The information status and discourse anchorage of non-nominal constituents: a case study on Spanish secondary predicates

[**preprint** of: Heidinger, Steffen. 2015. The information status and discourse anchorage of non-nominal constituents: A case study on Spanish secondary predicates. Journal of Pragmatics 81. 52–73. http://doi.org/10.1016/j.pragma.2015.03.012]

Abstract: The information status of nominal constituents has been a recurring topic of research over recent decades. Little is known, however, about the information status of non-nominal constituents such as secondary predicates. In the present paper, we present a corpus based analysis of the information status (and also the discourse anchorage) of Spanish secondary predicates. We will show that secondary predicates are anchored both in the preceding and the following context, but to a much lesser extent than their subjects (i.e., the nominal basis of comparison in our study). The most frequent information status of secondary predicates is neither *given* nor *new*, but one where the state denoted by the secondary predicate is evoked by some element of the preceding context (e.g., *to cry* evokes the state *sad*). *Evoked* is an intermediate information status between *given* and *new*. While the lack of given secondary predicates is not surprising, the existing literature would suggest that secondary predicates express predominately new information. In this respect, our study shows not only that information status is a relevant category for non-nominal constituents, but also that its analysis requires a more elaborate inventory than a binary given–new distinction.

Key words: information structure, givenness, secondary predicates, Spanish

1 Introduction

Much linguistic data can be better understood once we take into account the information status of linguistic expressions, i.e., whether they express *given* or *new* information. It is therefore not surprising that information status has received considerable attention. One domain where information status has proven particularly useful is that concerning the formal types of nominal constituents, such as pronouns, definite NPs, indefinite NPs. It has been

shown that information status is a good predictor for the occurrence of the various formal types (cf. Prince, 1981; Gundel et al., 1993; Ariel, 2001, 2014; Roberts, 2003).[1]

This close relation between form and information status in the nominal domain is, however, not the main reason why most research on information status is devoted to nominal constituents (while at the same time, the information status of non-nominal constituents has received very little attention). The main reason seems to be that information status is often taken as a category which is relevant for nominal constituents only. In her seminal paper on information status, Prince (1981: 235) explicitly states that only nominal constituents express discourse entities – which explains why she does not consider non-nominal constituents (cf. also Lambrecht (1994: 37, 75) for a similar view). In the same vein, Dipper et al. (2007: 151) state in their annotation guidelines for information structure that information status is only relevant for "[a] constituent which refers to a discourse entity; mostly referential NPs or PPs, or their pronominal counterparts, unless part of an idiom […].". Yet another annotation scheme where only NPs are considered is Nissim et al. (2004).

In other studies, however, non-nominal constituents have been included in the study of information status. Chafe (1976: 28) and Riester (2008: 85), for example, explicitly state that non-nominal constituents such as verbs have information status as well. Further, Ariel (1985) analyzes the discourse functions of given information and considers not only nominal, but also non-nominal constituents such as VPs or clauses. Horn (1991) analyzes informationally redundant (i.e., *given*) VPs. In the present paper, we will investigate the information status (and also the discourse anchorage) of another non-nominal constituent, namely secondary predicates.

Secondary predicates are constituents such as *drunk* in (1), which describe the state of a participant during the event denoted by the verb.[2]

(1) John$_i$ bought his new car drunk$_i$.

As concerns the information structural properties of secondary predicates, it has been repeatedly argued in the literature on Spanish that secondary predicates have a strong

---

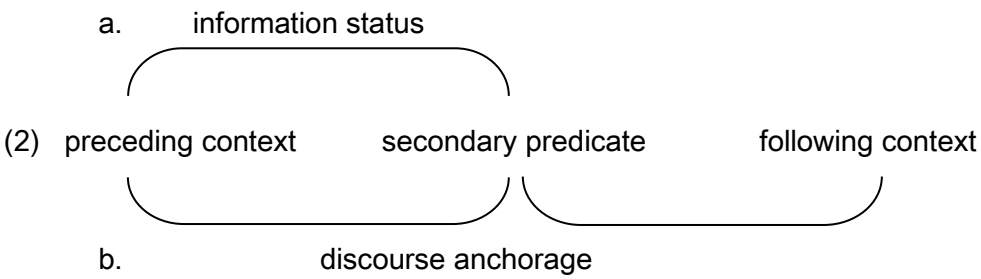
[1] Other domains where information status is relevant are the linear ordering (cf. Arnold et al., 2000) and the prosodic realization of constituents (cf. Baumann, 2006).

