**The Effects of Semantic Role Predictability on the Production of Overt Pronouns in Spanish**

Ana Medina Fetterman a\*, Natasha Vazquezb, Jennifer E. Arnoldc

aDepartment of Psychology and Neuroscience, UNC Chapel Hill, Chapel Hill, USA & Department of Psychology, OSU Main Campus, Columbus, USA; bcDepartment of Psychology and Neuroscience, UNC Chapel Hill, Chapel Hill, USA

Address all correspondence to: Jennifer Arnold, Dept of Psychology and Neuroscience, Davie Hall #337B, UNC Chapel Hill, Chapel Hill, NC, 27599-3270

**The Effects of Semantic Role Predictability on the Production of Overt Pronouns in Spanish**

In order to refer in any language, speakers must choose between explicit forms of expression, such as names or descriptions, or more ambiguous forms like pronouns. Current models suggest that reference form is driven by subjecthood, where speakers in English choose pronouns for the subject, and speakers of null pronoun languages like Spanish or Italian use null pronouns. We test this generalization by examining the effect of a different factor, thematic role predictability, on reference production in Spanish. In stories about transfer events (e.g., *Ana gave a ball to Liz)*, speakers prefer to use pronouns more for reference to goals (Liz) than sources (Ana; (Rosa & Arnold, 2017). However, this has not been examined for null pronoun languages. In two experiments, we demonstrate that Spanish speakers arealso sensitive to thematic role, but it primarily affects the rate of overt pronouns (ella, el) rather than null pronouns. These results highlight the need to include semantic constraints in models of reference production for null-pronoun languages.

# Introduction

Every time speakers refer, they have to decide how explicit to be, for example using a name (*Spot*) or a description (*the dog*), or an underspecified expression like a pronoun (*he*). While all languages offer a wide range of choices, languages vary in the specific forms available. In English, it is common to use a name or description when first mentioning someone (*Pamela washed her dog*), and use a pronoun for later reference (*Then she went to work*). By contrast, in null subject languages, such as Spanish or Italian, speakers tend to use null expressions in contexts where English speakers would use pronouns (*Pamela lavó su perro*. *Después Ø fue a trabajar / Pamela washed her dog. After* Ø went to work), although they can also use overt pronouns like “ella” (she) (…. *Después ella fue a trabajar*). Our study investigates how speakers choose referential expressions in Spanish. In particular, we examine the effect of semantic role, which has not figured prominently in models of reference processing for null subject languages.

The participants in events can be described in terms of their semantic, or thematic role, such as agent, patient, goal, source, and other roles. For example, in the transfer event in (1), Lady Mannerly is the source of the transfer, while Sir Barnes is the goal. Stevenson, Crawley, and Kleinman (1994) found that the structure of recent events creates expectations about who might be mentioned next, where characters in certain thematic roles are more likely to be mentioned than others. For example, following transfer verbs (as in example 1), people are more likely to continue talking about the goal (Sir Barnes in 1a; Lady Mannerly in 1b).

(1a.) Lady Mannerly gave the painting of the two of them to Sir Barnes.

(1b). Lady Mannerly borrowed a picnic basket from the chef.

In addition, some thematic roles – particularly those associated with transfer verbs – influence speakers’ decisions about how to refer, where pronouns are also used more for goals than sources (Arnold, 2001; Rosa & Arnold, 2017; Zerkle & Arnold, 2019). For example, in a story-telling task, speakers were more likely to use a pronoun to refer to goal characters (e.g., Lady Mannerly in 1b), than source characters (e.g., Lady Mannerly in 1a). Here we ask whether the same effect arises in Spanish. This question is relevant to two more general theoretical questions about pronoun production.

Our first question is whether reference production can be explained by purely syntactic constraints. Despite evidence for a goal bias, several studies on pronoun production in English suggest that reference production is driven entirely by a subject bias, where pronouns are preferentially selected for reference to the subject, and names or descriptions are used for objects and obliques (Fukumura & van Gompel 2010; Kehler, Kertz, Rohde, & Elman, 2008; Kehler & Rohde, 2013; Rohde & Kehler, 2014; see also van Rij, van Rijn, & Hendricks, 2012 for a similar model in comprehension). Support from this claim comes from several experiments in English in which the authors manipulate the predictability of the subject and nonsubject characters. For example, in *John amazed Bill*, participants expect that John was the cause of the event, whereas *John admired Bill* produces the opposite causality judgment (Brown & Fish, 1983). In this type of event, people tend to expect the implicit cause to be mentioned again, making it predictable. This predictability is associated with the thematic roles of the characters, where John is the stimulus in the *amaze* sentence, and Bill is the stimulus in the *admire* sentence. In both cases, the stimulus is the more expected referent, but several studies found that speakers have no preference to use pronouns more for the stimulus compared to the experiencer (Kehler & Rohde, 2013; Kehler et al., 2008; Kravtchenko, Modi, Demberg, Titov, & Pinkal, 2017; Rohde & Kehler, 2014).

By contrast, to our knowledge there is no work testing whether semantic constraints affect reference production in null subject languages. We therefore instead turn to research on comprehension, which emphasizes the role of the syntactic context. One clear-cut example of this is Carminati’s Position of Antecedent Hypothesis (PAH), which proposes that in Italian (a null subject language), pronoun comprehension is driven by only syntactic position, where null pronouns are preferred for reference to subjects, and overt pronouns are preferred for a lower syntactic position (Carminati, 2002). Violating this bias leads to difficulty in processing for comprehenders, and referential status, morphological or lexical properties are irrelevant.

Other comprehension research follows the broader principle that referential expressions are pragmatically constrained. Several models propose that reference form reflects the cognitive status of the referent, which has been described as accessibility, givenness, salience, or prominence (e.g. Ariel, 1990; 2001; Arnold, 2010; 2016; Chafe, 1976, 2004; Givón, 1983). Work on comprehension shows that more reduced forms (e.g., null pronouns) are appropriate for the most salient or accessible referents, whereas the more explicit forms are appropriate for reference to less accessible information (e.g., Ariel, 1990; Gundel et al., 1993); overt pronouns fall in the middle (Filiaci et al., 2013).

While numerous factors could potentially guide salience/accessibility, one of the constraints that has received substantial support from experimental work on pronoun comprehension is the role of subjecthood. For example, Gelormini and Almor (2011, 2013; Gelormini, 2018) demonstrate that in Spanish, readers slow down when an overt pronoun is used to refer to the prior subject, compared to sentences with a null pronoun. They term this effect the Overt Pronoun Penalty (OPP), and they interpret this as evidence that referential forms are pragmatically specialized, such that the more explicit forms are more felicitous when they refer to entities that are less prominent in the context (for similar findings see Alonso-Ovalle, Fernández-Solera, Frazier, & Clifton, 2002; Filiaci, 2011).

While this work focuses on comprehension, if the same factors affect production then we should expect that speakers should also follow a strategy of selecting null pronouns for subjects. On the other hand, recent work suggests that thematic roles also affect pronoun comprehension in Italian (Fedele & Kaiser, 2015), and that people have similar causality biases across multiple languages, including Spanish and Italian (Hartshorne, Sudo, & Uruwashi, 2013). This raises questions about whether semantic constraints also affect production.

Our second question is whether the same production model accounts for referential choices in all languages. Some previous work supports this idea. Several studies show that cross-linguistically, reference form choices tend to pattern with the information status of the referent: less-specified expressions, like pronouns or null reference, tend to be used for salient/accessible referents, while more explicit expressions tend to be used for less salient/accessible entities (e.g., Arnold, 1998 Gundel et al., 1993; Givon, 1983), which has led to theoretical proposals that referential forms are pragmatically specialized for cognitive status (Ariel, 1990; Gundel et al., 1993). While these proposals are not processing models of production per se, they have often been used to generate predictions for biases during language production. For example, as Arnold (2016) argues, a natural extension of these proposals would be to assume that during production, the cognitive status of the referent guides the selection of a referential form during production. A simple version of this model might propose that pronouns are selected when the referent is high in accessibility, while more explicit forms are selected when the referent is low in accessibility.

