Asymmetries between interpretation and production in Catalan pronouns

Laia Mayol (Universitat Pompeu Fabra)

Introduction. The literature on Romance null-subject languages has often postulated a division of labor between Null and Overt pronouns: Nulls prefer to retrieve an antecedent in subject position, whereas Overts prefer an antecedent in a lower syntactic position (Carminati, 2002). However, recent research on English pronouns (Rohde and Kehler, 2014) has shown that pronoun interpretation and production are sensitive to different set of factors and, instead of being mirror images of each other, are related probabilistically in a Bayesian fashion. It is an open question whether null-subject language pronouns can be accounted in the same way. We show that both Null and Overts exhibit the same asymmetry in Catalan, by means of two discourse completion studies: production is only sensitive to grammatical factors, while interpretation is also sensitive to pragmatic factors (i.e. context type and rhetorical relation).

Experiment 1. A discourse-completion study was carried out in which the context sentence used a transfer of possession verb (TPV), see (1). TPV contexts are interesting because a pronoun following it may not have the usual subject preference (Stevenson et al., 1994): i.e. a very natural follow-up of (1) would be to explain what Pere did with the book. Experiment 1 had 3 conditions: Free (participants could continue as they wished), Overt and Null (participants had to continue with an Overt and a Null, respectively). Three lists were generated with 18 items and 18 fillers. 90 Catalan native speakers participated in the experiment, yielding 1620 completions (see (2) for an example with a Null). We will only discuss the 1152 completions in which the subject unambiguously referred either to the subject or object of the context sentence.

(1)   El Robert li va passar un llibre al Pere.   (2)   Va donar -li   les gràcies.
       'Robert passed a book to Peter.'   Gave   him the thanks.

The data in the Null and Overt conditions, in Table 1, are interpretation probabilities: i.e. of how participants understood the pronouns presented to them. In the Null condition, the pronoun displayed a mild object preference, and in the Overt condition, the pronoun exhibited a very strong object preference. In the Free condition, we coded what type of expression was chosen to refer to either the subject or the object. These data, in Table 2, are, thus, production probabilities.

Table 1 (*)[1]	Null	Overt	Free
Subject	38	10	24
Object	62	90	76

Table 2 (*)	Subject	Object
Null	72	29
Overt	6	15
Proper Name	21	56

The results have uncovered an asymmetry between interpretation and production. Although Nulls displayed a mild interpretation bias towards the object, they have strong production bias towards the subject. When participants could choose a form they overwhelmingly chose a Null to refer to the subject. However, when they had to interpret a pronoun, the pragmatic biases of VTPs came into play and the Null pronoun lost its subject bias. The data has also uncovered another asymmetry concerning Overts. Although interpreters have a very strong bias to interpret an Overt as referring to the object, this is not how they use it when they can choose which form to use. The results also show that, while Nulls are more subject biased than Overts (as expected), it is possible for Nulls to have an overall Object preference, depending on pragmatic factors.

We further coded the rhetorical relation established between the context and the completion sentence to examine the role of pragmatic factors. Here we discuss the results of two rhetorical relations which have been shown to display opposed biases: Occasion and Elaboration (Kehler et. al. 2008). In an Occasion, the events described by the two sentences are temporally ordered, while in an Elaboration, the two sentences provide descriptions of the same eventuality. In the

---

[1]* indicates statistically significant differences at the .05 level in a mixed-effect logistic regression with item and participant as random effects.

former the end state for the first eventuality is important for the coherence relation, while in the latter it is not. This affects pronoun interpretation: pronouns in an Occasion , (3-a), are biased towards the object, since it is the most salient referent at the end state of the VTP sentence. In contrast, pronouns in an Elaboration, (3-b), are biased towards the subject, since Elaborations do not focus on the end state of the first event.

(3) a. John handed a book to Bob. He began reading it.
      b. John handed a book to Bob. He did so slowly and carefully.

Table 3 shows the subject bias of Nulls and Overts in the two rhetorical relations. As before, Nulls are more subject biased than Overts, but their biases are greatly affected by the rhetorical relation: Elaboration is much more subject-biased than Occasion. Table 4 shows the production data: the subject bias of Null pronouns by rhetorical relation. The difference is not statistically significant showing again that production, unlike interpretation, is insensitive to pragmatic factors.

Table 3 (*)	Null	Overt
Occasion	11	0
Elaboration	75	41

Table 4	%
Occasion	80
Elaboration	71

Experiment 2 had the same design as the previous experiments, but included implicit causality verbs, which impute the cause of the event they denote either to the subject (ICV1; *surprise*) or the the object (ICV2; *congratulate*). The experiment contained 30 critical items (15 ICV1s and 15 ICV2s) and 20 fillers. 78 native speakers of Catalan participated yielding 2340 completions, of which we analyze the 1963 that unambiguously referred to the previous subject or object.

Table 5 shows the percentage of subject reference of both pronouns depending on verb type: Nulls are more subject-biased than Overts, and ICV1s trigger more subject references than ICV2s. Thus, pronoun interpretation is sensitive to both grammatical and pragmatic factors. Table 6 shows the choice of referring expression in the Free condition for subject reference: Nulls are the favorite form regardless of whether the context is ICV1 or ICV2. Production is not affected by the pragmatic factors (i.e. verb type) that did affect interpretation.

Table 5 (*)	Null	Overt
All	65	20
VIC1	76	51
VIC2	52	15

Table 6	ICV1	ICV2
Null	87	93
Overt	4	2
Proper name	9	5

A Bayesian account. The data supports a model in which interpretation and production are related in a Bayesian fashion, as in (4). $P(subj \mid pronoun)$ is the interpretation probability: the probability that a pronoun refers to the subject. $P(pronoun \mid subj)$ is the production probability: the probability that the speaker uses a pronoun given that she wants to refer to the subject. These two probabilities are not mirror images of each other; a third probability plays a role, $P(subj)$, which is the probability that the subject will be mentioned regardless of the form (Null, Overt or Proper name).[2]

(4)

$$P(subj \mid pronoun) = \frac{P(pronoun \mid subj)\,P(subj)}{P(pronoun)}$$

In experiment 1, the observed probability $P(subject \mid null)$ was 38% (see table 1) and the expected probability applying the formula in (4) is 44%. In experiment 2, the observed probability $P(subject \mid null)$ was 65% (see Table 5) and the expected probability is 64%. For both experiments a linear regression test over item means was conducted, which yielded statistically significant correlations between observed and predicted means.

References. Carminati, M. N. (2002). *The processing of Italian subject pronouns*. PhD thesis, UMasss. ● Rohde, H. and Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, 29(8):912–927. ● Stevenson, R. J., Crawley, R. A., and Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, 9(4):519–548.

---

[2]$P(pronoun)$ is the probability that a pronoun is used and is computed summing the denominator over all possible referents. It contributes a constant factor and normalizes the probabilities over all possible referents to 1.