[2] There exist different types of secondary predicates. Distinctions have been made with respect to the participant of the event that the secondary predicate predicates over (subject- vs. object-oriented secondary predicates) and the syntactic status of the secondary predicate (argument vs. adjunct) (cf. Schultze-Berndt and Himmelmann, 2004; Himmelmann and Schultze-Berndt, 2005; Irimia, 2005; Rothstein, 2006, 2011; Winkler 1997). In the present study, only the type of secondary predicate illustrated in (1) will be considered: a depictive subject-oriented secondary predicate which is an adjunct. General descriptions of Spanish secondary predicates can be found in Hernanz (1988), Guemann (1990), Suñer (1990), Demonte (1991), Demonte & Masullo (1999), Hummel (2000), Gumiel (2005, 2008) and the references cited there.

affinity to focus. Guemann (1990: 200) states that in sentences with secondary predicates, the secondary predicate is the obvious rhematic element, and therefore it typically occurs in sentence final position and bears the main stress. Similarly, Rodríguez Espiñeira (1992: 53) states that the secondary predicate expresses the most important information of the sentence. For Porroche Ballesteros (1990: 157) the secondary predicate is the constituent of the sentence with the highest rhematic value; reflexes of this rhematic value are (i) the frequent occurrence in sentence final position and (ii) the fact that if the sentence has a negative particle, it has scope over the secondary predicate (cf. also Hummel (2000: 147) for this latter aspect).[3]

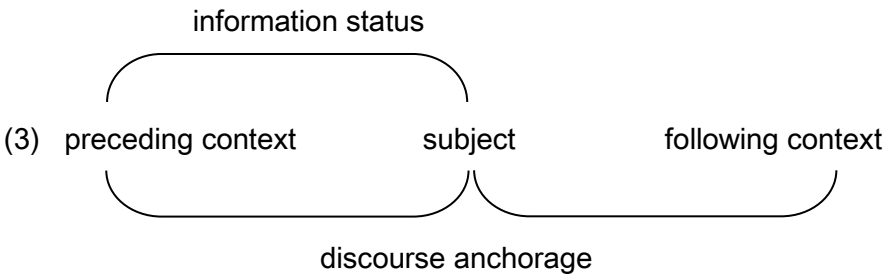
Although the above statements from the literature are about focus, they may also be instructive with respect to information status. Given that information status and focus typically align (cf. also Section 2.1), we may assume that constituents with affinity to focus also show an affinity to expressing new information. But neither the affinity of SPs to focus nor their affinity to expressing new information has been systematically investigated. In order to investigate the information status of secondary predicates, we will analyze corpus examples of sentences with secondary predicates and determine to what extent the secondary predicate is already anchored in the preceding context (cf. (2a.)); simplifying somewhat, secondary predicates which are already anchored in the preceding context have the information status *given*, while those that are not are *new*.

In addition to the preceding context, we will also look at the context following the secondary predicate. The following context is relevant because it allows us to track the fate of the information expressed by the secondary predicate: Is it picked up in the following context or not? The event boundness of secondary predicates might suggest that the information will not have a long lifespan after its mention as a secondary predicate. Preceding and following context taken together show us the discourse anchorage of the secondary predicate (cf. (2b.)).



[3] In this paper we focus on what might be called intrinsic information structural properties, such as the affinity to new (vs. given) information. The extrinsic information structural properties, i.e., the syntactic, prosodic and morphological behavior of secondary predicates under various information structural conditions, are not a topic of this paper.

Besides investigating the information status and discourse anchorage of secondary predicates, we will also look at another element in the sentence. We will analyze the information status and discourse anchorage of the sentence's subject, i.e., the element that the secondary predicate predicates over (cf. *John* in (1)). By taking into account the sentence's subject, we can compare the behavior of secondary predicates, which are non-nominal constituents, with that of a nominal constituent (cf. (3)).



The research goal of this paper is to investigate the information status and the discourse anchorage of secondary predicates, and to compare them with the information status and the discourse anchorage of the subjects of sentences with secondary predicates. By doing so, we will take a first step towards the systematic study of the information status of secondary predicates.