Yet other evidence suggests that this simple model would not work, because a single dimension of salience cannot account for referential choice in all languages. Kaiser & Trueswell (2008) demonstrated that Finnish has referential forms that follow different constraints, and moreover that they are influenced by more than salience level. Finnish is different from English in two ways that are relevant to referential form choice. First, it has two kinds of third person pronouns, one of which is gender neutral. Second, it has a flexible word order. Although the standard word order remains SVO, all six combinations can be admissible. Kaiser and Trueswell found that the pronoun *hän* is used primarily in reference to subjects, regardless of word order. In contrast, the demonstrative *tämä* is most commonly used in reference to post-verbal referents, although this preference is also modulated by syntactic role.

Of particular relevance to the current work, Filiaci et al. (2013) demonstrate that while null pronouns show a strong subject bias in both Spanish and Italian, overt pronouns are processed differently, where Italian overt pronouns are more strongly biased toward a non-subject referent than Spanish overt pronouns are. Thus, these results support a multi-dimensional approach, suggesting that different referential forms are constrained by different factors and at varying levels (Fedele & Kaiser, 2015; Filiaci et al., 2013; Kaiser & Trueswell, 2008). Different languages offer different types of referential forms, raising the possibility that different constraints apply cross-linguistically.

In sum, the current study tests whether thematic roles affect reference form in Spanish, which is relevant for shaping theories of reference production generally, given that some theories have either focused exclusively on structural constraints, or even explicitly claimed that thematic roles have no effect (e.g., Fukumura & van Gompel, 2010; Kehler et al., 2008; Kehler & Rohde, 2013; Rohde & Kehler, 2014). It is also relevant for understanding the extent to which reference form production is similar across languages. We test these questions by examining transfer verbs as in (2) in both written (Exp. 1) and spoken (Exp. 2) language.

2a.) Lady Mannerly le regaló una pintura de ellos dos a Sir Barnes.

(Lady Mannerly gave a painting of the two of them to Sir Barnes.)

2b) Sir Barnes recibió una pintura de ellos dos de Lady Mannerly.

(Sir Barnes received a painting of the two of them from Lady Mannerly.)

Here Lady Mannerly is the source of the transfer, and Sir Barnes is the goal. We control for the grammatical function of the referent by manipulating the verb; in 2a the source is in subject position, and in 2b the goal is in subject position. In both cases the other human character is mentioned second; for ease of exposition we refer to this position as the nonsubject.

***Experimental approach***

We examined reference form within a naturalistic narrative paradigm developed by Rosa & Arnold (2017) for English (stimuli available at https://jaapstimuli.web.unc.edu/). We involved participants in a role-playing game, where they were assigned the role of a tabloid photographer and instructed to describe the events occurring in a series of pictures. The pictures showcased a Clue-like murder mystery, and featured the same six participants throughout the study. There were three main male characters. (Sir Barnes, the chauffeur, and the butler), and three female (Lady Mannerly, the chef, and the maid; see Figure 1).



Figure 1. Characters in the event-retelling paradigm (from left-to-right: The butler and the maid, Sir Barnes, the chef, Lady Mannerly, the chauffeur).

The aim of this paradigm was to engage participants and provide a rich discourse context, which was hypothesized to strengthen their representations of discourse status. Thus, participants saw a series of experimental items, consisting of a pair of pictures and a prompt. This maintained the item-level control that is typical of psycholinguistic experiments. However, an innovation of this task was that the items together told a story about a small set of characters for the purpose of establishing the kind of rich discourse context that is typical of natural language. For further motivation of this task see Rosa & Arnold (2017).

One potential concern for our study is the variability across dialects, given that overt pronouns are more common in some dialects or regions than others (Lipski, 2002; Otheguy, Zentella, & Livert, 2007; Pešková, 2013). Typically, varieties of Spanish that use more verbal inflection, such as Spanish from Spain and Mexico, produce overt subject pronouns at a less frequent rate than Caribbean dialects, for example (Lipski, 2002; Pešková, 2013). Moreover, it is unclear if the use of overt pronouns is controlled by the same semantic conditions in different Spanish dialects. Both of our studies include speakers from multiple geographic regions, so we also provide secondary analyses that examine only those participants who produced one or more overt pronoun.

Another challenge for our experiment was the issue of bilingualism. While the participants of our online experiment all resided in Spanish-speaking countries, participants in Experiment 2 generally resided in the United States, and thus many of them spoke English or another language. Previous research suggests that children who are learning a null-subject language (like Spanish or Italian) at the same time as an overt subject language (like English) often overuse overt pronouns (Serratrice, 2008). Other research suggests that contact with English and generational differences among Spanish speakers living in the United States can lead to changes in Spanish pronoun production (Otheguy et al., 2007). In a related phenomenon, Basque-Spanish bilinguals were more likely to use null object pronouns than Spanish monolinguals (Sainzmaza-Lecanda & Schwenter, 2017). Since bilingualism is not the focus of the current study, we took steps to ensure that our participant samples were dominant Spanish speakers. In Experiment 2, we restricted our sample to only those people who had lived in a Spanish-speaking country until age 16, to ensure that all speakers were native speakers. Although the subjects in Experiment 2 likely also had extensive English exposure, the participants in Experiment 1 were all in Spanish-speaking countries at the time of their participation.

# Experiment 1

## Methods

### Participants

78 individuals participated in a one-hour online Spanish experiment posted on Amazon Mechanical Turk. Three additional submissions were ignored because they came from the same IP address as another submission. Each individual received $1 for their participation. Participation was restricted to individuals with IP addresses from countries or territories with Spanish as the official language. IP addresses from 15 different countries were represented in the final data set[[1]](#endnote-1).

35 participants were excluded from analysis. Our first criterion was that participants needed to have at least 2 trials with reduced expressions (grouping null and overt pronouns together for this constraint), and at least 2 with name/descriptions. We know that there is individual variability in the tendency to use reduced forms in this paradigm (Zerkle & Arnold, 2017a, 2019), where some participants pick a single reference form and stick with it for the entire experiment. We therefore limited our analysis to only those participants who used at least 2 reduced expressions (either null or overt pronouns), which ensures that our sample includes adequate variation to test our hypotheses (for a similar convention, see Buschkuehl, Hernandez-Garcia, Jaeggi, Bernard, & Jonides, 2014; Filmer, Mattingley, & Dux, 2015), We excluded 25 participants who used only name/descriptions on all critical trials, and 5 who used only one reduced form. 4 participants were excluded because fewer than 50% of their critical trials were usable (e.g., some participants re-wrote the prompt sentence, or produced ungrammatical or inappropriate responses). 1 participant from New York did not fit our eligibility criteria. 43 participants were included in the analysis, 22 from List A and 21 from List B.

### Materials and Design

In the main task, participants viewed pairs of picture, where the first picture was accompanied by a written sentence. They were instructed to read the first sentence of each pair and provide a written continuation that described what was occurring in the second image. They viewed two pictures like the ones shown in Figure 2 below.

Figure 2. Sample trial.

The pictures together told a story about six characters (Sir Barnes, Lady Mannerly, the chef, the chauffeur, the butler, and the maid). There were 53 trials in total, 24 of which were critical trials. Critical trials had images depicting actions with two of the characters. Filler trials described events occurring between one to three characters. Participants were exposed to one of two lists, either List A or List B. In both lists, participants saw the same illustrations, but read one of two versions of the given sentence. Materials for this experiment are available at https://....

Each of the critical trials depicted a transfer event with a goal and source character. Half of these (n=12) showed the goal character doing something in the second picture, these were termed “goal continuation” items. The “source continuation” items (n=12) showed the source character doing something in the second picture. Thus, source vs. goal continuation was manipulated between items. Within each item we manipulated the form of the first sentence to determine whether the goal or source was in subject position; see Table 1.

Table 1. Example stimuli

Goal continuation example (picture 2 shows Sir Barnes throwing the painting in the closet)

Source-goal verb:

Lady Mannerly le regaló una pintura de ellos dos a Sir Barnes.

(Lady Mannerly gave a painting of the two of them to Sir Barnes.)

Goal-source verb:

Sir Barnes recibió una pintura de ellos dos de Lady Mannerly.

(Sir Barnes received a painting of the two of them from Lady Mannerly.)

Source continuation example (picture 2 shows the butler lighting a fire)

Source-goal verb

El mayordomo le sirvió café a la criada.

 (The butler served coffee to the maid.)

Goal-source verb

La criada aceptó un café del mayordomo.

(The maid accepted a coffee from the butler.)

