
Refining Salience and the Position of Antecedent Hypothesis: a Study of Catalan Pronouns

Abstract

This paper investigates the referring preferences of null subject pronouns and overt subject pronouns in Catalan. In particular, it contributes the debate about whether the referring preferences of pronouns in null-subject languages are governed by syntactic factors or by information structure. A questionnaire experiment was performed, in which the word order of the experimental items (SVO vs OVS) was manipulated in order to distinguish syntactic and information structure factors. The results support a multiple-factor model of salience, in which grammatical as well as pragmatic function play a role and in which different forms are sensitive to different factors. Null pronouns refer to the most salient antecedent, which is always the subject, even if it is not the link of the sentence. In contrast, overt pronouns refer to non-salient antecedents (non-subject, non-link antecedents), if there such an antecedent, and fail to show a clear preference otherwise.

Refining Salience and the Position of Antecedent Hypothesis: a Study of Catalan Pronouns

Laia Mayol*

1 Introduction

The goal of this paper is to investigate the referring preferences of null-subject pronouns (NSPs) and overt subject pronouns (OSPs) in Catalan. Catalan, on par with other Romance languages, is a null-subject language in which both null and overt pronouns can generally occur in subject position, as shown in 1.¹

- (1) a. \emptyset Vivia als Estats Units.
“ \emptyset Lived in the United States.”
- b. Ella vivia als Estats Units.
“She lived in the United States.”

This work contributes to the debate about whether the referring preferences of pronouns in null-subject languages are governed by syntactic factors (see Carminati, 2002) or by information structure factors (see Samek-Lodovici, 1996; Frana, 2007). Also, since there is a correlation between reduced anaphoric forms and salient antecedents (Ariel, 1990), we can gain insight into which factors compose salience.

The main claims of this paper are the following: (i) salience is affected by both syntactic and information structure factors, and (ii) different anaphoric forms are sensitive to different factors: NSPs have a preference for subjects and OSPs have a preference for low-salience antecedents.

This paper is organized as follows. Section 2 presents some background on this topic. Section 3 presents an experiment for Catalan designed to tell apart syntactic and information structure factors. Section 4 discusses some related research and section 5 concludes.

2 Background

Carminati (2002) investigated the referring preferences of NSPs and OSPs in Italian and proposed that the variation between NSPs and OSPs is regulated by the Position of Antecedent Hypothesis, which is stated in 2.

- (2) Position of Antecedent Hypothesis: NSPs prefer to retrieve an antecedent in the (highest) Spec(IP), whereas OSPs prefer an antecedent in a lower syntactic position.

The PAH was tested in several intrasentential contexts, that is, the pronoun and its possible antecedents were always in the same sentence, as shown in 3.

- (3) a. Null Pronoun
Marta scriveva frequentemente a Piera quando \emptyset era negli Stati Uniti.
“Marta wrote frequently to Piera when \emptyset was in the United States.”
- b. Overt Pronoun
Marta scriveva frequentemente a Piera quando lei era negli Stati Uniti.
“Marta wrote frequently to Piera when she was in the United States.”

Carminati used items such as the one in 3 in a questionnaire study and she found a clear division of labor between the two pronouns. NSPs were mostly interpreted as referring to the subject of the previous sentence (Marta, in the example); this was the case for 80.7% of the experimental items.

*I wish to thank Robin Clark, Aviad Eilam, Satoshi Tomioka, Enric Vallduví, Charles Yang, three anonymous reviewers and the audience of PLC33 for very helpful comments and feedback.

¹I use the empty set to represent NSPs in examples.

In contrast, OSPs were mostly interpreted as referring to the object of the previous sentence (Piera, in the example); this was the case for 83.3% of the items. The same pattern was found for Catalan intersententially, that is, with the pronoun and the possible antecedents located in different sentences (Mayol and Clark, 2010), as shown in example 4: NSPs showed a clear subject preference (70.3%) and OSPs showed a clear object preference (64.5%).