In order to verify the extent to which secondary predicates and their subjects are anchored in the discourse, we will look for correspondents in the preceding and following contexts. Correspondents are (simple or complex) expressions which are semantically linked to the SP's or the subject's denotation. In Section 3.2 we will describe these semantic relations and thereby specify the expressions which count as correspondents. To give an example, an expression such as *cry a lot* in the preceding context of the SP *sad* would count as a correspondent because the expression *cry a lot* evokes the state denoted by the SP *sad*.

On the basis of a corpus study on Spanish secondary predicates we will defend the following empirical claims:[4]

– Secondary predicates are anchored both in the preceding and the following context.

[4] The choice of the investigated language, i.e. Spanish, and the specific non-nominal constituent, i.e. secondary predicate, is motivated by the fact that this paper is embedded in a larger research agenda on the information structural properties of Spanish secondary predicates. In this research we investigate different levels of information structure, and both phenomena which seem to be language specific and phenomena which are not. As concerns information status and discourse anchorage, we assume that the empirical results reported in Section 3 are not language specific.

– The discourse anchorage of secondary predicates differs both in quantity and quality from that of subjects: subjects have more correspondents in the context than secondary predicates and they maintain a closer semantic relationship with their correspondents than secondary predicates.
– In the majority of cases, secondary predicates have a correspondent in the preceding context, which shows that secondary predicates typically express non-new information. However, the secondary predicates and their correspondents maintain a rather weak semantic relationship and the secondary predicate hardly ever expresses given information.
– The typical configuration is that a secondary predicate which is evoked (and thus neither new or given) predicates over a given and highly accessible subject.
In the discussion of these empirical findings we will focus on two aspects: (i) How does the finding that secondary predicates often express non-new information fit with the above cited literature on the information structural properties of secondary predicates?; (ii) Why do secondary predicates hardly ever express given information?
The remainder of the paper is structured as follows. In Section 2, we will lay out the theoretical basis for our investigation and introduce in further detail the concepts of information status and discourse anchorage, discuss the relation between the two concepts, and argue that information status and discourse anchorage are also relevant for non-nominal constituents such as secondary predicates. Section 3 forms the core of this paper and is devoted to a corpus based empirical study. In this section, we will lay out the method and the empirical basis, describe the annotation procedure, and present and discuss the empirical results. The paper ends with a brief summary of the main findings in Section 4.

2 Basic concepts

2.1 Information status in the nominal domain

The information expressed by the various parts of a sentence usually does not have the same informational value. In (4B), for example, the information expressed by *a man from Florida* has a different informational value than the information expressed by *bought* and *it.* While *a man from Florida* is new information, *bought* and *it* are given (or old) information.

(4) A: Who bought this car?
B: A man from Florida bought it.

The level of description on which differences in informational value are captured is commonly referred to as *information structure*. An important advance in the study of information structure over recent years has been the refinement of the descriptive inventory. In a survey article on information structure, Krifka (2007) convincingly shows that different levels of information structure need to be distinguished: focus (as indicating alternatives), information

status (as indicating whether or not a denotation is already present in the Common Ground of speaker and hearer), topic (as specifying what a statement is about) and frame setting/delimitation.

The information structural partitions of a given sentence on the first three levels often align (i.e. *topic ~ background ~ given* and *comment ~ focus ~ new*); but since this is not always the case the levels must still be considered separately. In (5B) the focus of the sentence contains given information; this shows that *focus* and *new* do not always align. In (6B) the comment of the sentence contains both focal (*in 1968*) and background information (*married her*); this shows that focus and comment do not always align.

(5) A: I know that John stole a cookie. What did he do then?
B: He [returned [the cookie]$_{\text{Given}}$]$_{\text{Focus}}$
(Krifka 2007: 40)

(6) A: When did [Aristotle Onassis]$_{\text{Topic}}$ marry Jacqueline Kennedy?
B: [He]$_{\text{Topic}}$ [married her [in 1968]$_{\text{Focus}}$]$_{\text{Comment}}$
(Krifka 2007: 42)

The sub-level of information structure that is at the center of this paper is information status. In example (4) we have used the distinction between given and new information to introduce the level of information status (cf. Haviland and Clark, 1974; Chafe, 1976). However, over recent decades the inventory of information status has been considerably augmented. Prince (1981: 235-237) makes a threefold distinction between *new*, *inferable* and *given* (*evoked* in her terms) discourse entities – with further subdivisions of all three statuses.[5] New entities are those that are newly introduced into the discourse by the speaker. Given entities are those that are already part of the discourse, and thus correspond to what has been labeled as *old* or *given* information in other inventories. Inferable entities are those for which the speaker assumes that the hearer can infer it from other discourse entities. In (7), for example, the entity DRIVER is inferable from BUS through the knowledge that busses have drivers.