Thus, there were four key conditions in the experiment: (1) goal, subject reference; (2) goal, nonsubject reference; (3) source, subject reference; and (4) source, nonsubject reference. In half of each category, the two characters had the same gender, and in the other half they had different gender. This controls for the known tendency in English to use pronouns more often when the pronouns are unambiguous, in the different-gender condition. This resulted in a 2 (subject vs. nonsubject) x 2 (goal vs. source), x 2 (different vs. same-gendered characters) design. Thematic role and gender were manipulated between items, such that 6 of our 24 stimuli fell in each category (goal/same-gender; goal/different-gender; source/same-gender; source/different gender. Subjecthood was manipulated within item, such that there were two versions of each context sentence. The items were rotated across these two conditions and combined into two lists. In sum, each participant saw 3 items in each of the 8 conditions, for a total of 24 critical items.

Our dependent variable measured the referential form used to refer to the target character in the participant’s continuation. We coded the first clause of each continuation, classifying the form of the grammatical subject in one of three categories: description/name [El mayordomo cantó / The butler sang], overt pronoun [él cantó / he sang], or null pronoun [Ø cantó / Ø sang]. Trials were considered unusable if the subject NP referred to the wrong referent (a referent other than the character mentioned first) or to multiple referents. We also excluded continuations with verbs like gustar (“to like”), in which the target was most naturally continued as the object of the verb instead of the subject[[2]](#endnote-2). Responses were also excluded for using a *who* or *which* pronoun instead of a personal pronoun (*ella* (she) or *él* (he)), using quoted dialogue, or incomplete or missing responses.

### Procedure

Participants were directed to a Qualtrics survey and presented with a short introduction to their role as a tabloid photographer who had unknowingly witnessed a murder in a mansion. All aspects of the survey were in Spanish. Participants were presented with pairs of illustrations representing the pictures taken at the mansion and asked to reconstruct the story. The first illustration of each pair, which always had two characters in the image for critical items, was accompanied by a sentence describing the event. Participants were instructed to provide a written continuation for the second illustration of each pair, which contained only the target character.

## Results

The primary aim of our research was to determine if semantic roles affect the use of referential form in Spanish language production. We therefore modeled our data in terms of three binary choices: 1) name or description vs. other; 2) overt pronoun vs. other; 3) null pronoun vs. other. Data was analyzed using multilevel logistic regressions with SAS proc glimmix, with centered predictors and maximal random effects. If the model estimated a slope to be zero, it was removed from the model and noted below the table reporting the model’s results. Although participants were included as long as 50% of their responses were usable, participants contributed on average 86% of their trials. The total number of items included in the final analysis was 886. All statistical results are provided in the Appendix, and the significance patterns are summarized in Table 2.

As shown in Figures 3 and 4, participants used null pronouns more often to refer to the subject than to the nonsubject characters, and conversely, used overt pronouns and descriptions more for the nonsubject than subject characters. This emerged as a main effect of subjecthood for all three dependent measures.

In addition, we saw a smaller effect of thematic role. This effect emerged most strongly for overt pronouns, which were more likely for reference to goals than sources in nonsubject position, but were highly infrequent overall for reference to subject characters. This emerged as an interaction between subject and goal in the pronoun model. We used estimate statements to probe the effect of thematic role (goal minus source) for both subjects and objects, and confirmed that there was a significant goal effect for nonsubjects (β = 0.99 (SE = 0.41), t = 2.4, p = .02, but not for subjects (β = -0.85 (SE = 0.52), t = - 1.64, p = 0.11).

There was also a three-way interaction in the name/description model. Again we used estimates to probe the effect of thematic role for each of the other conditions (subject/same-gender; subject/different-gender; nonsubject/same; nonsubject/different). However, the goal effect was not significant in any of these conditions (|t|’s < 1.76; p’s > .09, suggesting that it had a limited effect on the production of names/descriptions in this experiment.

 Our gender manipulation (different vs. same) did not have a significant effect, but there was a marginal tendency to use more names in the same-gender condition, and more overt pronouns in the different-gender condition. These effects are in the predicted direction for a disambiguation strategy, since overt pronouns specify the gender of the referent.

Figure 3. Experiment 1, overall pattern of responses

Table 2. Experiment 1 results: Summary of effects from three models

|  |  |  |  |
| --- | --- | --- | --- |
|  | Name/Description model | Overt pronoun model | Null pronoun model |
| subject vs. nonsubject | \*\*\* | \*\* | \*\*\* |
| goal vs. source |  |  |  |
| different vs. same gender | (\*) | (\*) |  |
| subject x goal |  | \* |  |
| goal x gender |  |  |  |
| subject x gender |  |  |  |
| subject x goal x gender | \* |  |  |

\*\*\* = p < .001

\*\* = .001 < p < .01

\* = .01 < p < .05

(\*) = .05 < p < 0.1

Figure 4. Experiment 1: Average rate of using null and overt pronouns by subjecthood and thematic role.

### Pronoun users vs. No-pronoun users

Out of our 43 participants, 15 (35%) used only name/description and null pronouns, while 28 used all three forms. The pronoun-user group had a higher overall rate of using reduced forms (34% nulls; 17% overt pronouns) than the no-pronoun group (31% nulls). Table 3 lists the number of participants by country who fell in each group. There are several countries (e.g., Spain and Mexico) with individuals who fall into both groups, suggesting that this variation is not entirely dialect-driven, and may be based on individual differences in how people respond to this task. This is consistent with the finding that English speakers also exhibit variability in whether they produce pronouns or continue the sentence with an elliptical form (e.g., …*and threw the painting away;* Zerkle & Arnold, 2017a).

Table 3. Participants who used pronouns (or not) based on their location by country at time of participation.

|  |  |
| --- | --- |
| Pronoun users | No-pronoun users |
| Argentina | 3 | Argentina | 1 |
| Chile | 2 | Colombia | 1 |
| Costa Rica | 1 | Costa Rica | 1 |
| Dominican Republic | 1 | Ecuador | 1 |
| El Salvador | 1 | Guatemala | 1 |
| Guatemala | 1 | Mexico | 2 |
| Mexico | 6 | Puerto Rico | 2 |
| Panama | 1 | Spain | 6 |
| Peru | 3 |   |   |
| Puerto Rico | 2 |  |  |
| Spain | 2 |   |   |
| Venezuela | 5 |  |  |

We asked whether the thematic role effect occurred in both groups by analyzing each group separately. For these models, we excluded gender and its interactions, because these effects were either absent or marginal in the overall model. Results showed that the thematic role effect only emerged for the pronoun-user group. Their data looked very much like the data for the group overall, where they tended to use null pronouns more for subjects than nonsubjects, and pronouns more for subject/sources and nonsubject/goals. Models with only this subgroup critically showed the interaction between thematic role and subjecthood for pronoun production. This group also had an interaction between thematic role and subjecthood for null pronoun production, exhibiting the inverse effect: more nulls for subject goals than subject sources, and more nulls for nonsubject sources than subject goals.

On the other hand, the 15 no-pronoun subjects had a very strong subjecthood bias, where they produced null pronouns on 51% of the subject/goal trials; 52% of the subject/source trials; 8% of the nonsubject/goal trials, and 8% of the nonsubject/source trials. Only the main effect of subjecthood was significant in this model.

## Discussion

Experiment 1 revealed two main findings. First, grammatical function has a strong influence on reference form. As many models predict, speakers tend to avoid using explicit expressions (names/descriptions) when referring to a character last mentioned in grammatical subject position, and instead use reduced expressions (null or overt pronouns). For reference to nonsubjects we instead observed that overall names/descriptions were the most common choice. In addition, overt pronouns were relatively more likely for nonsubjects than subjects.

Our second finding was that thematic role modulated reference form choice as well. This was a much smaller effect, and only emerged for overt pronouns when used to refer to nonsubject characters. Stevenson et al. (1994) suggested that speakers focus their attention on the endpoint of transfer events, i.e. on the goal. This predicts that goals would be considered more salient/accessible than sources, and should be realized with reduced forms. In Spanish, there speakers can choose either overt pronouns (somewhat reduced) or null pronouns (even more reduced). Our results showed that speakers do avoid the most explicit form for goals in Spanish, supporting the contribution of thematic role information in reference choices. However, this also raises questions about why it does not have a broader effect.

