, each passage was rated in all six conditions.

Procedure. Participants rated each passage on a 1-7 Likert scale with 1 indicating “very bad” and 7 indicating “very good.” Passages were printed consecutively on a three-page questionnaire and participants were orally instructed to rate the passages according to how acceptable they found them in BP. They marked their ratings by circling a number from one to seven on a scale printed next to each item.

Insert Figure 2 about here

Results

Figure 2a shows histograms of ratings of the experimental passages in all conditions tested in Experiment 2 and Figure 2b shows pairwise comparisons of the densities of the distributions of ratings as well as the median rating for each condition. As can be seen in the figure, the null subject with object antecedent was rated unacceptable, likely reflecting a strong constraint on null subjects to only co-refer with the subject of the preceding sentence. The other forms were rated as largely acceptable but to different degrees. Our focus here however is on the interaction between antecedent salience and reference form on acceptability.

To analyze these ordinal ratings, we used cumulative link mixed models (Agresti, 2012) with a cloglog linking function (chosen due to the positive skew of the ratings). Using R package *ordinal* (ver 2015.6-28), we analyzed the data using models with

Antecedent (Subject vs. Object) and Anaphor (Null, Overt Pronoun, Repeated Name) as fixed factors and items and subjects as random factors. This analysis showed that the interaction of the two fixed factors made a significant contribution to model fit, $\chi^2(2) = 27.458, p < .001$. The parameter estimates of the chosen model are shown in Table 3.

We again used treatment coding with the reference levels Subject and Null for the Antecedent and Anaphor factors respectively so as to facilitate the examination of the differences between the three reference forms in the subject antecedent conditions. The coefficients in Table 3 thus represent differences from this reference level. As can be seen in the table, when the antecedent was the subject, overt pronouns were rated as more acceptable than null pronouns but repeated names were not. As can also be seen in the table, the two interaction coefficients were significant, indicating, not surprisingly, that null pronouns were even less preferred in the object antecedent conditions.

Finally, to directly test whether the meta-linguistic judgments expressed in the ratings can account for the reading times patterns we observed in Experiment 1, we reran the analysis of Experiment 1 after adding as a predictor the Z-transformed average rating across all subjects for each item in each condition. It is noteworthy that while the ratings were recorded on an ordinal scale, their average across subjects provides a continuous interval scale measure that can be included as a meaningful continuous regressor. Model comparisons indicated that although Rating did not interact with Antecedent or Anaphor, $\chi^2(2) = 5.628, p = .344$, it did have a main effect in improving model fit, $\chi^2(1) = 6.808, p = .009$. The coefficients of the model with Rating included are shown in Table 1, where it can be seen that the addition of Rating to the model mitigated the effect of the RNP (the Antecedent-Object:Anaphor-Name interaction is no longer significant) but did not change the significance of any of the coefficients involving overt pronouns. Critically, the

Antecedent-Object:Anaphor-Pronoun interaction coefficient, which represents the OPP, was still significant. Finally, the coefficient representing the difference in reading times between repeated names and null pronouns when the antecedent was the subject was significant in this analysis whereas it was only marginally significant in the original analysis.

Insert Table 3 about here

Discussion

The results of this experiment reveal that BP speakers prefer overt pronouns to null pronouns in the subject position of a sentence referring back to the subject of the previous sentence. This result becomes accentuated when the antecedent is in object position. Thus, while the null subject reference is much less acceptable than the overt pronoun for references to the object of the previous sentence, mirroring the reading times in the object conditions, it is also less acceptable for references to the subject of the previous sentence, in contrast to the reading times in the subject conditions.

Because this finding was obtained with the same items we used in Experiment 1 and with participants from the same population, the OPP we observed in that experiment cannot be attributed to the acceptability of the different conditions. Together, the results of the two experiments clearly show that despite the shift in both use and acceptability of null-subject pronouns, the OPP still occurs in BP. We note here that, in contrast to overt pronouns, repeated names were not rated as significantly more acceptable than null pronouns when the antecedent was the subject. In addition, adding the ratings as a predictor to the analysis

of the reading times from Experiment 1, eliminated the RNP but not the OPP. We discuss this result in depth in the General Discussion.

Corpus study

Our finding of an OPP in BP is remarkable because it is inconsistent with BP speakers' metalinguistic preferences, as revealed by Experiment 2, as well as with the documented shift away from null subjects in BP (Cavalcante & Duarte, 2008; Duarte, 1995, 1996, 2000, 2003; Kato, 2000; Paredes Silva, 1988; Tarallo, 1983). As such, this result appears to go against usage-based accounts of language processing (e.g., MacDonald, 2013).

According to such accounts, the decreasing rate of null-subject pronouns in BP should have resulted in greater difficulty in the processing of these forms in comparison to the processing of the more frequent overt pronouns. However, it may be that the intuitions of speakers as well as the previous work on null subjects in BP have underestimated the actual frequency of these forms in the language. If null subjects are still frequently encountered, then our results are not squarely incompatible with usage-based views. Therefore, the last question we address in this paper is whether it is indeed the case that null subjects are overall less frequent than overt subjects in BP in constructions similar to the ones we tested here, and whether this holds true across different language genres.

Previous studies on pronoun use in modern BP have argued that null subject pronouns have already become infrequent in BP. However, the corpora used in many of these studies suffer from limitations in both size and representativeness. For example, the corpus used by Cavalcante & Duarte (2008) had less than 300 samples and the one used by Duarte (1996) consisted of less than 200 samples per historical period, gathered exclusively

from theater plays.

