Romanian subject alternation: the position of antecedent hypothesis

Fabian Istrate, Anne Abeillé, Barbara Hemforth
istratemariusfabian@gmail.com
Paris Graduate School of Linguistics,
Laboratoire de Linguistique Formelle, Université Paris Cité-CNRS

The *Position of Antecedent Hypothesis* (Carminati 2002, 2005) is one of the main factors accounting for preferences in languages with null / pronominal subject alternation. According to this hypothesis, null subjects favor a subject antecedent and pronominal subjects, a non-subject antecedent. We ran a first experiment in which we found that null subjects are prefered over pronominal ones in Romanian. Additionally, negative polarity slightly increased the number of pronominal subjects. Our second experiment revealed that preferences for a subject or an object antecedent for null vs pronominal subject are less categorical than in Italian, and this tendency might be linked to a less widespread use of null subjects in Romanian, compared to Italian.

1. Introduction

Like other Romance languages (except French), Romanian is a pro-drop language, allowing pronominal subject omission with any verbal form. The choice between null and pronominal subjects has been shown to be sensitive to several factors in Romance languages: word order (Mayol 2010, Catalan), person and animacy (Correa Soares et al., 2019, 2020; Fernandes et al 2018, Brazilian Portuguese), discourse status (Runner & Ibarra 2016, Spanish), function of the antecedent (Carminati 2005, Italian), etc. According to the *Position of Antecedent Hypothesis* (Carminati 2002, 2005), subject preferences come in two parts: (i) null subjects favor a subject antecedent, (ii) overt (pronominal) subjects favor a non-subject antecedent. These tendencies have been shown to play a role both in production and comprehension (anaphora resolution). Experiments carried out on other Romance languages (Spanish, Catalan, Portuguese) have shown that, whereas the first part of the hypothesis seems to be rather robust cross-linguistically (e.g. Filiaci et al. 2013, de la Fuente et al. 2016), the second part is subject to variation: e.g. in Spanish, the object preference associated with pronominal subjects is often weaker than in Italian (Sorace & Filiaci 2006, Sorace et al. 2009, Filiaci et al. 2014, Chamorro 2018, Contemori et al. 2019). In our research project, we want to find out which factors favor the use of null vs pronominal subjects in Romanian, starting with the role of contrast (realized as polarity in Experiment 1) and in how far the PAH applies to Romanian (Experiment 2).

2. Empirical evidence

2.1 Null subjects in Romanian

Following earlier work on the role of discourse status (Correa Soares et al. 2020), we looked at the role of polarity in this experiment. We assume a contrastive role of negation as for example in Farkas (2010). Thus, in an example like (1), with a negative answer, the pronominal subject may be interpreted as a contrastive topic (Krifka 2006). We consider that there may be a presupposition that the person already mentioned (Maria) represents one of several possible alternatives in the communication situation.

We ran a first binary forced-choice experiment on IbexFarm (20 experimental items, 20 distractors and 10 comprehension questions) with one experimental factor Polarity, with two values (*affirmative, negative*) cf. (1). 49 Romanian native speakers (6 male, 43 female, mean age=21), students from the University of Bucharest, read 20 items consisting of a polar question and a choice of two answers with either a null or pronominal subject. As prescribed by the Romanian norm, participants showed a clear preference for null subjects, across

Polarity conditions (Figure 1, Table 1). However, we also found a main effect of polarity: for a negative answer participants chose more frequently the pronominal subject than for a positive answer.

(1) Question: *A   ascultat Maria sfaturile   părinților   săi?*
              AUX.3SG listen.PST Maria advices.DEF   parents.GEN.DEF POSS.3SG
              'Did Mary listen to her parents' advice?'
     Answer 1: *Da, (ea) ascultă mereu sfaturile părinților săi.*
              'Yes, (she) always listens to her parents' advice.'
     Answer 2: *Nu, (ea) a luat decizia pe cont propriu.*
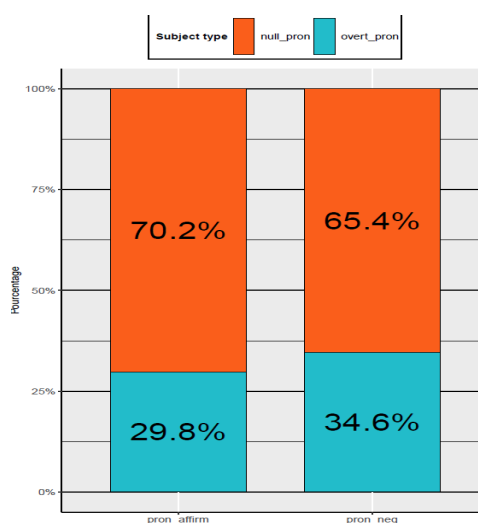   'No, (she) made the decision on her own.'


```
Fixed effects:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)        1.5240     0.4414   3.452 0.000555 ***
conditionpron_neg -0.4327     0.1902  -2.276 0.022860 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr)
cndtnprn_ng -0.234
```