However, in this experiment the effect was limited to only nonsubject references. This may be due to the fact that overt pronouns were rare in our sample, and were used mostly for nonsubject reference. On the other hand, it is notable that overt pronouns were not the dominant response for nonsubjects – instead, names were.

We did not observe more than a small, marginal effect of gender. In English, pronouns are more common when they are unambiguous, such as when the situation contains only one character that matches the pronoun’s gender (Arnold & Griffin, 2007; Rosa & Arnold, 2017). In Spanish, gender is only encoded on overt pronouns, which also mark gender. Given that overt pronouns are relatively uncommon in our task, we are not in a good position to detect any effect of gender that may contribute to reference selection in Spanish.

We also saw variation amongst individuals, where some people were more likely to use pronouns than others. This is consistent with findings from English versions of this task, where there is great individual variation in the rate of producing pronouns. Zerkle & Arnold (2017a) found that some participants treated the task as a simple “describe the picture” task, ignoring the fact that their response formed part of the story. Those participants used only names/descriptions, and also had a much lower use of connector words (e.g., “and”, or “so”), which signal discourse connectivity. Other participants treated the task as a story, and produced both pronouns and connector words. Likewise, our no-pronoun group in the Spanish experiment used far fewer reduced expressions overall than the pronoun group, suggesting that they may have engaged with the story aspect of our task less than the pronoun-user group.

Thematic roles clearly affected pronoun use for the pronoun-user group, but had no effect for the no-pronoun group. The most likely explanation for this pattern is that thematic roles most strongly affect the use of overt, not null pronouns. The no-pronoun group almost always used name/descriptions for nonsubjects, which is the condition where overt pronouns are most likely to occur for other participants. Without overt pronouns, we did not have the opportunity to observe a thematic role effect for these subjects. Alternatively, thematic roles may have a stronger effect in dialects where pronouns are more common, but this possibility is unmotivated and theoretically unlikely. Either way, these results suggest that thematic role affects overt pronouns more than other forms of reference in Spanish.

The most theoretically important finding in this experiment is that thematic roles can affect reference form in Spanish, albeit as a weak effect. This has implications for models of reference production both for Spanish particularly, and overall. Before we consider these implications, it is important to establish the robustness of this pattern of findings, particularly since the thematic role effect was small, and interacted with subjecthood. An additional question is whether thematic role effects can be attributed to production processes. We tested both of these questions in Experiment 2, where participants performed a spoken version of our task.

# Experiment 2

We extended our task to a spoken version for several reasons. First, this allows us to test whether thematic role effects are robust enough to emerge with a different population, and in a slightly different task. Second, experiments with live addressees permit stronger control over the experimental setting, in comparison with Amazon Mechanical Turk studies. Third, this allows us to ensure that participants are native speakers of Spanish, in that the experimenter could informally assess the fluency of participants in person.

By asking participants to respond verbally, we can also examine how performance across conditions relates to language production processing, specifically the time needed to plan a response. In our spoken task, we asked participants to read aloud the detective’s sentence, describing the first picture. They then used the second picture to guide their continuation of the story. This is similar to Rosa and Arnold’s (2017) Experiment 1, except that in their experiment, an experimenter spoke the context sentence aloud. As a test of planning time, Rosa and Arnold measured the latency to begin speaking as the time between the end of the context sentence and the participant’s response. They found that participants responded more quickly for trials in which they mentioned the goal, compared with the source (see also Zerkle & Arnold, 2015). Our goal here is to test whether the same effect occurs in Spanish. A secondary question is whether planning time correlates with reference form choice, although it does not for English (Rosa & Arnold, 2017; Zerkle & Arnold, 2017a, 2019; Zerkle & Arnold, 2015).

## Methods

### Participants

33 participants residing in the Chapel Hill and Durham areas participated in an in-person experiment. Participants were recruited through flyers, by email, and in person. 7 participants were excluded. 3 participants were excluded for technical and/or experimenter error: 1 received an incorrect version of the experiment, 1 participant's audio was not recorded, and 1 participant did not fit the eligibility criteria. 4 other participates were excluded on the basis of their responses: 1 participant used only names in their continuations, 1 participant referred to the target less than 50% of the time, 1 participant had more than 50% unusable responses, and 1 participant had difficulty reading while doing the task.

Experiments took place either in an on-campus research lab or in a study room at Durham County Library, main branch. Participants received $10 for completion of the hour-long experiment. Participants were individuals who grew up in a Spanish-speaking country, spoke Spanish as their first language, and came to the US at age 16 or later. Our participants were originally from 7 different countries/territories[[3]](#endnote-3). One participant grew up moving between Colombia and Argentina. Participants had been living in the United States anywhere from 0-21 years, and 19 individuals spoke at least one other language. All interactions with the participants, including participant recruitment, were conducted in Spanish to reduce English-language interference.

### Materials and Design

The introductory video was recorded by a native Spanish speaker who was from Mexico; this video informed participants of their role in the story. The same set of images was used as in Experiment 1 and participants were assigned to List A or List B on an alternating basis. As such, our independent variables were the same. Our stimuli were the same, except we replaced the names Sir Barnes and Lady Mannerly with the more generic terms El Duque (the Duke) and La Duquesa (The Duchess). As in Experiment 1, our dependent variable was the type of referential form produced to refer to the target character.

### Procedure

Participants sat in front of a computer and were presented with a video describing the characters, the story background and their role as a tabloid photographer. Before starting the experimental phase, participants were given a sample trial and listened to a sample response. They were also fitted with a microphone to record their audio. Participants then saw a preview of the images presented in the experimental phase of the experiment. Each pair of images was shown on the screen for 5 seconds. Participants were then shown the same set of images as in Experiment 1. The first image had a sentence written below it; participants read this sentence and then continued the story based on the second panel. Both sentences were recorded through the headset. Participants continued the experiment at their own pace, and pressed the space bar when they were ready to continue on to the next trial.

As in Experiment 1, we excluded responses that referred to the wrong referent, and responses with verbs like “gustar” (to like). We included data from all participants who had at least one name/description, and at least one null or overt pronoun. We also excluded participants if fewer than 50% of their responses were usable.

## Results

### Reference form choice

In Experiment 2, we modeled our data in terms of the same binary choices as in Experiment 1: 1) name or description vs. other; 2) overt pronoun vs. other; 3) null pronoun vs. other. Again, we used multilevel logistic regressions with SAS proc glimmix with centered predictors and maximal random effects. 481 items were included in the final analysis. All statistical results are provided in the Appendix, and the significance patterns are summarized in Table 4.

As illustrated in figure 5, participants tended to use null pronouns more for subjects than nonsubjects, and names more for nonsubjects than subjects. These patterns emerged as a main effect of subjecthood for both name and null pronoun dependent measures. In addition, speakers tended to use more overt pronouns and fewer names for goals than sources, as reflected by a main effect of thematic role for the name and overt pronoun models. There was a marginal tendency to use overt pronouns more in different-gender contexts.

Figure 5. Experiment 2, overall pattern of responses

Table 4. Experiment 2 results: Summary of effects from three models

|  |  |  |  |
| --- | --- | --- | --- |
|  | Name model | Overt pronoun model | Null pronoun model |
| subject vs. nonsubject | \*\*\* |  | \*\*\* |
| goal vs. source | \* | \*\* |  |
| different vs. same gender |  | (\*) |  |
| subject x goal |  |  |  |
| goal x gender |  |  |  |
| subject x gender |  |  |  |
| subject x goal x gender |  |  |  |

\*\*\* = p < .001

\*\* = .001 < p < .01

\* = .01 < p < .05

(\*) = .05 < p < 0.1

Figure 6. Experiment 2: Average rate of using null and overt pronouns by subjecthood and thematic role. Overt pronouns were more frequent for goals than sources (See Table 3).

### Pronoun users vs. No-pronoun users

Out of our 26 participants, 9 (35%) used only names/descriptions and null pronouns, while 17 used all three forms. Table 5 lists the number of participants by country of origin who fell in each group. As in Experiment 1, there were many countries with individuals in both groups. This overlap was more salient in Experiment 2 data, as demonstrated in Table 4.