(7) inferable
I got on <u>a bus</u> yesterday and <u>the driver</u> was drunk.
(Prince 1981: 233)

The most interesting domain of information status is the intermediate category between *new* and *given*, i.e., cases where a discourse entity is neither new to the discourse nor given by the preceding context. This is the domain of *bridging*-phenomena as in (7) (cf. Clark, 1975;

[5] Cf. Nissim et al. (2004), Dipper et al. (2007) and Riester (2008) for other inventories of information status.

Irmer, 2011) where a discourse entity maintains an indirect relation to an entity of the preceding context (e.g., *driver* + *bus*). In the case of non-nominal constituents, very little is known about such indirect relations to the preceding context. In Section 3, we will see that secondary predicates frequently maintain such indirect relations to entities in the preceding context.

2.2 Discourse anchorage in the nominal domain

A discourse typically consists of various discourse entities. In the discourse in (8), for example, we may identify as discourse entities the person named *John* (JOHN), the person named *Mary* (MARY), a blue shirt (BLUE.SHIRT) and a green shirt (GREEN.SHIRT).

(8) John and Mary are dancing. He wears his blue shirt and she wears her green one.

Further, we observe that such discourse entities can be mentioned more than once in a given discourse. In (8) JOHN is mentioned twice (as *John* and as *he*), MARY is mentioned twice (as *Mary* and as *she*), BLUE.SHIRT is mentioned once (as *his blue shirt*), and GREEN.SHIRT is mentioned once (as *her green one*). Discourse entities thus may differ with respect to how often they are mentioned in a given discourse. In the present paper, we use the term *discourse anchorage* for the dimension on which JOHN and MARY on the one hand and BLUE.SHIRT and GREEN.SHIRT on the other differ from each other. JOHN and MARY are anchored more deeply in the discourse than BLUE.SHIRT and GREEN.SHIRT because they are mentioned more often. But discourse anchorage not only has this quantitative, but also a qualitative side to it.

The necessity to consider the qualitative aspects of discourse anchorage as well can be seen in example (9).

(9) The Smith family often goes to the park. The son likes to play soccer there.

In this discourse, the discourse entity SON is mentioned by means of *the son*, but it can be argued that it is mentioned also by means of *the Smith family* (of which it is a part). If one accepts that SON is mentioned twice in the discourse, this entity would have the same quantitative discourse anchorage as, for example, JOHN in (8). But there is an obvious difference between the two mentions of JOHN in (8) and SON in (9): While the two mentions of JOHN are coreferent (i.e., referentially identical), this is not the case for SON. Since the qualitative relation is closer in the case of JOHN, one might argue that, on the qualitative side, JOHN is more deeply anchored in the discourse in (8) than SON is in the discourse in (9).

2.3 The relation between information status and discourse anchorage

Information status and discourse anchorage are related in a peculiar way. In order to illustrate this, let us assume two discourse entities A and B which have a very different

quantitative discourse anchorage; the first entity is mentioned only once, the second is mentioned ten times in the discourse. Entity B is thus anchored much more deeply in the discourse than A. The consequence for the information status of the two entities is as follows: A and B are new in their first mentions, and since the first mention is the only mention in the case of A, this entity is new in all its uses. Entity B, however, is mentioned nine more times after its first mention and in these nine further mentions it has the information status of being given (or old) information. Hence, the discourse anchorage of an entity has a direct impact on its information status; if an entity is deeply anchored in a discourse in the sense that it is often mentioned, then this entity often has the status of being given information. In our hypothetical example, A has a givenness-rate of 0% while B has one of 90%.

A case where the impact of discourse anchorage on information status becomes more interesting than in the above hypothetical example is when it can account for systematic differences in the information status of types of linguistic constituents. Such a systematic difference supposedly exists on the level of syntactic functions between secondary predicates (typically new information) and subjects (typically given information).