A recent study (Smith, 2013) addressed some of these limitations by examining the 1900s BP section of the larger Corpus do Português (CDP; Davies & Ferreira, 2006), which is the largest fully annotated on-line corpus of the Portuguese language, consisting of more than 45 million words in nearly 57,000 European and Brazilian Portuguese texts from the 14th to the 20th century sampled from four genres (academic, newspapers, fiction, and spoken). Smith found that, overall, about 30% of independent non-embedded sentences in the corpus contained null subjects with the remaining 70% having an overt subject of some form, and that the distribution varied across genres.

While the results of Smith's analysis may be informative about the general frequency of null subjects in BP, they are not informative for our purposes because Smith's sample was based on grammatical structures that are different from the simple transitive sentences that we are interested in and that we used in our experiments. In particular, his sample included cases in which the null subject is still commonly or even obligatorily used in BP, possibly resulting in an overestimation of the number of null subjects. A second reason is that Smith only distinguished between sentences that had no overt subject and sentences that did. This could have resulted in underestimating the number of null pronouns since sentences with grammatically complex subjects were counted as overt subjects. These are clearly different from the overt pronouns and the repeated names we used in our experiments and may be subject to very different processing and pragmatic constraints. A third reason is that, while verb type was not a sampling criterion in Smith's study, there is evidence that the main verb of a sentence is likely to influence referential form choice (e.g., Caramazza et al., 1977; Kehler et al, 2008).

To assess the frequency of null and overt subject forms in structures that closely

resemble the ones used in our BP experiments, one should preferably focus on the verbs used in these experiments. We therefore conducted a corpus study in which we measured the frequency of repeated names, overt pronouns, and null pronouns in BP sentences that contained the same verbs we used in Experiments 1 and 2. We repeated this study twice, first looking at the overall frequency of the different forms appearing in subject position in the corpus, and second, looking at only those instances in which the forms of interest were used as anaphors.

Methods

Materials. Because the CDP distinguishes between Brazilian and European Portuguese only for texts published in the 20th century and because this section of the corpus is not further annotated for time, we analyzed only the 20th century Brazilian subsection of the CDP. This partial corpus consisted of around 10 million words, of which 1 million words were from the spoken register and the remaining 9 million were evenly divided between the academic, newspaper and fiction corpus genres.

The spoken section was collected from radio and television interviews, as well as from one-on-one conversations conducted by the CDP creators (Davies & Ferreira, 2006). The other genre sections were gathered from a variety of academic journals/textbooks, newspapers, and literary works, respectively. Also, with the exception of the fiction genre, the majority of the 20th century section is comprised of texts from the 1980s and on (Smith, 2013).

Procedure. The CDP is lemmatized and tagged for part of speech but is not parsed. In addition, the interface only allows searches for words, word classes, lemmas, and phrases. Consequently, in order to retrieve words with specific grammatical functions (e.g., subject), it was necessary to use one of the above categories as the search term, and then

manually parse the search results (Smith, 2013).

In order to assess the frequency of overt and null subjects in structures similar to the ones employed in our experiments, we performed separate searches for each of the 36 transitive verbs we used in Experiments 1 and 2. Each query was done with a search term consisting of the target verb in the simple past, as it appeared in the stimuli set, plus a wildcard to allow for the possibility of weak pronouns attached to the verbs (e.g., “venceu-*”). From the search results, we selected sentences with the plain version of the verbs and with the verbs attached to the weak pronouns (“o, a”) as used in our experiments (e.g., “venceu,” “venceu-o,” “venceu-a”). This process resulted in a set of 4,624 sentences, which were then processed as follows.

Of the initial sample, 12 sentences were excluded because they either (i) lacked an expanded context that would allow the identification of the expression in subject position; or (ii) represented duplications of the same sentence in the CDP. Additionally, 1,182 sentences were removed because they had grammatical subjects that differed structurally from those used in our experiments (e.g., embedded and relative clauses, coordinated sentences, cleft structures, passive structures, wh-questions, indefinite and demonstrative pronouns).

Our final filtered sample consisted of 3,430 sentences, which were manually classified on the basis of the expression appearing in subject position (noun phrase, overt pronoun, or null pronoun). The category “noun phrase” included proper names and other types of definite and indefinite descriptions. The “overt pronoun” category included all attested instances of personal pronouns (regardless of person) and the “null pronoun” category included only null pronouns of definite reference. These choices were motivated by our desire to generate a reasonably sized corpus that represented the general distribution

of null pronouns, overt personal pronouns, and fuller noun phrases in contexts similar to the ones used in our experiments. Focusing on only proper names and third-person pronouns would have resulted in an impractically small sample.

We also created a sub-corpus consisting of sentences in which the subject was an anaphor. The criterion we used to identify these cases was the presence of an overt, direct antecedent in the expanded context (172 words on average) retrieved from the CDP for each sample. Indirect anaphoric cases, in which there was no considerable formal overlap between the head-noun of the anaphor and the head-noun of its antecedent, were excluded from our anaphoric sub-corpus. This anaphor-only sub corpus consisted of 1,845 sentences.

All data filtering and categorization was performed by one of the native BP speaking co-authors of this paper (JCM).

Insert Table 4 about here

Insert Figure 3 about here

Insert Figure 4 about here

Results

The raw frequencies and relative percentages of the subject forms investigated in all genres of the CDP are shown in Table 4. Because previous studies found considerable differences between genres within the corpus, we included genre as a factor in our analyses.