Table 5. Participants who used pronouns (or not) based on their country of origin.

|  |  |
| --- | --- |
| Pronoun users | No-pronoun users |
| Colombia/Argentina | 1 | Colombia  | 1 |
| Colombia | 1 | Peru  | 1 |
| Guatemala | 1 | Guatemala | 1 |
| Mexico | 9 | Mexico  | 4 |
| Puerto Rico | 2 | Venezuela | 1 |
| Venezuela | 2 |  |  |

We tested each subgroup separately, examining only main effects (subject continuation, goal continuation, and different gender), since the interactions were not significant in the full models. Unlike in Experiment 1, we found that the goal bias emerged for both subgroups. For pronoun users, the thematic role effect only showed up for the use of pronouns, such that this group used significantly more pronouns for goals (37%) than sources (24%), and marginally fewer pronouns for subjects than nonsubjects. By contrast, their null pronoun and name/description production was only driven by a preference to use more null pronouns and fewer names for subjects vs. nonsubjects. For the no-pronoun users, we found both main effects of subjecthood and thematic role, such that they used null pronouns more for goals (40%) vs. sources (24%), and more for subjects (61%) than nonsubjects (3%). Just like in Experiment 1, the pronoun-users had a higher overall rate of using reduced forms (30% nulls; 31% overt pronouns) than the no-pronoun group (32% nulls).

### Latency

Our second question was whether reference form choice was related to the amount of time that participants spent preparing their response. To test this, we measured the time that elapsed between the moment when the participant finished speaking the first sentence and beginning of the first word in the second sentence. On 14 trials, the participant clicked ahead to the stimulus trial while still speaking; on these trials we used the end of the previous trial response as the start of the latency measurement. On 38 trials (out of 481), the participant produced connector words like “y” (*and),* and then paused before continuing the main response. In those cases, we included the pause in the total latency. We excluded latencies that were over 2.5 standard deviations above the mean (n = 15).

The first question we asked was whether the goal or source responses were easier to plan, as measured by latency. To assess this, we used log latency as our dependent measure, and modeled the same centered predictors as used for the reference form models[[4]](#endnote-4). In addition, we included the duration of the detective sentence as a predictor.[[5]](#endnote-5) Participants read the detective sentence, which means that its duration varied from trial to trial. The amount of time spent reading could potentially have been used to pre-plan the response, thus leading to shorter latencies. In addition, the time spent reading the detective sentence varied across conditions (goal / subject: 3875 ms; goal / object: 3228 ms; source/ subject: 3125 ms; source/object: 3468 ms).

As shown in Figure 7 (see also the Appendix, Table 12), the latency to respond was somewhat shorter for goal than source responses for the prepositional object continuations, but not for the subject continuations. This resulted in a marginal interaction between subject and goal continuation (β = 0.14 (.08), t = 1.9, p = .06); contrasts revealed a significant goal effect for nonsubjects (t = -2.29, p = .023) but not subjects (t = 0.02, p = .98). No other effects or interactions were significant.

Figure 7. Average latency to respond in the four critical conditions. This graph shows latency in ms for readability; models used log latency.

The second question we asked was whether latency predicted reference form, when added to the models. That is, are speakers more likely to produce null pronouns when sentence planning is easy? However, there was no such effect. We added latency to the reference form models, first eliminating any effects or interactions that were not significant or marginal in the original reference form model (but keeping detective latency as a control). As shown in the Appendix, tables 13-15, there were no effects of either latency or the duration of the detective sentence.

# Discussion

Experiment 2 replicated the two main findings of Experiment 1. First, null pronouns were produced more often for subjects than nonsubjects, and names/descriptions were preferred for nonsubjects over subjects. Remarkably, the rate of null pronoun use was highly consistent across our written and spoken experiments (Exp. 1: 56% for subjects; 8% for nonsubjects; Exp. 2: 61% for subjects; 2% for nonsubjects).

Second, and of critical importance, thematic roles modulated reference form production. Moreover, the effect of thematic roles was stronger in Experiment 2 than in Experiment 1. Like in Experiment 1, the strongest effect was on the use of overt pronouns, but it emerged as a main effect instead of an interaction. That is, overt pronouns were more common for goals than sources overall, for both subjects and nonsubjects. In addition, there was a preference to use names for sources more than for goals. Finally, even the no-pronoun-user group had an effect of thematic role, which emerged for their production of null pronouns. Thus, in this experiment the effect of thematic roles was not limited to just the production of overt pronouns.

Experiment 2, which was performed orally, also provided an opportunity to test whether thematic role effects are related to ease of planning, where latency is a metric of the time needed to plan a response. In an English version of this task, goal responses were faster to initiate than source responses (Rosa & Arnold, 2017). We found only partial support for this pattern in Spanish, in that the latency to respond was faster for goals than sources, but only in the nonsubject condition. In addition, just as for English, we found that latency was unrelated to the choice of reference form. Thus, the latency analysis argues against the hypothesis that production planning explains the thematic role effect.

# General Discussion

In two experiments, we examined whether reference production in Spanish can be explained purely as a function of the antecedent’s position (subject vs. nonsubject), or whether thematic roles also matter. We found that in both experiments, thematic role modulated reference form choices, in addition to the expected tendency to use null pronouns for subjects. In both experiments, speakers used more overt pronouns for goals than sources; in Experiment 1 this emerged only for nonsubjects, but in Experiment 2 it was a main effect.

Our findings are relevant to two theoretical questions. First, we asked whether reference production should be modelled on the basis of purely structural constraints. This idea has emerged in models of reference production in English (Fukumura & van Gompel, 2010; Kehler & Rohde, 2013; Rohde & Kehler, 2014). The importance of structural constraints might be inferred from evidence that they play a large role in the comprehension of null and overt pronouns in Romance Languages (Alonso-Ovalle et al., 2002; Carminati, 2002; de Carvalho Maia et al., 2017; Filiaci et al., 2013; Gelormini-Lezama et al., 2011, 2013; Gelormini-Lezama, 2018). Consistent with these claims, we also saw that subjecthood was a strong predictor of reference form.

Nevertheless, our findings also suggest that subjecthood alone does not provide a full explanation of reference production. Critically, we found that goal and source roles influence reference form in Spanish, as they do in English. This contrasts with claims that thematic role has no effect on pronoun production (Fukumura & van Gompel, 2010; Kehler et al., 2008; Kehler & Rohde, 2013; Rohde & Kehler, 2014). Notably, evidence against a role for thematic role mostly stems from work examining a different verb type, namely verbs with strong implicit causality biases, suggesting that there may be stronger thematic role effects with some verbs than others. On the other hand, studies with English consistently find that speakers use pronouns more for goals than sources (Arnold, 2001; Rosa & Arnold, 2017; Zerkle & Arnold (2018); Zerkle, Rosa, & Arnold, 2015), and even that implicit causality can also guide pronoun production in English (Weatherford & Arnold, 2021) and German (Bott & Solstad, under review). Thus, the current findings provide converging evidence for the idea that thematic roles can affect reference form.

Our work also extends research on reference comprehension in Spanish, which has also focused on the influence of syntactic role on comprehension biases (Alonso-Ovalle et al., 2002; Carminati, 2002; de Carvalho Maia et al., 2017; Gelormini-Lezama & Almor, 2011, 2014; Gelormini, 2018). Syntactic role is assumed to reflect the salience/accessibility of entities in the discourse context, where the more highly ranked subject is more accessible than lower-ranked objects or obliques (e.g., Brennan, Friedman, & Pollard, 1987). Thus, this body of work approach is consistent with both psycholinguistic evidence that pronouns are used more for subjects than other entities (e.g., Arnold, 1998, 2001; Kehler et al., 2008; Rohde & Kehler, 2014, Stevenson et al., 1994) and with theories that more broadly claim that reduced expressions are preferred for more salient or accessible referents (e.g., Ariel, 1990; Chafe, 1976; Gundel et al., 1993). Our work extends this comprehension work to production, and demonstrates that subjecthood is also a key factor in choosing referential expressions in Spanish. However, syntactic role is not the only relevant factor, and thematic role also affects the selection of overt pronouns.

The second question that we asked was whether the same factors guide reference production cross-linguistically. Comprehension research suggests that different constraints may apply to different forms, based on work in Finnish (Kaiser & Trueswell, 2008) and Spanish/Italian (Filiaci et al., 2013). Our work supports this idea. On the one hand, we found that the goal/source distinction affects reference form in both Spanish and English. But critically, the effect is different. In English, goal status increases the likelihood that the speaker will use a pronoun, on top of other constraints like subjecthood (Rosa & Arnold, 2017). This would be consistent with a salience model where both subjecthood and goal status increase the salience of the referent.