- (4) a. La Marta escrivia sovint a la Raquel. \emptyset Vivia als Estats Units.
 “Marta wrote frequently to Raquel. \emptyset Lived in the United States.”
 b. La Marta escrivia sovint a la Raquel. Ella vivia als Estats Units.
 “Marta wrote frequently to Raquel. She lived in the United States.”

While Carminati framed her proposal in purely syntactic terms, other researchers have proposed that it is the information structure of the previous discourse that is responsible for the pronouns’ preferences. This is the position of Samek-Lodovici (1996) and Frana (2007)². In fact, the experiments carried out by Carminati (2002) and replicated by Mayol and Clark (2010) do not allow us to distinguish between the two hypotheses, due to the way information structure is encoded in Romance languages.

In this paper, I adopt Vallduví’s (1992) model of information structure. In this approach, an utterance is divided into focus and ground, which is further subdivided into link and tail. The focus is the update potential of an utterance, its new information, while the ground is old information. The link (a notion similar to that of ‘topic’) is the linguistic material which indicates where the new information must be updated (i.e., in which file) and the tail indicates *how* this new information must be updated.

In Catalan, information structure is encoded through word order, which is relatively flexible: subjects can be preverbal or postverbal, and the other arguments of the verb can appear in situ or left and right dislocated. Links appear preverbally: both preverbal subjects and left-dislocated constituents are links. Postverbal constituents are focal information, while right-dislocated constituents are tails. Generally, there is a correlation between subjects and links: links are usually, but not necessarily, realized as subjects and the resulting sentence is SVO. However, in a left-dislocation with a postverbal subject, the roles are inverted: the left-dislocated constituent acts as a link and the postverbal subject is part of the focus.

Since in the experiments just reported all sentences were SVO (i.e., all subjects were preverbal), their results are compatible with two explanations: (1) NSPs and OSPs are sensitive to syntactic factors: NSPs have a preference for subject antecedents and OSPs a preference for object antecedents and (2) NSPs and OSPs are sensitive to information structure: NSPs have a preference for link antecedents and OSPs a preference for non-link antecedents. The goal of the experiment that I present in the next section is to find out which hypothesis is the correct one.

3 Experiment

The experiment consisted of a questionnaire study in which participants had to choose their preferred interpretation for NSPs and OSPs. The goal of the experiment is to test whether the referring preferences of NSPs and OSPs are regulated by syntactic or pragmatic factors. This was tested by manipulating the syntactic structure of the experimental items so that some conditions did not present the usual correlation between link and subject. §3.1 presents the methodological details of the experiment and §3.2 presents and discusses the results.

3.1 Methodology

Materials: the materials consisted of sixteen three-sentence discourses with four conditions created by crossing word-order (SVO vs. OVS) and type of pronoun (null vs. overt).

The discourse starts with a question about the referent that appears preverbally in the second sentence. The second sentence mentions two individuals by means of two proper names of the same

²I postpone the discussion of Frana’s experiments until section 4.

grammatical gender, one in subject position and the other in object position. In conditions 1 and 2, this sentence has an SVO order in which the subject is the link of the sentence (these conditions are the same as conditions 1 and 2 in Mayol and Clark, 2010). In the other two conditions (conditions 3 and 4), this sentence has an OVS order; the object has been left-dislocated and is the link, and the subject is postverbal and represents new information. Crucially, in conditions 3 and 4 the usual association between links and subjects is broken and the object (not the subject) is the link of the sentence.

Finally, the third sentence contains either an NSP or an OSP, which could in principle refer both to the previous subject or to the previous object. That is, the content of the third sentence is not pragmatically biased to refer to one of the two referents of the discourse.

An example in the four conditions can be seen in 5.