Figure 1. Results of Experiment 1   *Table 1. Statistical analysis for Exp. 1*

2.2   Experiment 2: The PAH in Romanian

We ran a second binary forced-choice task on IbexFarm (20 experimental items, 30 distractors, comprehension questions after 20% of the trials), by comparing antecedent choices in two conditions: *null, overt*, cf. (2,3). Each experimental item contained a transitive verb with two human arguments (subject and object), followed by a temporal embedded clause containing a null or pronominal subject. The task of the participants was to determine the antecedent of the embedded subject (either matrix subject or object). Half of our items contained female arguments, cf. (2), and another half male arguments, cf. (3) (no effect of gender was found in this experiment). We tested 64 Romanian native speakers (25 male, 39 female, mean age=24), all students from the University of Bucharest. As shown in Figure 2, participants mostly prefer to associate the embedded null subject with the matrix subject antecedent (64.2%), and the embedded pronominal subject with the matrix object antecedent (58.2%), confirming Carminati's hypothesis. For the statistical analysis, we used a general linear mixed model as shown in Table 2. The factor type of subject (null vs. overt) is statistically significant (p < .001).

(3) *Ioana a   văzut-o   pe Alexandra când (ea) s-a*
     Ioana AUX.3SG see.PST-CL.3SG.F.ACC   DOM Alexandra when (she) REFL-AUX.3SG
     urcat   în   autobuz.
     get.PST in   bus
     'Ioana saw Alexandra when she got on the bus.'
     Question: *Cine s-a   urcat   în   autobuz?*
              who REFL.AUX.3SG get.PST on bus
              'Who get on the bus?'
     Answer 1: *Ioana.*

Answer 2: *Alexandra*.

(3) *Matei l-a*   *întâlnit*   *pe*   *Cătălin*   *când*   *(el) a*   *ajuns*
Matei CL.3SG.M.ACC-AUX.3SG meet.PST DOM Cătălin when he AUX.3SG arrive.PST
la Bușteni.
at Bușteni
'Matei met Cătălin when he arrived in Bușteni.'
Question: *Cine a ajuns în Bușteni?*
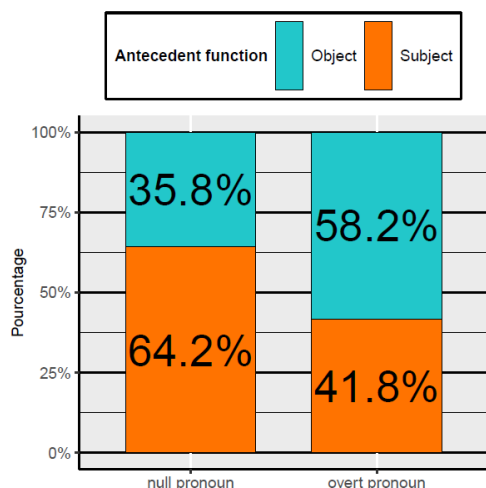   'Who arrived in Bușteni?'
Answer 1: *Matei.*
Answer 2: *Cătălin.*

Figure 2. Results of Experiment 2	*Table 2. Statistical analysis for Exp. 2*

3. The Position of Antecedent Hypothesis in Romance languages

In a Romance perspective (cf. Table 3), our experiments show that (i) Romanian has a preference for null subjects (70%), (ii) a null subject favors a subject antecedent (64%), as in Spanish and Catalan (though not categorically as in Italian or European Portuguese), and (iii) a pronominal subject favors an object antecedent (58%), as in Spanish (Filiaci et al. 2013, de la Fuente et al. 2013) and Catalan (Mayol 2010). In contrast to this group of languages (Romanian, Spanish, Catalan), Italian and European Portuguese are characterized by a stronger tendency to interpret null subjects as referring to a subject antecedent and a stronger tendency to link the pronominal subjects to an object antecedent (about 80%). We thus observe some variation across Romance languages in the preference for subject/object, since Italian is characterized by a stronger tendency to link the pronominal subject to an object antecedent (about 80%), while in Romanian and Catalan, for instance, preferences seem to be less categorical.