By contrast, in Spanish we found that the goal bias primarily affected overt pronoun production. This is surprising, because overt pronouns are otherwise used for the **less** salient option (i.e., nonsubjects). This pattern of results suggests that thematic role and grammatical function play different roles in Spanish. As for Finnish pronouns (Kaiser & Trueswell, 2008), the constraints for null and overt pronouns cannot be captured by a single factor like salience/accessibility (see also Filiaci et al., 2013; Fedele & Kaiser, 2015; Torregrossa, Bongart, & Tsimpli, 2015). Our data are mostly consistent with claims that subjecthood drives the use of null pronouns, which are primarily used for reference to the character last mentioned in subject position. On the other hand, the prominence of goals increases the use of overt pronouns, and decreases the use of names. In Experiment 2 this effect was consistent across all conditions, but in Experiment 1 it was only observed for nonsubject references.

However, our results also broadly suggest a high degree of similarity between reference production in Spanish and English. Speakers choose forms in both languages on the basis of grammatical function and thematic role. The difference between languages stems from the presence of two different reduced options in Spanish (overt and null pronouns), but just one in English (overt pronouns)[[6]](#endnote-6). Generally the Spanish null pronoun functions similarly to the English overt pronoun, and both are preferred for talking about subject referents. However, while the goal bias affects pronouns in English, it instead primarily affects overt pronouns in Spanish.

Our results are also broadly consistent with findings from Russian. Kravtchenko (2014) found that the likelihood of omitting the subject in Russian was correlated with the contextual probability of the referent, controlling for discourse factors like previous mention in subject position. Thematic role may contribute to contextual probability, in which case Russian provides an example of predictability affecting null reference, not just overt pronouns.

One might be concerned about the fact that our thematic role effect was not exactly the same across our two experiments. In Experiment 1, the difference between goal and source conditions was only significant for the nonsubject items, as indicated by contrasts in the model. In Experiment 2, the effect was more general. However, this difference should not be taken to indicate a difference in the underlying production mechanism. Reference production in experimental tasks exhibits extensive variation. The use of explicit vs. reduced forms is sensitive to many factors that are not explicitly manipulated, such as whether the participants think that it is their job to tell a story, as opposed to simply describe the second picture in isolation. Zerkle & Arnold (2017a) encountered extreme variation on this dimension. It was clear that some participants paid more attention to the context. Those “context-users” were also more likely to produce connector words, and their eye movements suggested that they used a different timecourse of planning the utterance than those participants who never used any pronouns. It is possible that the spoken task in Experiment 2 encouraged participants to use more reduced forms overall, or more referential form variation. In Experiment 1, the rate of overt pronoun production was quite low (6% for subject references, 16% for nonsubject references), which limits our ability to detect systematic variation, especially in the subject-reference condition. By contrast, overt pronoun use was higher in Experiment 2 (13% for subject references, 25% for nonsubject references), and we observed the thematic role effect in both conditions.

 Another potential concern about our findings is that our participants came from a variety of geographical areas, and may have spoken different dialects of Spanish. Some participants never used overt pronouns. This variation was not aligned with country of origin, but it could potentially reflect variation in dialect zones within each country. Indeed, the rate of overt pronouns is known to vary by dialect (Lipski, 2002; Otheguy et al., 2007; Pešková, 2013). However, it is not clear that dialect variation can explain individual variation in our study either. Participants who provided more form variation overall were those who used overt pronouns. In addition, the majority of participants in both experiments produced all three forms, suggesting extensive overlap in the referential system used for all our participants.

The next question is why thematic roles matter. Many scholars have pointed out that some thematic roles are more likely to be mentioned than others, and therefore those references are more predictable (e.g., Arnold, 2001; Guan & Arnold, 2021; Kehler & Rohde, 2013; Stevenson et al., 1994). This predictability consistently affects pronoun comprehension in both English (Kehler et al., 2008; Rohde & Kehler, 2014; Stevenson et al., 1994), and in null-pronoun languages like Italian (Kaiser & Fedele, 2015). But it is not clear why predictability should affect production. Speakers do not need to predict their own productions, because they are in control of them. Instead, speakers plan what they will say. This suggests that instead, the goal bias may be explained by something that is associated with predictability.

Here we asked whether the effect could be explained by planning effects. Perhaps goal continuations are easier to plan, and perhaps easier planning facilitates the selection of null and pronominal forms. To test this question, we examined latency. However, there was little evidence that goal continuations were easier to plan: responses were faster for goals than sources, but only for nonsubjects. More importantly, we found no evidence that latency predicted reference form (similar to results for English, Rosa & Arnold, 2017; Zerkle & Arnold 2019). In sum, planning does not offer a natural explanation for the thematic role effect.

Instead, we speculate that thematic roles contribute to the speaker’s representation of story events. When the goal character is continued, it provides a predictable continuation to the story. This may result in stronger connections between events in the speaker’s mental representation, and support the discourse constraints for using reduced forms (see Arnold & Nozari, 2017, for a similar effect). However, it is unclear why predictability has a less consistent effect for stories using implicit causality verbs, e.g. *The cook annoyed the butler because….*. One possibility is that these verbs require a temporal reversal to discuss the event’s causes, which may decrease representations of event connection. Further work is needed to understand why thematic roles modulate reference form.

The contribution of the current paper is to provide a systematic experimental test of the conditions under which Spanish speakers use null pronouns, overt pronouns, and names. While many studies only compare null vs. overt pronouns (e.g., Filiaci, et al., 2013), our study examines reference production as a contrast between three levels of referential explicitness. We demonstrate effects of both subjecthood and thematic role, in parallel with English. At the same time, we find that each effect applies to different reference forms in Spanish, suggesting that different constraints apply in each language. Most critically, our findings argue for a model of reference production that includes more than just subjecthood as an explanation for why people use pronouns.

# Author note

This research was supported by NSF grant 1348549 to J. Arnold.

**References**

Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L.F., & Clifton, C. Jr. (2002). Null and overt pronouns and the topic-focus articulation in Spanish. *Journal of Italian Linguistics*, *14*, 151-169.

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. Psychological Review, 106(4), 748 765. doi:10.1037/

0033-295X.106.4.748Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schliperoord & W. Spooren (Eds.), *Text representation* (pp. 29-87). Amsterdam/Philadelphia: John Benjamins.

Arnold, J. E. (1998). *Reference form and discourse patterns* (Unpublished doctoral dissertation). Stanford University, Stanford, CA.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation*. Discourse Processes, 31*, 137-162. doi:10.1207/S15326950DP3102\_02

Arnold, J.E. (2010). How speakers refer: the role of accessibility. *Language and Linguistic Compass, 4,* 187-203. <https://doi.org/10.1111/j.1749-818X.2010.00193.x>

Arnold, J. E. (2016). Explicit and emergent mechanisms of information status. Topics in Cognitive Science, *8*, 722-736. doi: 10.1111/tops.12220

Arnold, J.E., & Griffin, Z. (2007). The effect of additional characters on choice of referring expression: Everyone competes. *Journal of Memory and Language, 56*, 521-536. doi: [10.1016/j.jml.2006.09.007](https://dx.doi.org/10.1016/j.jml.2006.09.007)

Arnold, J.E., & Nozari, N. (2017). The effects of utterance timing and stimulation of left prefrontal cortex on the production of referential expressions. *Cognition, 160*, 127-144. <https://doi.org/10.1016/j.cognition.2016.12.008>

Bott, O. & Solstad, T. (under review). The Production of Referring Expressions is Influenced by the Likelihood of Next Mention. Ms., University of Bielefeld, Germany.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition, 14*, 237-273. [https://doi.org/10.1016/0010-0277(83)90006-9](https://doi.org/10.1016/0010-0277%2883%2990006-9)

Cacciari, C., Corradini, P., Padovani, R., & Carreiras, M. (2011). Pronoun resolution in Italian: The role of grammatical gender and context. *Journal of Cognitive Psychology,* *23*, 416-434, DOI: 10.1080/20445911.2011.526599

Carminati, M. N. (2002). *The processing of Italian subject pronouns* (Doctoral dissertation). Retrieved from Proquest (AAI3039345).

de Carvalho Maia, J., Vernice, M., Gelormini-Lezama, C., Cunha Lima, M. L., Almor, A. (2017). Co-referential processing of pronouns and repeated names in Italian. *Journal of Psycholinguistic Research*, *46*, 497–506.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–56). New York, NY: Academic Press.