- (5) a. Cond 1: SVO + Null
 A: Què li va passar a la Marta?
 A: "What happened to Marta?"
 B: La Marta escrivia sovint a la Raquel. ∅ Vivia als Estats Units.
 B: "Marta wrote frequently to Raquel. ∅ Lived in the United States."
- b. Cond 2: SVO + Overt
 A: Què li va passar a la Marta?
 A: "What happened to Marta?"
 B: La Marta escrivia sovint a la Raquel. Ella vivia als Estats Units.
 B: "Marta wrote frequently to Raquel. She lived in the United States."
- c. Cond 3: OVS + Null
 A: Què li va passar a la Raquel?
 A: "What happened to Raquel?"
 B: A la Raquel, l'escrivia sovint la Marta. ∅ Vivia als Estats Units.
 B: "To Raquel, Marta wrote (to her) frequently. ∅ Lived in the United States."
- d. Cond 4: OVS + Overt
 A: Què li va passar a la Raquel?
 A: "What happened to Raquel?"
 B: A la Raquel, l'escrivia sovint la Marta. Ella vivia als Estats Units.
 B: "To Raquel, Marta wrote (to her) frequently. She lived in the United States."

OVS sentences are quite marked and unnatural in the absence of context. The role of the first sentence in the discourses (i.e., the question) was to provide some context so that OVS sentences would sound more natural.

The conditions for each item set were counterbalanced and incorporated into a questionnaire experiment together with 24 filler items and 5 practice items. Eight counterbalanced lists were constructed (the last four lists with the items in reverse order), with a single randomization for all lists.

Procedure: The experiment was administered using a laptop equipped with EPrime software. Before starting the experimental session proper, subjects read a set of written instructions which explained the experimental procedure. Participants went through a practice session, so that they could get familiar with the keyboard and the procedure, and the experiment subsequently began. The discourses were presented on the computer screen. Subjects were asked to indicate which interpretation of the second sentence they preferred, i.e., whether they thought it was a statement about the subject of the first sentence, or the object of the first sentence. Under each experimental sentence, two paraphrases of the second sentence were given, such as the following, corresponding to the example items presented above.

- (6) a. Marta lived in the United States
 b. Raquel lived in the United States

Subjects: Thirty-two members from the Universitat Pompeu Fabra community took part in this experiment.

Predictions: If syntactic factors drive the preferences of NSPs and OSPs, we expect conditions 1 and 3 to show a subject preference, and conditions 2 and 4 to show an object preference. If information structure factors drive the preferences of NSPs and OSPs, we expect conditions 1 and 4 to show a subject preference and conditions 2 and 3 to show an object preference.

3.2 Results and discussion

The results of the experiment can be seen in Table 1.

	subject antecedent	object antecedent
Cond 1: SVO + null	59.1	40.9
Cond 2: SVO + overt	35.2	64.8
Cond 3: OVS + null	58.0	42.0
Cond 4: OVS + overt	51.1	48.9

Table 1: Results for the experiment

Conditions 1 and 2 follow the pattern predicted by the PAH, although the results for condition 1 in the current experiment are not as strong as the ones previously reported in Carminati (2002) and Mayol and Clark (2010). Since in these conditions the subject is the link, these results are compatible both with a syntactic and an information structure hypothesis.

By contrast, the results for conditions 3 and 4, in which the subject is not the link, are not expected under either hypothesis. In condition 3, the NSP still shows a preference for the subject, and not for the object, which crucially is the link in these sentences. In condition 4, the OSP has no clear preference towards either the object or the subject. While conditions 1 and 2 are mirror images of each other, this is not the case for conditions 3 and 4. An ANOVA analysis of the frequency with which the subject antecedent was chosen in the four conditions was performed with subjects and items as random effects. The ANOVA analysis shows that whether an anaphoric element is interpreted as referring to the preceding subject depends on the type of pronoun (null or overt). That is, NSPs have a greater preference towards subject antecedents than OSPs ($F(1,31) = 7.02$, $p = 0.01$; $F(1,15) = 5.07$, $p = 0.04$). There is also a significant interaction between type of pronoun and word order, although it is only marginally significant by subjects ($F(1,31) = 2.32$, $p = 0.07$; $F(1,15) = 3.06$, $p = 0.04$).

As noted, the preferences in the SVO conditions are less clear in this experiment than in the experiment reported in Mayol and Clark (2010), although both results do point in the same direction. This can be attributed to the fact that this experiment added OVS items, which require a more developed discourse context than simple SVO sentences. The experiment presented here did include some more context than the experiments in Carminati (2002) and Mayol and Clark (2010) so that the OVS sentences would not sound completely unnatural. However, it seems that this was not enough and contributed to raise the overall variation found in the results.