2.4 Information status and discourse anchorage: Which constituents can have it?

In the previous subsections we have introduced information status and discourse anchorage using nominal constituents. But the main interest of this paper is the information status and discourse anchorage of a non-nominal constituent, namely secondary predicates. Thus, we need to come back to an issue that has already been briefly alluded to in Section 1: For which types of constituents are information status and discourse anchorage relevant categories at all? Information status and discourse anchorage are relevant for constituents expressing discourse entities, and the concept *discourse entity* figures prominently in influential work on information status (e.g. Prince 1981). Although it becomes apparent below that expressing a discourse entity may not be a necessary condition for having information status, we initially provide several hints that non-nominal constituents might express discourse entities.

Riester (2008: 85) discusses the information status of the verb *evakuieren* 'evacuate' in (10) and argues that it is related to *Abtransport* 'evacuation' in the preceding discourse because the events expressed by the noun and the verb are linked. One might thus assume that the noun and the verb are linked via a common discourse entity.

(10) Der Abtransport der Zivilbevölkerung aus der von den bosnischen Serben besetzten Uno-Schutzzone Cepa ist jetzt fast abgeschlossen. Zwischen 4000 und 5000 Menschen sind bereits evakuiert.
'The transportation of the civilian population from the UN protected area Cepa, occupied by the Bosnian Serbs has almost been completed by now. In between 4000 and 5000 people have already been evacuated.'
(Riester 2008: 85; mod. StH)

Hegarty (2003: 923) points out several properties of events (as expressed by verbs) that they share with entities: "An event [...] occurs at a particular time and place, or it has particular spatiotemporal boundaries." Similarly, Maienborn (2005: 303) describes Kimian states as abstract objects for the exemplification of a property P at a holder x and a time t (cf. Asher, 1993 on abstract objects).
Another hint might be that anaphoric (or cataphoric) expressions are not limited to nominal constituents, but are also found with verbs and adjectives. *So did,* in the second part of example (11), picks up the information given by the preceding verb *dance*. In the same way, *así* in (12) is linked to the state expressed by the following adjectival constituents.

(11) John danced and so did Mary.

(12) pasaron la noche así: entre despiertos y dormidos, solos en el mundo, [...]
'They spent the night like that: between awake and asleep, alone in the world'
(Rodoreda; cited after Hummel 2000: 175)

Despite these similarities between nominal and non-nominal constituents, there are important differences which have to be kept in mind. First, differences in the status of the respective information do not have the same consequences for the morphosyntactic form of the constituent in the nominal and the non-nominal domain. Crucially, we do not find for non-nominal constituents an equivalent to the definite–indefinite distinction which is pervasive throughout languages in the nominal domain. Hence, the identifiability of an entity, which is crucial in the nominal domain, does not seem to be relevant for non-nominal constituents (cf. Lambrecht 1994: 111). Second, differences between nominal and non-nominal constituents also become apparent in the use of anaphoric expressions (cf. 13).

(13) a. John$_i$ is currently running in the woods and he$_i$ is wearing a blue shirt.
b. John [is running in the woods]$_i$ and Paul is doing [it]$_i$ as well.

In (13a.), we interpret the anaphora to be coreferent with the antecedent, i.e., they refer to the same entity in the world and we see the highest possible referential identity. In the case of the event-anaphora *it*, however, such a high degree of identity is not required. It might be that John and Paul are at the same time in the same woods doing the same thing, but

nevertheless do not participate in the same event. In fact, (13b.) would still be a true proposition if John and Paul are running miles apart from each other; a case in which they would definitely not be part of the same instantiation of a predication. Thus, the use of pro-forms is less restricted in the non-nominal domain. This might suggest that the uniqueness of an event or state, and thus also its identifiability, is less important in the non-nominal than in the nominal domain. Similarly, Dimroth et al. (2010: 3329) argue that what is maintained in the repeated use of predicative expressions are "[…] the properties characterizing a given situation, not the situation in the external world", while in the case of referential expressions repeated uses mean "[…] co-reference to the exact same referent."

As concerns the relevance of information status and discourse anchorage for non-nominal constituents we can draw the following conclusion: The denotation of non-nominal constituents can be semantically linked to the preceding context (cf. (10)) and the denotation expressed by non-nominal constituents can also be semantically linked to the following context (cf. (11) and (12)). Put in simplified terms, non-nominal constituents can pick up information from the preceding context and the information expressed by non-nominal constituents can be picked up in the following context. As a consequence, non-nominal constituents can be (at least) *new* or *given*, and they can be more or less anchored in a discourse. This makes both information status and discourse anchorage relevant descriptive categories for non-nominal constituents − irrespective of whether they express discourse entities or not. As concerns the latter question, the observed differences between nominal and non-nominal constituents may suggest that only nominal constituents express discourse entities. The differences between the secondary predicates and their subjects which are presented in the following section may also be interpreted in this way.