Chafe, W. (2004). Discourse effects of polysynthesis. In C. L. Moder & A. Martinovic-Zic (Eds.), *Discourse Across Languages and Cultures* (pp. 37-52). Amsterdam/Philadelphia: John Benjamins.

de Carvalho Maia, J., Vernice, M., Gelormini-Lezama, C.,Cunha Lima, M. L., Almor, A. (2017). Co-referential processing of pronouns and repeated names in Italian. *Journal of Psycholinguistic Research*, 46:497–506. DOI 10.1007/s10936-016-9450-2

Fedele, E. & Kaiser, E. (2015, October). Resolving null and overt pronouns in Italian. In C. Brown, Q. Gu, C. Loos, J. Mielens, G. Neveu (Eds.), *Proceedings of the 15th Texas Linguistic Society*. Paper presented at the 15th Texas Linguistic Society Conference, University of Texas, Austin, TX (pp. 53-72).

 Filiaci, F. (2011). Anaphoric preferences of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. In C. Borgonovo, M. Español-Echevarría, and P. Prévost (Eds.), *Selected Proceedings of the 12th Hispanic Linguistics Symposium.* Paper presented at 12th Hispanic Linguistics Symposium, Laval University (pp. 171-182). Somerville, MA: Cascadilla Proceedings Project.

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, 29*, 825–843. doi: 10.1080/01690965.2013.801502

Fukumura, K., & van Gompel, R.P.G. (2010). Choosing anaphoric expression: Do people take into account likelihood of reference? *Journal of Memory and Language, 62*, 52-66. <https://doi.org/10.1016/j.jml.2009.09.001>

Garnham, A., Oakhill, J., Ehrlich, M., & Carreiras, M. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, *34*, 41-62. doi:10.1006/jmla.1995.1003

Gelormini-Lezama, C. (2018). Exploring the repeated name penalty and the overt pronoun penalty in Spanish. *Journal of Psycholinguistic Research*, *47*(2), 377–389. https://doi-org.libproxy.lib.unc.edu/10.1007/s10936-017-9529-4

Gelormini-Lezama, C., & Almor, A. (2011). Repeated names, overt pronouns, and null pronouns in Spanish. *Language and Cognitive Processes*, *26*, 437–454. <http://doi.org/10.1080/01690965.2010.495234>

Gelormini-Lezama, C. & Almor, A. (2014). Singular and Plural Pronominal Reference in Spanish. Journal of Psycholinguistic Research, 43(3) 299-313.

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), Topic continuity in discourse: A quantitative cross-language study (pp. 1–42). Amsterdam/Philadelphia: John Benjamins.

[Gundel, J. K.](http://psycnet.apa.org/index.cfm?fa=search.searchResults&latSearchType=a&term=Hartshorne,%20Joshua%20K.), Hedberg N., & Zacharski R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*, 274-307. doi: 10.2307/416535

Guan, S. & Arnold, J. E. (2021). The predictability of implicit causes: testing frequency and topicality explanations. *Discourse Processes*. DOI:
10.1080/0163853X.2021.1974690

Hartshorne, J. K., Sudo Y., & Uruwashi M. (2013). Are implicit causality pronoun resolution biases consistent across languages and cultures? *Experimental Psychology*, *60*, 179-196. doi: 10.1027/1618-3169/a000187

Kaiser, E., & Trueswell, J.C. (2008) Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes, 23*, 709-748. <https://doi.org/10.1080/01690960701771220>

Kehler, A., Kertz, L., Rohde, H., & Elman, J. (2008). Coherence and coreference revisted. *Journal of Semantics, 25*, 1-44. doi: [10.1093/jos/ffm018](https://dx.doi.org/10.1093/jos/ffm018)

Kehler, A., & Rohde, H. (2013). A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics, 39*, 1-37. <https://doi.org/10.1515/tl-2013-0001>

Kravtchenko, E. (2014). Predictability and syntactic production: Evidence from subject omission in Russian. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Chairs), *Program of the 36th Annual Meeting of the Cognitive Science Society.* Paper presented at the 36th Annual Meeting of the Cognitive Science Society, Quebec City, Canada.

Kravtchenko, E., Modi, A., Demberg, V., Titov, I., & Pinkal, M. (2017, March). *Does discourse predictability affect rate of pronominalization?* Poster presented at the 30th Annual CUNY Conference on Human Sentencing Processing, Massachusetts Institute of Technology, Cambridge, MA.

Lipski, J. (2002). *Subject pronoun usage among Spanish dialects*. Unpublished manuscript, Department of Linguistics, Pennsylvania State University, State College, PA.

Navarrete, E., & Costa A. (2009).  [The naming of gender-marked pronouns supports interactivity in models of lexical access](http://www.uv.es/psicologica/articulos2.09/9NAVARRETE.pdf). *Psicológica, 30,* 301–321.

Otheguy, R., Zentella, A.C., & Livert, D. (2007). Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language,* *83*, 770-802. doi: [10.1353/lan.2008.0019](https://doi.org/10.1353/lan.2008.0019)

Pešková, A. (2013). Experimenting with pro-drop in Spanish. *SKY Journal of Linguistics*, *26*, 117-149.

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production*. Language, Cognition, and Neuroscience*, *29*, 1-16. <http://dx.doi.org/10.1080/01690965.2013.854918>

Rosa, E. C., & Arnold, J. E. (2017). Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language*, *94,* 43-60. doi: 10.1016/j.jml.2016.07.007

Sainzmaza-Lecanda, L., & Schwenter, S.A. (2017). Null objects with and without bilingualism in the Portuguese-and Spanish-speaking world. In K. Bellami, M. Child, P. González, A. Mutendam, & M. C. Parafita Couto (Eds.), *Issues in Hispanic and lusophone linguistics* (Vol. 13, pp. 95-113). Amsterdam/Philadelphia: John Benjamins.

Serratrice, L. (2008). Null and overt subjects at the syntax-discourse interface: Evidence from monolingual and bilingual language acquisition. In van Kampen, J., Bauw, S. & Pinto, M. (Eds.), *LOT Occasional Series* (Vol. 8, pp. 181-200) Utrecht, the Netherlands: LOT.

Stevenson, R.J., Crawley R., & Kleinman D. (1994). Thematic roles, focusing and the representation of events. *Language and Cognitive Processes, 9*, 519–548. <https://doi.org/10.1080/01690969408402130>

Torregrossa, J., Bongartz, C., Tsimpli, I. (2015). Testing accessibility: A cross-linguistic comparison of the syntax of referring expressions, extended abstract for the *Proceedings of the 89th Annual Meeting of the Linguistic Society of America.*

van Rij. J., van Rijn, H., & Hendriks, P. (2012, April). How WM load influences pronoun interpretation. In N. Rußwinkel, U. Drewitz & H. van Rijn (Eds.) *Proceedings of the 11th International Conference on Cognitive Modeling*. Paper presented at the 11th International Conference on Cognitive Modeling, Universitaetsverlag der TU Berlin, Berlin, Germany.

Weatherford, K., & Arnold, J. E. (2021). Semantic predictability of implicit causality can affect referential form choice. *Cognition*, 214, 104759. DOI: https://doi.org/10.1016/j.cognition.2021.104759

Zerkle, S., Rosa, E. C. & Arnold, J. E. (2015, March). Do addressee gestures influence predictability of spoken reference form?Poster presented at the 28th Annual CUNY conference on Human Sentence Processing, Los Angeles, CA.