To test whether differences between the frequency counts of null pronouns, overt pronouns, and noun phrases were significant within the different genres, we performed a linear mixed-effects analysis on log transformed frequencies with Subject Form, Genre, and their interaction as fixed effects and verb as a random factor (Baayen, 2008). Since models with random slopes did not converge, our final model only included an intercept effect of the random factor. The significance of each fixed main and interaction effect was assessed by means of model comparison between the full model and an identical model without the effect in question (Baayen, 2008). We report the analyses for the full and anaphor-only corpora separately.

For the full corpus, the analyses showed a main effect of genre, $\chi^2(9) = 90.721, p < 0.001$, and of subject form, $\chi^2(8) = 84.931, p < 0.001$, as well as a significant interaction between the two factors, $\chi^2(6) = 16.322, p < 0.05$. The complete data are shown in Figure 3. Tukey contrasts of pairwise differences in frequency revealed (a) no significant differences between the frequency of null pronouns and noun phrases in any genre (academic, $Z = -0.508, p = 1.00$, spoken, $Z = -1.842, p = 0.79$, news, $Z = -2.481, p = 0.34$, and fiction, $Z = -0.420, p = 1.00$); (b) greater frequency of null pronouns than overt pronouns in all genre (academic, $Z = -3.996, p < 0.01$, news, $Z = -4.089, p < 0.01$, and fiction, $Z = -5.64, p < 0.01$) except for the spoken genre, $Z = 0.087, p = 1.00$.

For the anaphor-only corpus, the analyses similarly showed a main effect of genre, $\chi^2(9) = 74.146, p < 0.001$, a main effect of subject form, $\chi^2(8) = 48.482, p < 0.001$, and a significant interaction between the two factors, $\chi^2(6) = 13.037, p < 0.05$. The data from the anaphor only corpus are shown in Figure 4. Tukey contrasts of pairwise differences revealed that (a) the frequency of null pronouns was not significantly different than that of noun phrases in any of the genres analyzed (academic, $Z = 1.115, p = 0.9916$, spoken, $Z =$

0.054, $p = 1.00$, news, $Z = -1.186$, $p = 0.9893$, and fiction, $Z = 1.702$, $p = 0.8602$); (b) The frequency of null pronouns was significantly greater than the frequency of overt pronouns in the academic, $Z = -3.441$, $p < 0.05$, and fiction genre, $Z = -4.777$, $p < 0.01$, but not in the news, $Z = -2.763$, $p = 0.1876$, or spoken genres, $Z = -0.45$, $p = 1.00$. Thus, the only difference between the full corpus and the anaphor-only corpus is that null pronouns were more frequent than overt pronouns in the news genre in the former but not in the latter.

Similar to the meta-linguistic judgments from Experiment 2, we tested whether the language patterns reflected in the corpus data can account for the reading time results of Experiment 1. We thus reran the reading time analysis of Experiment 1 with corpus bias calculated for each verb and subject form as an additional predictor. We repeated this twice, once with biases calculated on the basis of the full corpus, and once with biases based on the anaphor-only corpus. We only included items with verbs which had more than 30 occurrences in the full corpus (25 verbs). Corpus bias did not contribute significantly to model fit in either the full corpus analysis, $\chi^2(6) = 3.809$, $p = .70$, or the anaphor-only corpus analysis, $\chi^2(2) = 3.348$, $p = .76$. Unfortunately, our corpus was too small for separate similar analyses for individual genre.

Discussion

Our results revealed considerable variability between genres in the relative distribution of null vs. overt subjects. In our sample, the spoken and newspaper genres showed the strongest preference for overt forms over null pronouns. This is not surprising as spoken language is most likely to show the recent changes in BP. By the same token, formal academic writing is the least likely to show these changes. Since newspapers tend to use less formal language that generally resembles spoken language, and since the newspaper samples in the corpus are from the later part of the 20th century the newspaper

samples were more similar to the spoken samples than to the academic samples. Because the fiction section in the corpus consists of mixed work from all periods from the last century, it is likely less representative of the recent changes in BP.

Overall, these corpus findings support the general claim that BP is shifting towards overt subjects. However, our corpus analyses also suggest that this change may not have been as dramatic as linguists have argued in that null pronouns are still frequently encountered and overt pronouns are relatively rare. In the genres that are most likely to reflect recent changes, and in which null pronouns have become rarer, it is not overt pronouns but rather other overt forms that are taking the place of null pronouns.

The corpus results also raise the possibility that our experiments, which involved students in an academic institution tested in a university lab, may have elicited a processing pattern that is appropriate for an academic setting, in which null subjects are still frequently encountered. This interpretation, however, does not explain why our participants showed an OPP despite the similar frequencies of null and overt subjects in the academic genre, and why their acceptability ratings in Experiment 2, which were collected in the same environment, showed a different pattern.

One important limitation of our corpus analysis is that we did not distinguish between anaphors with subject and object antecedents. Unfortunately, the spoken genre, which being most recent is the most relevant part of the corpus for our question, was not large enough to allow for a meaningful analysis of the role of antecedent status. We thus note this limitation which we hope to address in a future study.

Finally, corpus biases did not account for any variance in the reading times in Experiment 1. While we do not consider this null finding very informative, we report it because we feel that readers would be interested to know what this analysis might reveal.

General discussion

The experiments reported in this paper showed that the OPP occurs in BP despite claims by linguists that null subjects are no longer frequently used, and despite speakers' intuitions that agree with these claims. Our corpus analyses, however, showed that null subjects are still frequently encountered and that overt pronouns are still relatively rare in BP.