3 Empirical study

3.1 Method and data basis

In this study, we investigate the information status and discourse anchorage of secondary predicates and their subjects on the basis of a corpus of 50 sentences with a subject-oriented non-argumental secondary predicate (cf. (1)). In addition to the sentence itself, we consider a context window of about 500 words (250 before and 250 after the secondary predicate) and analyze what is going on in this context window: We will search for correspondents of the secondary predicate and the subject and code the respective example with respect to the quantity and quality of these correspondents. Correspondents are (simple or complex) expressions which are semantically linked to the SP's or the subject's denotation. In Section 3.2., we describe the semantic relations which must hold between an

expression and the SP or the subject in order to count the respective expression as a correspondent.

The 50 examples for our study stem from the corpus of examples underlying Hummel (2000). The examples in his corpus stem from present-day novels by authors of Iberian Spanish. This corpus was chosen as a data source because information in novels is potentially maintained and picked up again (length of text, narrative structure). Our selection of the examples was random except for the following factors: examples should not contain coordination of secondary predicates; the secondary predicate in the example should not lean towards an adverbial interpretation; not more than ten examples per author.

3.2 Annotation categories and coding decisions

Each of the 50 examples contains a secondary predicate (SP) and its subject (S) that it predicates over. The coding of the examples for further analysis refers to these two elements and their correspondents in the preceding and following context, henceforth *pre-text* and *post-text* (context window: 500 words).

In the analysis of the data, we are interested in quantitative as well as in qualitative aspects. As concerns quantity, we are interested in how many correspondents the secondary predicate and the subject have. As concerns quality, we look at the semantic and the formal properties of the correspondents. For the number of correspondents, we count all correspondents within the 500-word window. For the qualitative properties of the correspondents, we only consider the closest two correspondents in pre- and post-text respectively. While the number of correspondents is determined by simple counting, the annotation of the formal and semantic properties of the correspondents needs some further description. Generally speaking, their semantic properties are determined by the semantic relation between the correspondent and SP or S. The formal properties are the morphological or syntactic form of the correspondent.

We will begin with the annotation categories of SP, and here with the semantic properties. The semantic properties of the correspondents rely on the semantic relation between the correspondent and the secondary predicate. In the annotation of the correspondents we make a threefold distinction between *same*, *similar* and *evoked*.

(14) semantic properties of correspondents of SP

a. same

b. similar

c. evoked

There is a basic difference between *same* and *similar* on the one hand and *evoked* on the other. While correspondents with the relation *same* and *similar* denote states which are the

same or similar to the ones denoted by the secondary predicate, correspondents with the relation *evoked* merely evoke the state denoted by the secondary predicate.[6] For the sake of illustration let us compare the following three contexts for the secondary predication (*he*) *went home sad.*

(15) a. He was unhappy$_{CORR}$ and he went home sad$_{SP}$. same

b. He felt disappointed$_{CORR}$ and he went home sad$_{SP}$. similar

b. He [cried a lot]$_{CORR}$ and then he went home sad$_{SP}$. evoked

In (15a.) the secondary predicate *sad* has a correspondent in the preceding context which denotes the same state as the secondary predicate itself (namely *unhappy*). *Disappointed* in (15b.) on the other hand denotes not the same, but a similar state as the secondary predicate *sad*. Finally, *cry a lot* in (15c.) does not directly denote a state that is the same or is similar to the one denoted by the secondary predicate *sad*, but instead evokes the state *sad* since crying is typically related to the emotion of sadness.

Several remarks are in order with respect to this threefold distinction. First, *same* and *similar* are more direct relations than *evoked*. Second, we are aware that the distinction between *same* and *similar* is probably a gradual and not a categorical one. Under the assumption that no full synonyms exist, it would even be impossible to have the relation *same* other than in the case of repetitions of the same lexeme such as *He was very sad and he went home very sad*. Third, since the state that is evoked may either be the same as that denoted by the secondary predicate or may be similar to the secondary predicate, it would be desirable to distinguish between *same* and *similar* within the cases of *evoked*. However, since the classification of an evoked state with respect to the gradual distinction between *same* and *similar* is rather uncertain, we refrain from this further classification.