Language	Subject/ Object + null subject	Subject/ Object + overt subject
Italian	76.8 % / 23.1%	17% / 82.9%
European Portuguese	79 % / 21%	24 % / 76%
Spanish	66.3 % / 36.8%	33.6 % / 63.1%
Catalan	59.1 % / 40.9%	35.2 % / 64.8%
Romanian	**64.2** % / 35.8%	**41.8** % / 58.2 %

Table 3. Romanian & other Romance languages

4. Conclusion

We conclude that subject alternation in Romanian is sensitive to the antecedent function as in other Romance languages. As in Spanish and Catalan, the first part of Carminati's hypothesis is more robust than the second part. Although our results will have to be tested in a larger scale fully parallel crosslinguistic study, we assume that the less categorical choice between an object or a subject antecedent for a pronominal subject might be related to the generally weaker preference for null subjects in Romanian compared to Italian or European Portuguese (Fernandes et al. 2018). As suggested by the first experiment, the choice of a null or pronominal subject is also sensitive to semantic features like polarity, an issue which has to be further investigated. The baseline of null subjects in Romanian (70%), compared to Italian (77%, see Torregrossa et al. 2020) and Catalan (62%, see Casanova 1999) is likely to trigger gradual preferences when associating the type of subject with its antecedent (subject or object): the higher the percentage of the null subject baseline, the stronger we expect the preferences to link a null subject to subject antecedent and the pronominal subject to a non-subject antecedent to be. Ongoing research (corpus analyses and experiments) extends our work on the constraints for the usage of null vs pronominal subjects and their role in pronoun resolution.

Selected references

Carminati, M. N., 2005, Processing reflexes of the Feature Hierarchy (Person> Number> Gender) and implications for linguistic theory, *Lingua*, 115, 259-285.

Correa Soares, E., Miller,P., & Hemforth, B.. (2019). The effect of verbal agreement marking on the use of null and overt subjects: a quantitative study of first person singular in Brazilian Portuguese. *Fórum Linguístico* 16, no. 1 (2019): 3579-3600. https://periodicos.ufsc.br/index.php/forum/article/view/1984-8412.2019v16n1p3579.

Correa Soares, E., Miller, P., & Hemforth, B. (2020). The Effect of Semantic and Discourse Features on the Use of Null and Overt Subjects - A Quantitative Study of Third Person Subjects in Brazilian Portuguese. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, *36*(1), 2020360107. Epub July 06, 2020.https://dx.doi.org/10.1590/1678-460x2020360107

de la Fuente, I. & Hemforth, B., 2013, *Effects of Clefting and Left-Dislocation on Subject and Object Pronoun Resolution in Spanish*, in *Proceedings of the 16th Hispanic Linguistics Symposium*, 27-45. Cascadilla Proceedings Project.

de la Fuente, I., Hemforth, B., Colonna,S., & Schimke, S. (2016). The role of syntax, semantics, and pragmatics in pronoun resolution: A cross-linguistic overview. In *Empirical perspectives on anaphor resolution*, 11-32. Anke Holler, Katja Suckow (eds). Berlin/Boston: De Gruyter.

Farkas, D., 2010, The grammar of polarity particles in Romanian, in A. M. Di Sciullo & V. Hill (eds.), *Edges, Heads, and Projections: Interface properties*, John Benjamins, 87-124.

Fernandes, E. G., Luegi, P., Correa Soares, E., de la Fuente, I. & Hemforth, B., 2018, *Adaptation in pronoun resolution: Evidence from Brazilian and European Portuguese*. *Learning, memory, and cognition*, 44(12), 1986–2008.

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.

Mayol, L., 2010, *Refining Salience and the Position of Antecedent Hypothesis: a Study of Catalan Pronouns*, Proceedings of the 33rd Annual Penn Linguistics Colloquium.