Experiment 1 found a robust OPP in BP. There was also a clear RNP, but the difference between null pronouns and repeated name anaphors with a subject antecedent was not significant, suggesting that the RNP in BP may be weaker than the OPP. Experiment 2, in turn, confirmed that BP speakers from the same population we tested in Experiment 1 considered null pronouns less acceptable than overt pronouns (but not repeated names) when making reference to the subject of the previous sentence. Finally, our corpus study showed that while null subjects are still found in BP in constructions similar to the ones we tested, they appear more frequently in academic writings and works of fiction and less frequently in news and spoken language.

This last finding supports the claim that BP is changing towards using overt subjects (full noun phrases or overt pronouns) and suggests that this transition is ongoing. The fact that overt pronouns were least frequent in academic writings allows usage-based theories to account for the OPP because our testing was carried out in an academic environment. However, this leaves open the question of why participants found the null pronoun items less acceptable in the same academic environment. Nevertheless, we recognize that the question of whether the OPP would occur in BP in non-academic settings is important and should be addressed by future research.

Overall, our findings provide the following answers to the three initial questions we set out to investigate. First, they show that the RNP occurs in other languages besides English, Spanish, and Chinese, and therefore lends further support to the claim that this phenomenon may be universal in nature. Second, they show that the OPP is not unique to Spanish and occurs in other null-subject Romance languages. Finally, although our results are not outright incompatible with usage-based accounts, they present a serious challenge to such accounts.

In particular, the results of Experiment 2 show that the OPP is not a processing delay that merely reflects metalinguistic intuitions. Because usage-based accounts attribute both acceptability judgments and processing during comprehension to the same underlying statistical patterns, it is unclear how extant usage-based theories can account for all our findings. In the following sections, we attempt to offer an account of how usage-based and pragmatics-based accounts can be profitably combined to explain our body of results.

Pronoun avoidance in BP

The most remarkable aspect of our findings is that the OPP occurs in BP, despite the shift in preferences in favor of overt subjects. The preference for overt subject anaphors is supported by previous work in BP linguistics (Cavalcante & Duarte, 2008; Duarte, 1995, 1996, 2000, 2003; Kato, 2000; Paredes Silva, 1988; Tarallo, 1983) and by the ratings of speakers (Experiment 2) from the same population we tested in Experiment 1. To some extent, this preference is also supported by our corpus study, which showed that the use of the null pronoun is least common in spoken language and in news. As the least formal genre we examined, speech is most likely to reflect ongoing linguistic changes. It should be noted, however, that the overt personal pronoun is still overall quite infrequent in BP, as

shown by our corpus analysis. This suggests that, instead of switching to overt pronouns, BP speakers may prefer other forms of full reference.

We suggest that, as BP undergoes change, speakers are responding to two separate constraints. The first is to avoid the null pronoun. This constraint results from the recent change to the language and is strongly reflected in metalinguistic judgments by both linguists and experimental subjects, and more weakly in available corpora. The second constraint is to avoid the full pronoun. This implicit constraint may reflect a lingering processing strategy (Duarte, 1995; Chomsky, 1981), or some remnants of a former pragmatic function of the overt pronoun, and is reflected most strongly in actual production (as shown by our corpus study), and in whole-sentence reading times during comprehension (as demonstrated by the OPP in Experiment 1), but not in metalinguistic judgments of speakers and linguists. This interpretation is consistent with usage-based theories, in that avoiding overt pronouns could be a production constraint that is then mirrored by the comprehension system because of the statistical patterns it creates.

We now turn to pragmatic explanations such as the ILH (Almor, 1999; Almor & Nair, 2007). According to such theories, the processing of more informative references is initially costlier than the processing of less informative expressions. Despite this difference in initial processing cost, more informative referential forms can be processed more easily overall (as reflected in whole sentence reading times) when their additional cost is justified by some discourse function that could not have been achieved by a less informative expression. Gelormini-Lezama and Almor (2011) applied this explanation to the OPP in Spanish and argued that the OPP reflects the use of an overly informative expression (the overt pronoun) when it serves no discourse function.

While the same explanation could account for our BP findings, a possible objection

could be raised in that the recent changes to the BP pronominal system and verb morphology could be argued to have resulted in the overt pronoun becoming more functional in BP for resolving referential ambiguities in contexts where a null pronoun would be sufficient in the Spanish equivalent. By this argument, in BP contexts like the ones we used, in which the potential referents are of opposite genders, and in which the verb carries no gender information, overt pronouns could now serve a new function of helping to identify the correct antecedent, and should, as a result, not elicit the OPP. This objection however is problematic for two reasons. First, it ignores the fact that subjects antecedents are the “default” antecedents and are not likely to require additional agreement information to be identified as the antecedent of a subject anaphor in the next sentence. Second, previous research questioned the importance of the information carried by overt pronouns for disambiguating an antecedent, although this research has focused on intra- rather than inter-sentential discourse anaphora (e.g, Cameron, 1992; Filiaci, 2001). However, because our experiments did not directly manipulate referent ambiguity this explanation will have to remain speculative until more data are collected.

Whether or not overt pronouns in BP have gained increased discourse functionality , we argue that the processing cost of overt pronouns has also increased. According to our hypothesis of a “pronoun avoidance strategy” reflecting the transitory state of BP, overt pronouns would accrue a greater processing cost, and it is this high processing cost and not the lack of functionality that would underlie the OPP in BP. This type of explanation has not been previously considered in pragmatic frameworks of reference processing and thus requires the extension of the notion of processing cost in such theories. We propose that in cases of ongoing language change, usage frequency may become more important than the discourse and memory constraints that normally underlie processing cost. In this revised

framework, both usage-based constraints and pragmatic constraints can operate at the same time and their relative influence can be modulated by processes of language change. In the present case of BP, one of the by-products of this process would be the strengthening of the influence exerted by usage-base constraints in relation to memory and discourse constraints. Note that although our corpus analysis was not fine-grained enough reveal the frequency of overt pronouns specifically when their antecedent was in the subject position, the percent of overt pronouns was low overall, suggesting that they were rare following any type of antecedent.