These results clearly show that the two pronominal forms are not sensitive to the same factors: while NSPs have a simple subject preference, regardless of the pragmatic function of the subject, OSPs have an object preference only when the object is not the link and show no clear preference when the object is the link. In other words, NSPs are only sensitive to syntactic function, while OSPs are sensitive to both. These results support a notion of salience in which different factors play a role and different referential expressions are sensitive to different factors. In particular, both subjecthood and linkhood add to salience, but the former has a larger weight than the latter. NSPs, being the default pronominal form, have a preference for the most salient entity, the subject, which remains the most salient entity even if it is not the link (that is, both in condition 1 and condition 3). In other words, NSPs follow the Position of Antecedent Hypothesis. In contrast, OSPs have more constrained preferences and the Position of Antecedent Hypothesis needs to be modified: they are constrained to refer to low salience entities. When both factors contributing to salience (syntax and information structure) agree in marking a referent as low in salience, this is the one the OSP will prefer. When both factors do not agree (one potential antecedent is subject but non-link and the

other is non-subject but link), both potential antecedents have an intermediate degree of salience, there is no low salience antecedent and, therefore, OSPs do not show a clear preference for any of the candidates. This explains the contrast between condition 2, in which the OSP shows a clear preference for the object, non-link referent, and condition 4, in which the OSP does not exhibit a clear preference.³

The crucial difference between the two types of pronouns is that NSPs have a simple preference for previous subjects, while OSPs have more complex preference involving both syntactic and pragmatic factors. This proposal is very much in the spirit of Kaiser and Trueswell (2008), who propose, based on data from Finnish, that multiple constraints affect anaphora resolution (see Section 4.1 for discussion of their experiment.).

This experimental data points to the following interpretation: in Catalan, there are particular anaphoric forms that refer to antecedents at opposite ends of a scale of salience for activated referents. That is, NSPs refer to maximally salient antecedents (i.e., subjects) and OSPs to low salience antecedents in the immediate context (i.e., non-subjects, non-link constituents). In contrast, there is no dedicated pronominal anaphoric form to refer to an antecedent with an intermediate degree of salience.

The results of this experiment can be interpreted as a sign that hearers do not behave in a completely Gricean manner. That is, given that in condition 3 (OVS + null), there is a subject preference, we would expect to observe the usual division of labor and find an object preference in condition 4 (OVS + Overt). However, as argued, OSPs seem to be highly constrained to refer only to low salience antecedents and not to object links. It is plausible that the Gricean division of labor takes place only when the two interpretations are distinct enough. Opposite ends of a scale of salience for activated referents would fulfill this condition. In contrast, the intermediate degree of a scale of salience, represented by link, non-subject referents, is not distinct enough from the other two degrees of salience and, in consequence, no pronominal form shows a clear preference for referents with this intermediate degree of salience.

4 Related Experiments

In this section I review two related experiments for two different languages: Finnish and Italian. The experiment for Finnish was very similar to the experiment presented here and the results are strikingly similar, despite the fact that Catalan and Finnish are not related to each other. In contrast, the experiment for Italian, a language closely related to Catalan, yielded very different results. However, the manipulations carried out in the experimental items of both experiments are very different and this can explain the differences in the results.

4.1 Finnish

Kaiser and Trueswell (2008) have studied the interpretation of pronouns and demonstratives in Finnish. Finnish is a partial null-subject language (Holmberg, Nayudu, and Sheehan, 2009), which allows third person null-subjects only in very restricted circumstances.⁴ However, Finnish, like Catalan, has two types of third person anaphors, the pronoun *hän*, ‘s/he’ and the demonstrative *tämä*, ‘this’. Additionally, like in Catalan, word order in Finnish is flexible: SVO is the default word order, but OVS sentences are also possible and felicitous when the object is discourse-old information and

³The results of this experiment could also be explained by appealing to the different syntactic structures of SVO and OVS sentences. In particular, the behavior of the OSP could be explained as follows: OSPs have a preference for non-subject antecedents in a position lower than Spec(IP). In SVO conditions, this results in a preference for the object, while in OVS conditions there is no candidate that fulfills the conditions, the object having been left-dislocated in a position which cannot be lower than Spec(IP). In contrast, NSPs have a preference for subject antecedents, regardless of their syntactic position: in Spec(IP) for SVO sentences and in a lower position in OVS sentences. I thank Charles Yang for this observation.