Zerkle, S., & Arnold, J. E. (2017a). Discourse attention during utterance planning affects referential form choice. *Linguistics Vanguard, 2*(S1). <https://doi.org/10.1515/lingvan-2016-0067>

Zerkle, S., Rosa, E. C., & Arnold, J. E. (2017b). Thematic role predictability and planning affect word duration. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8*, 1-28. doi: <http://doi.org/10.5334/labphon.98>

Zerkle, S. & Arnold, J. E. (2019). Does planning explain why predictability affects reference production? *Dialogue and Discourse, 10(2),* 34-55*.*

# Appendix

## Experiment 1 Models

*Table 6*. Experiment 1: Name/description rate model

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate (Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | -2.30 (0.28) | -8.08 | <.0001 |
| **goal continuation** | -0.25 (0.31) | -0.79 | 0.44 |
| **different gender (vs. same gender)** | -0.56 (0.31) | -1.81 | 0.09 |
| **subject\*goal**  | 0.51 (0.44) | 1.16 | 0.26 |
| **goal\*same gender** | -0.85 (0.62) | -1.37 | 0.19 |
| **subject\*same gender** | -0.11 (0.43) | -0.25 | 0.80 |
| **subject\*goal\*same gender**  |  -1.94 (0.87) | -2.24 | 0.04 |

\* The random slope for different gender by participant was estimated to be zero by the model and therefore excluded.

*Table 7.* Experiment 1: Overt pronoun rate model

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | -1.12 (0.30) | -3.79 | 0.0014 |
| **goal continuation** | 0.06 (0.37) | 0.17 | 0.86 |
| **different gender (vs. same gender)** | 0.70 (0.35) | 2.0 | 0.06 |
| **subject\*goal**  | -1.85 (0.58)  | -3.2 | 0.0038 |
| **goal\*same gender**  | 0.22 (0.70) | 0.32 | 0.75 |
| **subject\*same gender**  | 0.07 (0.57) | 0.13 | 0.90 |
| **subject\*goal\*same gender**  | 0.63 (1.15) | 0.55 | 0.59 |

\* The random slope for different gender by participant was estimated to be zero by the model and therefore excluded.

*Table 8.* Experiment 1: Null pronoun rate model

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | 3.37 (0.39) | 8.66 | <.0001 |
| **goal continuation** | -0.21 (0.31) | -0.67 | 0.51 |
| **different gender (vs. same gender)** | 0.20 (0.31) | 0.65 | 0.53 |
| **subject\*goal**  | 1.06 (0.64) | 1.64 | 0.12 |
| **goal\*same gender**  | 0.79 (0.63) | 1.26 | 0.22 |
| **subject\*same gender**  | -0.07 (0.64) | -0.12 | 0.91 |
| **subject\*goal\*same gender**  | 1.24 (1.28) | 0.97 | 0.34 |

\* The random slopes for different gender and goal continuation by participant were estimated to be zero by the model and therefore excluded.

**Experiment 2 Models**

*Table 9*. Experiment 2: Name/description rate model

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | -3.11 (0.4) | -7.82 | <.0001 |
| **goal continuation** | -1.18 (0.43) | -2.78 |  0.01 |
| **different gender (vs. same gender)** | -0.67 (0.42) | -1.61 | 0.13 |
| **subject\*goal**  | -0.4 (0.64) | -0.62 | 0.54 |
| **goal\*different gender** | 1.44 (0.84) | 1.72 | 0.10 |
| **subject\*different gender**  | 0.25 (0.63) | 0.4 | 0.69 |
| **subject\*goal\*different gender**  | 0.06 (1.25) | 0.05 | 0.96 |

\* subject continuation was a random slope by both participants and items; all other random slopes estimated to be zero by the model

*Table 10.* Experiment 2: Overt pronoun rate model

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | -0.75 (0.48) | -1.57 | 0.133 |
| **goal continuation** | 1.01 (0.35) | 2.9 | 0.009 |
| **different gender**  | 0.66 (0.34) | 1.93 | 0.07 |
| **subject\*goal**  | -0.41 (0.83) | -0.49 | 0.63 |
| **goal\*different gender** | -1.17 (0.69) | -1.7 | 0.107 |
| **subject\*different gender**  | -0.54 (0.82) | -0.66 | 0.519 |
| **subject\*goal\*different gender**  | 1.11 (1.65) | 0.67 | 0.508 |

\* subject continuation was a random slope by both participants and items; all other random slopes estimated to be zero by the model

*Table 11.* Experiment 2: Null pronoun rate model

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | 4.58 (0.63) | 7.25 | <.0001 |
| **goal continuation** | 0.03 (0.6) | 0.05 | 0.96 |
| **different gender (vs. same gender)** | 0 (0.59) | -0.01 | 0.99 |
| **subject\*goal**  | 1.03 (1.08) | 0.96 | 0.34 |
| **goal\*different gender** | 0.14 (1.16) | 0.12 | 0.90 |
| **subject\*different gender**  | 0.51 (1.08) | 0.47 | 0.64 |
| **subject\*goal\*different gender**  | -1.69 (2.15) | -0.79 | 0.44 |

\* maximal random slopes included

Table 12. Experiment 2 Latency model: how do experimental conditions affect latency to respond?

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | -0.036 (0.04) | -0.97 | 0.35 |
| **goal continuation** | -0.071 (0.05) | -1.42 | 0.17 |
| **different gender continuation** | 0.024 (0.05) | 0.46 | 0.65 |
| **subject\*goal continuation** | 0.14 (0.07) | 1.9 | 0.06 |
| **subject\* different gender continuation** | 0.017 (0.07) | 0.23 | 0.82 |
| **subject\*goal\*different gender continuation** | -0.08 (0.14) | -0.54 | 0.59 |
| **detective latency** | 0.000012(0.000011) | 1.12 | 0.27 |

\* maximal random effects included

*Table 13.* Experiment 2 Model for name/description responses, including latency as a predictor

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | -3.18 (0.38) | -8.45 | <.0001 |
| **goal continuation** |  -1.18 (0.47)  | -2.51 | 0.02 |
| **Latency** | 0.0001 (0.0001) | 0.77 | 0.44 |
| **detective sentence duration** | 0.0001 (0.0001) | 1.18 | 0.24 |

\* Subject continuation was a random slope by both participants and items; all other slopes estimated to be zero by the model

*Table 14.* Experiment 2 Model for null pronoun responses, including latency as a predictor

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | 4.74 (0.63) | 7.47 | <.0001 |
| **goal continuation** | 0.48 (0.44) | 1.11 | 0.2833 |
| **latency** | 0.00003 (0.0001) | 0.23 | 0.8181 |
| **detective sentence duration** | -0.00003 (0.00009) | -0.33 | 0.743 |

\*Subject continuation was a random slope by both participants and items; goal continuation was a random slope by participants; all other slopes estimated to be zero by the model

*Table 15.* Experiment 2 Model for overt pronoun responses, including latency as a predictor

|  |  |  |  |
| --- | --- | --- | --- |
| **Effect** | **Estimate** **(Standard Error)** | **t Value** | **Pr > |t|** |
| **subject continuation** | -0.82 (0.44) | -1.86 | 0.08 |
| **goal continuation** | 0.81 (0.37) | 2.21 | 0.04 |
| **latency** | -0.00015 (0.000157) | -0.93 | 0.35 |
| **detective sentence duration** | -0.00009 (0.000162) | -0.58 | 0.56 |

\* Subject continuation was a random slope by both participants and items; all other slopes estimated to be zero by the model

Notes

1. Argentina (4), Chile (2), Colombia (1), Costa Rica (2), Dominican Republic (1), Ecuador (1), El Salvador (1), Guatemala (2), Mexico (7), Panama (1), Peru (2), Puerto Rico (4), Spain (8), Uruguay (1), and Venezuela (6). [↑](#endnote-ref-1)
2. If the speaker wishes to say “He likes it”, in Spanish the most natural translation is “Le gusta”, which takes a form more similar to “it pleases him”. [↑](#endnote-ref-2)
3. Colombia/Argentina (1), Colombia (2), Guatemala (1), Mexico (13), Peru (1), Puerto Rico (5), and Venezuela (3) [↑](#endnote-ref-3)
4. subject continuation; goal continuation, different gender, subject x goal; subject x different gender; [↑](#endnote-ref-4)
5. Detective latency was included as a random slope only by participants. [↑](#endnote-ref-5)
6. English speakers can also indicate reference by using an elliptical, or VP-coordination structure (*Sir Barnes received a painting from Lady Mannerly…. And chucked it in the closet)*. This is not the same as null pronoun use, but because it is achieved by syntactic coordination, it also is most heavily influenced by grammatical function (Zerkle & Arnold, 2018). [↑](#endnote-ref-6)