In summary, we propose that BP speakers and readers implicitly follow a pronoun avoidance strategy in actual processing but not when making meta-linguistic judgments. Incorporating this strategy allows both usage-based and pragmatic theories to account for the OPP results. For usage-based theories, pronoun avoidance during production results in statistical patterns that underlie comprehension difficulty. We, however, propose that the statistical patterns generated by the pronoun avoidance strategy could at the same time be seen as generating an increased processing cost in pragmatic theories. Thus, unlike “pure” usage-based accounts, which attribute comprehension difficulty to input statistics, our revised pragmatic framework attributes comprehension difficulty to a combination of input statistics that can affect initial processing cost, and pragmatic principles. As we will argue next, this revised pragmatic framework provides a more complete and informative account of the entirety of our results than do usage-based theories.

The OPP vs. the RNP

In addition to an OPP, we also observed an RNP in our first experiment. It should be noted that the reading-time differences between the sentences with repeated name anaphors and sentences with null pronouns were not quite significant for subject

antecedents in Experiment 1. As a null result, we should exercise caution before interpreting it as reflecting the absence of a difference. In addition, we note that it is an interaction pattern rather than any differences between two means that is the marker of a clear RNP (Gelormini-Lezama, 2012). However, we also note that once we added the acceptability ratings from Experiment 2 as a predictor in the analysis of the reading times in Experiment 1, the interaction representing the RNP was no longer significant. As a result, we can argue that the RNP is weaker than the OPP in BP, and also that the RNP but not the OPP, could reflect the same factors that are captured by metalinguistic judgements.

It is here that we note the limitations of usage-based accounts. The problem is that usage-based accounts will not simply predict that the RNP will be weaker, but rather that it will not exist at all. If indeed our OPP results are explained by the much higher frequency of null pronouns than overt pronouns in the academic genre, then the RNP should be explained by a higher rate of null subjects than that of nouns in that genre as well. However, this was not the case as the null and the repeated name items with subject antecedents had similar frequencies of occurrence. Moreover, in Experiment 2, repeated-name items with subject antecedents were rated as being as acceptable as null-subject items. Therefore, the same usage-based account that explained the OPP cannot accommodate the RNP in a straightforward manner.

Looking at the Spanish results of Gelormini-Lezama and Almor (2011), it becomes clear that, in Spanish too, the RNP is numerically smaller than the OPP. Indeed, the same difference in magnitude between the RNP and the OPP occurs in Spanish, in which the overt pronoun serves a specific contrastive-emphatic function, and in BP, in which the overt pronoun appears to serve a broader disambiguation function similar to English. Despite the similarity between BP and Spanish in that both show a weaker RNP than OPP,

we propose that the reasons are different in BP than in Spanish. In BP, following our “pronoun avoidance strategy,” we argue that the cost of processing overt pronouns is likely higher than the cost of processing repeated names. In contrast, in Spanish, we believe that the weaker RNP reflects an overall reduced discourse function for overt pronouns than for repeated names when the antecedent is salient (except when establishing contrast or emphasis). In lieu of such a special function, overt pronouns in Spanish serve absolutely no discourse role and this results in an overall stronger penalty for overt pronouns than for repeated names, which can serve some function, possibly in highlighting, foregrounding, or emphasizing information.

In short, in a cost-balance pragmatic framework, both the OPP and the RNP represent a high processing cost that is not balanced by discourse function, though this imbalance can have different degrees. In Spanish and BP, due to either lower functionality or increased cost, overt pronouns associated with subject antecedents lead to a stronger imbalance than repeated names.

Language change, processing, and metalinguistic intuitions

Tracking the dynamics of an ongoing change within a single language, as in present BP, can lead to valuable insights about the intricate relationship among language processing, acceptability, and statistical distribution. The contrast between the results of the rating task in Experiment 2 and the self-paced reading task in Experiment 1 suggests that, when language change is in progress, speakers’ linguistic perceptions and intuitions precede shifts in language processing. In the case of BP, we speculate that if the transition away from null subjects continues, the overt pronoun avoidance strategy will disappear, resulting in the disappearance of the OPP in this language. Tracking this process will be highly informative about the dynamics of language change.

The differences we found between reading times and acceptability ratings also highlight the danger of relying on only one source of evidence in linguistic research. The use of speakers' intuitions as the primary source of data is common practice in many subfields in theoretical and descriptive linguistics (e.g., generative grammar, lexical semantics). Recently, as the standard methodologies of data-driven disciplines have made their way into linguistics, the exclusive reliance on individual speaker intuitions has come under criticism. Carefully designed, collected and analyzed acceptability judgment tasks are often put forth as an alternative to intuitions in linguistic argumentation, (e.g., Wasow & Arnold, 2005). Our results, however, show that acceptability judgments, even when collected and analyzed with the most rigorous methods, may have some important limitations, especially when used in the context of language processing. Crucially, by showing a clear dissociation between acceptability, statistical patterns, and processing, our study reinforces the need to treat intuitions and speakers' judgments as only one source of data and to test linguistic hypotheses using multiple types of evidence and empirical methods (Mitchell, 2004; Schütze, 1996; Wasow & Arnold, 2005).