⁴Namely, null-subjects are only allowed when the subject is a generic pronoun or when the subject is controlled by an argument in a higher clause.

the subject is discourse-new information. Therefore, Finnish and Catalan map informational status with sentence position very similarly. Kaiser and Trueswell carried out a sentence completion task. Participants were presented with small discourses of three sentences. The third sentence was either SVO or VOS and the fourth sentence started either with the pronoun *hän* or with the demonstrative *tämä*. The task of the participants was to complete this fourth sentence. A sample item is given in 7:

- (7) a. Sentence 1: Liisa astuu sisään erään suuren firman päätoimistoon.
 “Liisa steps into the main office of a big company.”
 Sentence 2: Hän huomaa sihteerin, joka puhuu puhelimessa.
 “She notices a secretary who is talking on the phone.”
- b. Sentence 3a. SVO: Hetken päästä sihtööri moittii juuri sisääntullutta liikemiestä.
 “After a moment the secretary-SUBJ criticizes a businessman-OBJ who has just walked in.”
 Sentence 3b. VOS: Hetken päästä sihtööriä moittii juuri sisään tullut liikemies.
 “After a moment the secretary-OBJ criticizes a businessman-SUBJ who has just walked in.”
- c. Sentence 4a. Hän ...
 Sentence 4b. Tämä ...

The results from this study can be seen in Table 2. Since *tämä* is a demonstrative, some of the continuations used it in this way. When the continuation was ambiguous, it was coded as ‘unclear’.

	Subject	Object	Demonstrative	Unclear
Cond 1: SVO + Hän	64	13	0	23
Cond 2: SVO + Tämä	0	88	9	3
Cond 3: OVS + Hän	64	13	0	23
Cond 4: OVS + Tämä	44	0	30	17

Table 2: Results in Kaiser and Trueswell (2008)

The results show that the pronoun *hän* is sensitive primarily to syntactic role and has a subject preference regardless of word order. In contrast, the demonstrative is sensitive to word order. It prefers postverbal referents, but this preference is modulated by the syntactic role of the antecedent: it prefers objects to subjects. It is remarkable that in the last condition, OVS-Tämä, there were many uses of *tämä* as a demonstrative (for example, ‘this was fun’), indicating that the anaphoric use of *tämä* is less felicitous in the VOS condition than in the SVO condition. Kaiser and Trueswell (2008) take these results to show that salience cannot be described by a single-factor concept; they point to a model with multiple constraints, in which referential forms can show different degrees of sensitivity to different factors.

Note that the results of the Finnish (*hän/tämä*) study are very similar to the results for Catalan: the most reduced form (NSPs and *hän*) shows a subject preference, while the other anaphoric form (OSPs and *tämä*) shows a preference for non-subjects, non-links. The main difference is that while the overt pronoun in the OVS condition in Catalan does not show a clear preference, the demonstrative *tämä* in Finnish showed a weak preference for the postverbal subject. However, the results for Finnish may have been affected by the fact that the demonstrative use of *tämä* acted as a ‘escape hatch’ in a situation where neither argument is a good antecedent. Catalan does not have such an escape hatch and participants just chose between the possible antecedents at random.

4.2 Italian

Frana (2007) also noticed that Carminati’s results were compatible with two explanations: with a syntactic account (i.e., the PAH) and with an information structure account. In her work, she aimed to test what she called the Discourse-Prominence Hypothesis of Antecedent Assignment (DPH), which states that NSPs have a link⁵ preference. Frana (2007) also entertains the Anti-Topic Hypothesis,

⁵She uses the term ‘topic’, instead of ‘link’.

according to which OSPs decrease their preference for non-subject antecedents when this position is occupied by a link. She tested the DPH by performing an experiment very similar to Carminati's experiment 1, but manipulating the items so that, according to her analysis, the immediately preceding subject does not always coincide with the link. The details of the experiment are the following:

Materials: the materials consisted of twenty two-sentence passages with four conditions. The first sentence introduces an individual by proper name (Referent 1). The second sentence is a complex sentence. In the subordinate clause, a new individual is introduced in subject position (Referent 2), while Referent 1 is repeated in object position, either by means of a full DP or a clitic. The main clause contains either an NSP or an OSP in subject position. The content of the second sentence is not pragmatically biased to refer to one of the two referents. Thus, the four conditions are:

- (8) a. Cond 1: full DP + null
 La signora Rossi è una persona molto maleducata che non merita alcun riguardo. Quando Maria incontra la signora Rossi per strada, \emptyset fa sempre finta di non vederla.
 "Mrs. Rossi is a very rude person that does not deserve any regard. When Maria sees Mrs. Rossi in the street, \emptyset pretends always not to see her."
- b. Cond 2: full DP + overt
 La signora Rossi è una persona molto maleducata che non merita alcun riguardo. Quando Maria incontra la signora Rossi per strada, lei fa sempre finta di non vederla.
 "Mrs. Rossi is a very rude person that does not deserve any regard. When Maria sees Mrs. Rossi in the street, she pretends always not to see her."
- c. Cond 3: clitic + null
 La signora Rossi è una persona molto maleducata che non merita alcun riguardo. Quando Maria la incontra per strada, \emptyset fa sempre finta di non vederla.
 "Mrs. Rossi is a very rude person that does not deserve any regard. When Maria her-sees in the street, \emptyset pretends always not to see her."
- d. Cond 4: clitic + overt
 La signora Rossi è una persona molto maleducata che non merita alcun riguardo. Quando Maria la incontra per strada, lei fa sempre finta di non vederla.
 "Mrs. Rossi is a very rude person that does not deserve any regard. When Maria her-sees in the street, she pretends always not to see her."

Frana assumes that this manipulation (clitic versus proper name) is able to distinguish subjects from links. The clitic in conditions 3 and 4 is supposed to reinforce the DP it corefers with and its discourse referent and, as a consequence, reinforce its topical status. In those cases, Frana assumes that the clitic is the link of the sentence, although it is not the subject. In contrast, the proper name in conditions 1 and 2 is supposed to not reinforce the DP and, consequently, the subject, and not the object, is supposed to be the link of the sentence. Carminati (2002) predicts that this manipulation should not produce any effect and that conditions 1 and 3, on the one hand, and conditions 2 and 4, on the other hand, should behave in the same way. In contrast, if the DPH is correct, the prediction is that conditions 1 and 3 should show a different pattern: the NSP should prefer the subject antecedent in condition 1 and the object antecedent in condition 3. As for condition 4, Carminati predicts object antecedent, while, in contrast, the Anti-Topic hypothesis would predict subject antecedent.

Procedure: Four counterbalanced versions of the questionnaire were created. 32 Italian native speakers completed the questionnaire via e-mail.

Results: The results of Frana's experiment can be seen in Table 3.

	Subject antecedent	Object antecedent
Cond 1: null + full DP	70	30
Cond 2: overt + full DP	27	73
Cond 3: null + clitic	35	65
Cond 4: overt + clitic	16	84

Table 3: Results from Frana (2007)

The results for condition 1 and condition 2 are parallel to Carminati's results for Italian and

my results for Catalan. When the subject acts as a link, NSPs have a subject preference and OSPs have an object preference. However, the pattern for condition 3 is different. The NSP does not refer to the immediately preceding subject, but to the subject of the first sentence, which appears as an object clitic in the subordinate clause of the second sentence. Frana (2007) takes this as supporting evidence for the DPH. In addition, the Anti-Topic hypothesis does not get support from the data, as shown in the results for condition 4. OSPs maintain their preference for the previous object, even when the object is acting as a link. Moreover, this preference is even stronger than when the object is not the link (condition 4 vs. condition 2). Thus, while in conditions 1 and 2, both types of pronouns have different preferences, in conditions 3 and 4, they behave in a similar way.