As concerns the formal properties of the correspondents of secondary predicates, we will distinguish the seven categories in (16). In the above examples, the correspondents with the direct relations *same* and *similar* have the form of an adjective while the correspondent with the relation *evoked* has the form of a verb. Note, however, that *same* and *similar* correspondents are not limited to the form of adjectives; e.g., the state *sad* can have non-adjectival direct correspondents such as *his sadness*.

[6] Our use of the term *evoked* thus differs from the one in Prince (1981).

(16) formal properties of correspondents of SP

a. SP (lexical)
b. SP (pro-form)
c. copula (lexical)
d. copula (pro-form)
e. NP internal adjective
f. state noun (e.g. *sadness*)
g. other

We now turn to the coding of the correspondents of S for which we cannot use the same coding categories as in the case of SP. As concerns the formal properties, the necessity for other categories is obvious. But the same holds for the semantic properties of the correspondents. Recall from Section 2.4 that we expect nominal and non-nominal categories to have different semantic relations to their correspondents. In the literature on the information status of nominal constituents, various inventories of relations have been proposed. These inventories differ in size and content. Since the information status and discourse anchorage serves mainly as a base for comparison for the secondary predicates for us we will use a rather reduced inventory of categories for correspondents of S.

(17) semantic properties of correspondents of S

a. same (= coreferent; e.g., *he* and *he* in (15a))
b. part-of-group (e.g. *the family* contains *the son* (as in (9), Section 2.2))
c. evoked (e.g. *a bus* evokes *the driver* (as in (7), Section 2.1))

As for the formal properties of the correspondents of S, we use the inventory of the six different coding categories given in (18).

(18) formal properties of correspondents of S

a. zero-pronoun
b. overt pronoun
c. proper name
d. definite or quantized NP
e. indefinite NP
f. bare NP

The total number of annotation categories amounts to 28. A table summarizing the categories can be found in the appendix.

3.3 Results

3.3.1 Frequency of correspondents and form of the SP's subject

We begin with some general observations on the frequency of correspondents for SP and S in the context windows. Table 1 and Figure 1 give the absolute frequency of correspondents in the context of the 50 examples.

	pre-text	post-text	total
SP	62	44	106
S	426	376	802
total	488	420	908

Table 1: Absolute frequency of correspondents

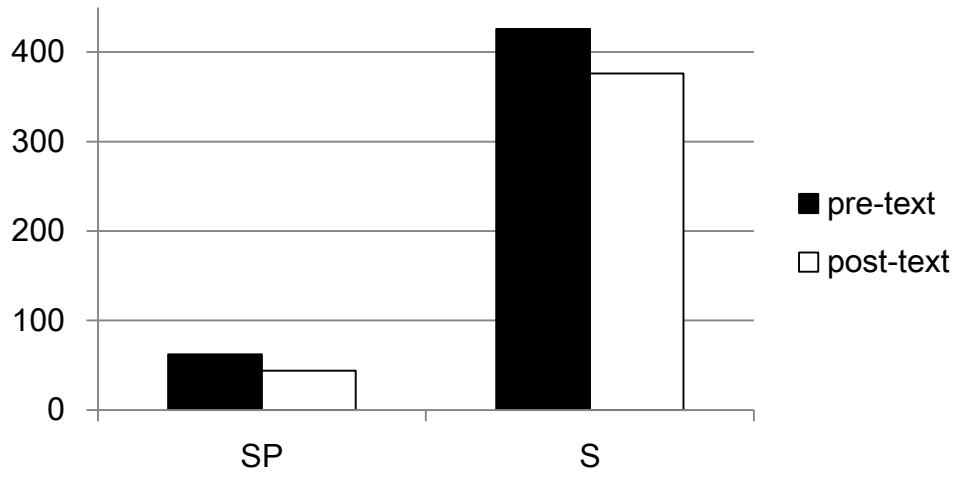


Figure 1: Absolute frequency of correspondents

For the 50 SPs a total of 106 correspondents and for the 50 Ss a total of 802 correspondents have been identified in the context windows. The average number of correspondents for each set of context windows amounts to 2.12 for SP and 16.04 for S (cf. Table 2). As concerns their distribution in the preceding and following context, the correspondents of both SP and S have a nearly equal frequency in the pre-text and the post-text with a slight dominance of the pre-text in both cases (cf. Figure 1).