Conclusion

The results we report here highlight the importance of cross-linguistic comparisons for addressing questions about language processing in general and discourse reference in particular. Overall, this study shows that actual language processing in comprehension could lag behind changing preferences and speakers' intuitions about the acceptability of different referential patterns. This result has important implications for interpreting acceptability and metalinguistic judgments in the context of language processing and language change.

Authors' Note

The authors thank Pablo Arantes and members of the *a*Lab at University of South Carolina for helpful comments on earlier versions of this work. This research was partially funded by NIH R21AG030445 and NSF BCS0822617 grants to AA and by a Brazilian National Council for Scientific and Technological Development fellowship to JCM.

References

- Agresti, A. (2012). *Categorical Data Analysis* (3rd ed.). Wiley.
- Almor, A. (1999). Noun-phrase anaphora and focus: the informational load hypothesis. *Psychological Review*, 106(4), 748-765.
- Almor, A., & Nair, V. A. (2007). The Form of Referential Expressions in Discourse. *Language and Linguistics Compass*, 1(1-2), 84-99.
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.
- Baayen, R. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. New York: Cambridge University Press.
- Barbosa, P. (1995). *Null Subjects* (Doctoral dissertation). MIT, Cambridge, MA.
- Barbosa, P. (2000). Clitics: a Window into the Null Subject Property. In J. Costa (Ed.), *Portuguese Syntax: Comparative Studies* (pp. 31-93). New York: Oxford University Press.
- Barbosa, P., Duarte, M. E. L., & Kato, M. A. (2005). Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4, 11-52.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structures for conformatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278.
- Cameron, R. (1992). *Pronominal and null subject variation in Spanish: Constraints, dialects, and functional compensation*. Ph.D. dissertation. University of Pennsylvania. Philadelphia, PA.
- Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behaviour*, 16, 601-609.

Carminati, M. N. (2002). *The processing of Italian subject pronouns*. Ph.D. Dissertation.

University of Massachusetts Amherst. Amherst: MA.

Cavalcante, S. R. O., & Duarte, M. E. L. (2008). The subject position in Brazilian

Portuguese: the embedding of a syntactic change. *University of Pennsylvania*

Working Papers in Linguistics, 14, 54-62.

Chambers, C. G., & Smyth, R. (1998). Structural parallelism and discourse coherence: a

test of centering theory. *Journal of Memory and Language*, 39, 593–608.

Cho, H. Y. (2010). *Coreference processing in L1 and L2* (Doctoral dissertation). University

of Illinois at Urbana-Champaign, Champaign, IL.

Chomsky, N. (1981). Lectures in Government and Binding. (Studies in generative grammar

9.) Dordrecht: Foris.

Crawley, R., Stevenson, R., & Kleinman, D. (1990). The use of heuristic strategies in the

interpretation of pronouns. *Journal of Psycholinguistic Research*, 4, 245-264.

Cyrino, S. M. L., Duarte, M. E. L., & Kato, M. A. (2000). Visible subjects and invisible

clitics in Brazilian Portuguese. In M. A. Kato, & E. V. Negrão (Eds.), *Brazilian*

Portuguese and the Null Subject Parameter (pp. 55-73). Frankfurt: Vervuert-

Iberoamericana.

Davies, M., & Ferreira, M. (2006). *Corpus do Português*: 45 million words, 1300s-1900s.

Available online at <http://www.corpusdoportugues.org>.

Duarte, M. E. L. (1995). *A perda do princípio “Evite Pronome” no Português Brasileiro*

(Doctoral dissertation). Universidade Estadual de Campinas, São Paulo, SP.

Duarte, M. E. L. (1996). Do pronome nulo ao pronome pleno: a trajetória do sujeito no

português do Brasil. In I. Roberts, & M. A. Kato (Eds.), *Português Brasileiro: uma*

viagem diacrônica (pp. 107-125). Campinas: Editora da UNICAMP.

Duarte, M. E. L. (2000). The loss of the Avoid Pronoun Principle in Brazilian Portuguese.

In M. A. Kato, & E. V. Negrão (Eds.), *Brazilian Portuguese and the Null Subject Parameter* (pp. 17-36). Frankfurt: Vervuert-Iberoamericana.

Duarte, M. E. L. (2003). A evolução na representação do sujeito pronominal em dois

tempos. In M. C. Paiva, & M. E. L. Duarte (Eds.), *Mudança linguística em tempo real* (pp. 115-128). Rio de Janeiro: Contracapa.

Filiaci, F. (2011). *Anaphoric preferences of null and overt subjects in Italian and Spanish: A cross-linguistic comparison*. Ph.D. dissertation. The University of Edinburgh.

Edinburgh, U.K.

Filiaci F., Sorace, A. & Carreiras, M. (2013). Anaphoric biases of null and overt subjects in

Italian and Spanish: a cross-linguistic comparison. *Language and Cognitive Processes*,

Flores-Ferrán, N. (2004). Spanish subject personal pronoun use in New York City Puerto

Ricans: Can we rest the case of English contact? *Language Variation and Change*, 16, 49-73.

Frederiksen, J. (1981). Understanding anaphora: Rules used by readers in assigning

pronominal referents. *Discourse Processes*, 4, 323-347.

Galves, C. C. (1996). O enfraquecimento da concordância no português brasileiro. In I.

Roberts, & M. A. Kato (Eds.), *Português Brasileiro: uma viagem diacrônica* (pp. 387-406). Campinas: Editora da UNICAMP.

Garrod, S., & Sanford, A. J. (1982a). Bridging inferences in the extended domain of

reference. In A. Baddeley, & J. Long (Eds.), *Attention and Performance IX* (pp. 331-346). Hillsdale, NJ: LEA.