Discussion: These results show that the relationship between syntactic position and type of pronoun is not as straightforward as proposed in Carminati (2002). However, it is disputable that they provide conclusive evidence for the Discourse-Prominence Hypothesis. It can be argued that Frana's experimental conditions are not manipulating the information structure of the sentence, which in Romance languages correlates with word order. Following Vallduví's approach, if the update of the information is done clause by clause, the link of the experimental items by the time the pronoun is reached is the subject, *Maria*, in all four conditions. Moreover, it is very striking that in Frana's experiment, there does not seem to be a way to refer to the subject of the subordinate clause in conditions 3 and 4. That is, when there is a clitic coreferential with a previous antecedent, both pronouns have a preference for this referent. In what follows, I present two hypotheses of why this is the case, one in reference to the previous discourse and the other in reference to the following discourse.

First, Carminati's experimental items were all concerned with intrasentential anaphora and all of them consisted of a single, complex sentence. My own experimental items address intersentential anaphora and all consist of multiple simple sentences. In contrast, Frana's items have multiple sentences, one of which is complex. Thus, in her items, both intersentential and intrasentential anaphora play a role. An alternative explanation of the results could be that a pronominal clitic in a subordinate clause is a sign for future anaphoric pronouns to ignore this clause and look for a referent in some previous point of the discourse. That is, clitics may signal that resolution needs to be intersentential. If this is so, there is only one referent available in the previous sentence and this is what both pronouns end up referring to. In other words, the choice of referring expressions may affect discourse segmentation. A DP or a full noun phrase signals a new discourse segment, while a clitic signals a continuing discourse segment.

Of course, it is possible to remove the complexity added by pronouns potentially coreferring within and across sentences by constructing items such as the ones in 9, in which all potential coreferential relations are across sentences.

- (9) a. La signora Rossi è una persona molto maleducata che non merita alcun riguardo. Maria incontra Mrs Rossi spesso. Lei / \emptyset fa sempre finta di non vederla.
 "Mrs. Rossi is a very rude person that does not deserve any regard. Maria sees Mrs Rossi often. She / \emptyset always ignores her."
 b. La signora Rossi è una persona molto maleducata che non merita alcun riguardo. Maria la incontra spesso. Lei / \emptyset fa sempre finta di non vederla.
 "Mrs. Rossi is a very rude person that does not deserve any regard. Maria sees her often. She / \emptyset always ignores her."

Second, it could be that clitics trigger certain expectations about how the discourse will continue. Since Frana's sentences have quite a broad context, the concept of discourse topic, or d-topic, may be playing an important role in determining pronoun preferences. A d-topic is what the discourse segment is about (see Asher, 2004), and it has been argued that d-topics are crucial in explaining some discourse relations: for example, Asher and Lascarides (2003) require that all elements belonging to the the discourse relation Narration have a common d-topic. In conditions 3 and 4, the clitic signals that the subject of the previous sentence, 'la signora Rossi', is the d-topic of the Narration. Thus, it is expected that this d-topic will be maintained and that the speaker will add some more information about it. In other words, the fact that 'la signora Rossi' is interpreted as a d-topic is responsible for the fact that in both conditions 3 and 4 the pronoun must be coreferential with it. In contrast, in conditions 1 and 2, there are not enough linguistic cues to construe 'la signora Rossi'

as a d-topic, since there is no clitic reinforcing the subject of the previous sentence and, moreover, the name is repeated. A more general d-topic (such as ‘what is happening in our neighborhood’) is constructed and the coreferential pattern follows Carminati’s Position of Antecedent Hypothesis. In other words, a clitic can act as some sort of cataphoric marker, triggering the expectation that something else will be added about its referent. Both the overt and null pronoun fulfill this expectation and are interpreted as adding information about the clitic referent. In order not to fulfill this expectation and change the d-topic, a stronger cue than a pronominal form would be needed. This stronger cue could be a definite description or a proper name, as shown in 10. The missing coreferential pattern is achieved by placing a proper name in the main clause and a null pronoun in the subordinate clause.

- (10) Quando \emptyset la incontra per strada, Maria fa sempre finta di non vederla.
 “When sees her on the street, Maria pretends not to see her.”

Finally, this manipulation could be not related to topicality at all, since it has been suggested that pronominalization is one of the factors that adds to the complex concept of ‘salience’. Kameyama (1999) claims that pronominalized non-subjects gain in salience by virtue of being pronominalized and that they compete in salience with a non-pronominalized entity in subject position. However, this claim is partially disconfirmed by one of Carminati’s experiments (experiment 9). In this experiment, she tested the reading times of non-ambiguous sentences which contained a clitic pronoun vs. name manipulation, as in 11.

- (11) a. Condition a: subject antecedent + name
 Quando Maria cerca Roberto, \emptyset diventa ansiosa.
 “When Maria looks for Roberto, \emptyset becomes anxious (fem).”
 b. Condition b⁶: object antecedent + clitic
 Quando Maria lo cerca, \emptyset diventa ansioso.
 “When Maria him looks for, \emptyset becomes anxious (masc).”

Condition A (subject DP antecedent) was read faster than condition b (object clitic antecedent): the average reading time for condition a was 1358 ms., while it was 1537 ms. for condition B. This is not consistent with the Discourse Prominence Hypothesis, which would predict the opposite: sentences with NSPs should be read faster when the antecedent is a clitic than when it is a proper name. By contrast, these results are consistent with the results of the experiment for Catalan, since it reinforces the idea that NSPs have a simple subject preference. Note that in these experimental items, Carminati introduced the clitic out of the blue and there was not enough context to create a d-topic.

In sum, I have argued that Frana’s experiment and the experiment presented here yielded different results because of the way experimental items were manipulated. I have suggested that the discourse structure in Frana’s experiments may be responsible for the pronouns’ preferences when a clitic is present in the sentence.

5 Conclusions

This paper presents an experiment to test which factors drive the referring preferences of null and overt subject pronouns in Catalan, a null-subject language. In order to be able to distinguish syntactic and pragmatic factors, the word order of the item sentences was manipulated. The results of the experiment point to a complex notion of saliency, in which both syntactic and pragmatic function play a role, the former having a greater weight than the later. Different forms are sensitive to different factors. NSPs have a preference for subject antecedents (which is the most salient antecedent even if it is not the link); that is, they are sensitive only to syntax. By contrast, OSPs have more complex preferences and have a preference for non-salient (non-subject, non-link) antecedents, if there is one: they are sensitive both to syntax and information structure.

⁶This was condition c in Carminati’s experiment. She tested other factors which are not relevant for our purposes here.

Neither the syntactic account (the PAH) nor the purely information structure hypothesis can account for the data obtained through the experiment, and a multi-factor notion of saliency is needed to accommodate the fact that both factors play a role, particularly for overt pronouns.

References

- Ariel, Mira. 1990. *Accessing Noun Phrase Antecedents*. London: Routledge.
- Asher, Nicholas. 2004. Discourse topic. *Theoretical Linguistics* 30:163–201.
- Asher, Nicholas, and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.
- Carminati, Maria Nella. 2002. The processing of Italian subject pronouns. Doctoral Dissertation, University of Massachusetts.
- Frana, Ilaria. 2007. The role of discourse prominence in the resolution of referential ambiguities. Evidence from co-reference in Italian. *UMOP 37: Semantics and Processing*.
- Holmberg, Anders, Aarti Nayudu, and Michelle Sheehan. 2009. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica* 63:59–97.
- Kaiser, Elsi, and John Trueswell. 2008. Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes* 23:709–748.
- Kameyama, Megumi. 1999. Stressed and unstressed pronouns: Complementary preferences. In *Focus: Linguistic, Cognitive, and Computational Perspectives*, 306–321. Cambridge University Press.
- Mayol, Laia, and Robin Clark. 2010. Pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics* 42.
- Samek-Lodovici, Vieri. 1996. Constraints on subjects: an Optimality Theoretic analysis. Doctoral Dissertation, Rutgers University.
- Vallduví, Enric. 1992. *The Informational Component*. New York: Garland.

University of Pennsylvania
619 Williams Hall
Philadelphia, PA 19104
laia.mayol@gmail.com