- Garrod, S., & Sanford, A. J. (1982b). The mental representation of discourse in a focused memory system: Implications for the interpretation of anaphoric noun-phrases. *Journal of Semantics*, 1, 21-41.
- Gelormini-Lezama, C. (2012). A review of the repeated name penalty: implications for null subject languages. *Revista Lingüística / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro*, 8(2), 22-34.
- Gelormini-Lezama, C. & Almor, A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Language and Cognitive Processes*, 26(3), 437-454.
- Gelormini-Lezama, C. & Almor, A. (2013). Singular and plural references in Spanish. *Journal of Psycholinguistic Research*, 43(3), 299-313.
- Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, 32, 99–156.
- Givón, T. (1987). *On understanding grammar*. New York, NY: 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.
- Gordon, P. C., & Searce, K. A. (1995). Pronominalization and discourse coherence, discourse coherence and pronoun interpretation. *Memory & Cognition*, 23, 313-323.
- Grosz, B. J., Joshi, A., K., & Weinstein, S. (1983). Providing a Unified Account of Definite Noun Phrases in Discourse. In Association for Computational Linguistics, 21st Annual Meeting of the Association for Computational Linguistics: Proceedings of the Conference (pp. 44-50). Cambridge, MA: Association for Computational Linguistics.
- Grosz, B. J., Joshi, A., K., & Weinstein, S. (1995). Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, 2(21), 203-225.

- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2), 274-307.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Holmberg, A., Nayudu, A., & Sheehan, M. (2009). Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, 63(1), 2009.
- Huang, Y. (1994). *The syntax and pragmatics of anaphora*. Cambridge: Cambridge University Press.
- Huang, Y. (2000). *Anaphora: A cross-linguistic study*. Oxford: Oxford University Press.
- Hubert, M., & Vandervieren, E. (2008). An adjusted boxplot for skewed distributions, *Computational Statistics & Data Analysis*, 52(12), 5186-5201
- Kato, M. A. (2000). The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. In M. A. Kato, & E. V. Negrão (Eds.), *Brazilian Portuguese and the Null Subject Parameter* (pp. 207-240). Frankfurt: Vervuert-Iberoamericana.
- Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, 25, 1-44.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Leitão, M. M. (2005). Processamento correferencial de nomes e pronomes em Português Brasileiro. *Revista Linguística (PPGL/UFRJ)*, 1, 235-258.
- Leitão, M. M., & Simões, A. B. G. (2011). A influência da distância no processamento correferencial de pronomes e nomes repetidos em português brasileiro. *Veredas (UFJF. Online)*, 1, 262-272.

- Leitão, M. M., Ribeiro, A. J. C., & Maia, M. (2012). Penalidade do nome repetido e rastreamento ocular em português brasileiro. *Revista Linguística / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro*, 8(2), 35-55.
- MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4, 1-16.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.
- Maia, M. A. R. (1997). The Processing of Object Anaphora in Brazilian Portuguese. *Revue Linguistique de Vincennes, França*, v. 26, p. 151-176.
- Mitchell, D. C. (2004). On-line methods in language processing: Introduction and historical review. In M. Carreiras & C. J. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs, and beyond*. Brighton, UK: Psychology Press.
- Montalbetti, M. (1984). *After Binding : On the Interpretation of Pronouns*. Doctoral Dissertation, MIT, Cambridge, Massachusetts.
- Nunes, J. M. (1996). Direção de ciclitização, objeto nulo e pronome tônico na posição de objeto em português brasileiro. In I. Roberts, & M. A. Kato (Eds.), *Português Brasileiro: uma viagem diacrônica* (pp. 207-22). Campinas: Editora da UNICAMP.
- Paredes Silva, V. L. (1988). *Cartas cariocas: a variação do sujeito na escrita informal* (Doctoral dissertation). Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.
- Prince, E. F. (1978). A comparison of wh-clefts and it-clefts in discourse. *Language*, 34, 883-906.
- Queiroz, K. L., & Leitão, M. M. (2008). Processamento do sujeito anafórico em Português Brasileiro. *Veredas – Revista de Estudos Linguísticos Online*, 2, 163-166.

- Roberts, I., & Kato, M. A. (1996). *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP.
- Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *SSLA*, 30, 333–357.
- Rohrbacher, W. (1999). *Morphology-driven syntax: a theory of V to I raising and pro-drop*. Amsterdam: John Benjamins.
- Sanford, A. J., & Garrod, S. C. (1981). *Understanding written language*. Chichester, England: Wiley.
- Schütze, C. (1996). *The empirical basis of linguistics*. Chicago: University of Chicago Press.
- Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23, 569–588.
- Smith, S. D. (2013). *Pro-drop and word-order variation in Brazilian Portuguese: A corpus study* (Master's thesis). Brigham Young University, Provo, UT.
- Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22, 339–368.
- Tarallo, F. (1983). *Relativization strategies in Brazilian Portuguese* (Doctoral dissertation). University of Pennsylvania, Philadelphia, PA.
- van-Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Wasow, T., & Arnold, J. (2005). Intuitions in linguistic argumentation. *Lingua*, 115, 1481–1496.
- Yang, C. L., Gordon, P. C., Hendrick, R., & Wu, J. T. (1999). Comprehension of referring expressions in Chinese. *Language and Cognitive Processes*, 14(5/6), 715–743.

Table 1: Sample Stimuli for Experiments 1 and 2

<u>Anaphor Form</u>	<u>Antecedent Salience</u>	
	Salient (subject)	Non-Salient (object)
Repeated Name	Carla desafiou Luís para uma partida de xadrez.	Luís desafiou Carla para uma partida de xadrez.
	Carla o venceu rapidamente e sem esforço.	Carla o venceu rapidamente e sem esforço.
	<i>Carla challenged Luís to a chess match.</i>	<i>Luís challenged Carla to a chess match.</i>
	<i>Carla beat him quickly and easily.</i>	<i>Carla beat him quickly and easily.</i>
Overt Pronoun	Carla desafiou Luís para uma partida de xadrez.	Luís desafiou Carla para uma partida de xadrez.
	Ela o venceu rapidamente e sem esforço.	Ela o venceu rapidamente e sem esforço.
	<i>Carla challenged Luís to a chess match.</i>	<i>Luís challenged Carla to a chess match.</i>
	<i>She beat him quickly and easily.</i>	<i>She beat him quickly and easily.</i>
Null Pronoun	Carla desafiou Luís para uma partida de xadrez.	Luís desafiou Carla para uma partida de xadrez.
	Venceu-o rapidamente e sem esforço.	Venceu-o rapidamente e sem esforço.
	<i>Carla challenged Luís to a chess match.</i>	<i>Luís challenged Carla to a chess match.</i>
	<i>NULL beat him quickly and easily.</i>	<i>NULL beat him quickly and easily.</i>

Table 2. Summary of Experiment 1 RT analyses.

	β	SE	df	t	p
Intercept	7.820	0.040	83.818	197.931	0.000 *
Antecedent-Object	0.212	0.030	187.563	7.016	0.000 *
Anaphor-Name	0.051	0.028	1416.974	1.822	0.069
Anaphor-Pronoun	0.095	0.028	1421.491	3.429	0.001 *
Antecedent-Object:Anaphor-Name	-0.120	0.039	1417.209	-3.050	0.002 *
Antecedent-Object:Anaphor-Pronoun	-0.125	0.040	1420.433	-3.158	0.002 *
<u>Model with Z-transformed Experiment 2 ratings as an additional predictor</u>					
Intercept	8.041	0.095	381.121	84.966	0.000 *
Antecedent-Object	0.156	0.037	252.939	4.244	0.000 *
Anaphor-Name	0.056	0.028	1420.552	2.009	0.045 *
Anaphor-Pronoun	0.108	0.028	1447.472	3.831	0.000 *
Rating	-0.039	0.015	348.831	-2.565	0.011 *
Antecedent-Object:Anaphor-Name	-0.070	0.044	1413.907	-1.577	0.115
Antecedent-Object:Anaphor-Pronoun	-0.086	0.042	1472.632	-2.038	0.042 *

Note: Only best fitting models shown. Both the Antecedent and Anaphor factors were coded using treatment coding. The null pronoun and subject antecedent were used as the reference levels (0) for the Anaphor and Antecedent factors respectively. Coefficients represent differences from the null pronoun/subject antecedent condition in log scale. Significant effects at a $p \leq .05$ level are marked with a *.

Table 3. Summary of Experiment 2 acceptability ratings analyses.

	β	SE	Z	p
Antecedent-Object	-0.526	.1746	-3.014	0.003 *
Anaphor-Name	-0.462	.1633	-2.832	0.005 *
Anaphor-Pronoun	-0.373	.1690	-2.210	0.027 *
Antecedent-Object:Anaphor-Name	0.556	.2199	2.528	0.011 *
Antecedent-Object:Anaphor-Pronoun	-0.956	.2190	-4.367	0.001 *

Note: Only best fitting models shown. Both factors were coded using treatment coding. The null pronoun and subject antecedent were used as the reference levels (0) for the Anaphor and Antecedent factors respectively. Coefficients represent differences from the null pronoun/Subject antecedent condition. Significant effects at a $p \leq .05$ level are marked with a *.

Table 4: Frequency Counts and Percentages of full NPs, Null Pronouns and Overt Pronouns in the CDP by Genre

<u>Genre</u>	<u>Whole corpus (N = 3,430)</u>						
	<u>Noun Phrases</u>		<u>Null Pronouns</u>		<u>Overt Pronouns</u>		<u>Total</u>
Academic	161	46%	177	51%	10	3%	348
Fiction	687	37%	1002	54%	158	9%	1847
News	669	63%	337	32%	58	5%	1064
Spoken	78	45%	51	30%	42	25%	171
Total	1595	46%	1567	46%	268	8%	3430
<u>Genre</u>	<u>Anaphor-only corpus (N = 1,845)</u>						
	<u>Noun Phrases</u>		<u>Null Pronouns</u>		<u>Overt Pronouns</u>		<u>Total</u>
Academic	79	37%	127	60%	7	3%	213
Fiction	313	29%	655	60%	115	11%	1083
News	260	54%	167	35%	55	11%	482
Spoken	22	33%	23	34%	22	33%	67
Total	674	36%	972	53%	199	11%	1845

Figure 1: Mean Reading Times for Experiment 1. Error bars show the Standard Error of the Mean.

(a)

(b)

Figure 2: Rating of items in the different conditions in Experiment 2. (a) histogram of percent ratings; (b) pairwise density plots of ratings in the different conditions. Condition median ratings are shown by vertical lines.

Figure 3: Least-square means of null pronouns, overt pronouns, and noun phrases in the linear mixed model (full corpus). Error bars show the Standard Error of the Mean.

Figure 4: Least-square means of null pronouns, overt pronouns, and noun phrases in the linear mixed model (anaphor-only corpus). Error bars show the Standard Error of the Mean.