	pre-text	post-text	total
SP	1.24	0.88	2.12
S	8.52	7.52	16.04

Table 2: Average frequency of correspondents (per example)

The results show that both SP and S have correspondents in the preceding and the following context, and that both of them tend to be anchored in the discourse. However, there is a

considerable quantitative difference in that correspondents of S are nearly eight times more frequent than correspondents of SP. Thus, our results clearly suggest that, from a merely quantitative perspective, Ss are more deeply anchored in the discourse than SPs. Despite this quantitative difference it needs to be stressed that discourse anchorage is not limited to the nominal constituent S, but can also be found in the case of the non-nominal SP.

As for information status, we need to compare the number of SP and S with a correspondent in the preceding context to the number of SP and S without one. Table 3 provides the respective numbers.

	none	1 or more	total
SP	17	33	50
S	4	46	50

Table 3: Correspondents in pre-text

Of the 50 SPs, 33 have at least one correspondent in the preceding context of 250 words. This shows that in the majority of cases the secondary predicate expresses non-new information. In the case of S, the tendency to express non-new information is even stronger.[7] In the next section and in the discussion of the results we will take a closer look at the qualitative side of this finding. Crucially, we will see that the SP and their correspondents usually maintain a rather weak semantic relation, and that therefore being a non-new SP still means to be rather new (cf. Section 3.3.2 and 3.4.1 on this matter).

Of the 50 Ss, 46 have at least one correspondent in the previous context of 250 words. Thus, the information status of the SP's subject is that it predominantly expresses non-new information. As we will see in the following section, the pre-text correspondents of S are typically coreferential with S (92% of the subjects have a coreferent antecedent in the previous context). An interesting formal correlate of this high number of coreferential antecedents can be found in the form of S. As Table 4 shows, zero-pronouns are by far the most frequent form of the secondary predicate's subject.

[7] The chi-square test shows that this difference between SP and S is statistically significant ($p<0.005$).

	abs.	%
zero-pronoun	27	54.00
overt pronoun	1	2.00
proper name	14	28.00
definite or quantized NP	8	16.00
indefinite NP	0	0.00
bare NP	0	0.00
total	50	100.00

Table 4: Form of S

The high frequency of zero-pronouns in Table 4, however, hints at more than just the presence of coreferent antecedents in the preceding context. Literature on the interpretation and resolution of anaphoric expressions has shown that a discourse often contains more than one potential antecedent for a given anaphoric expression (cf. Irmer 2011: 96). One way to restrict the interpretation of an anaphoric expression, i.e. to give the hearer some idea of where to look for the right antecedent, is to vary the form of the anaphoric expression. This is the basic idea of proposals by Ariel (2001, 2014) and Gundel et al. (1993).

Ariel (2001, 2014) holds that the various linguistic forms of referring expressions signal a discourse entity's accessibility in memory. Accordingly, these various forms are called *accessibility markers* and can be located on an *accessibility scale*. Ariel (2014: 28f.) argues that accessibility is determined by several factors, e.g. distance (between antecedent and anaphora) or competition (number of competitors on the role of antecedent). Gundel et al. (1993) propose an implicational hierarchy of six different cognitive statuses which depend on the location in memory and the attention state of the respective discourse entity (cf. Gundel et al. 1993: 274); e.g. referents that are *in focus* are in short-term memory and at the current center of attention (Gundel et al. 1993: 279). Both approaches have the zero-pronoun as the formal device signaling the highest degree of accessibility in common (cf. Ariel (2014: 73), (2001: 31); Gundel et al. (1993: 284)). The fact that zero-pronoun is the most frequent form of the SP's subject therefore tells us something specific about its status: Among the various discourse entities which might be part of the discourse at a certain point in time the SP's subject is quite often the discourse entity with the highest degree of accessibility. This status goes beyond the observation that S frequently has a coreferent antecedent in the previous context.

3.3.2 Semantic properties of correspondents

Following our research goals, we are not only interested in how often SP and S have correspondents in the context, but also in the semantic and formal properties of these correspondents. While the semantic properties are determined by the semantic relation between the correspondent and SP or S, the formal properties are the morphological or syntactic form of the correspondent. In the analysis of these qualitative aspects, not all correspondents are taken into consideration, but only the two closest to SP and S in the preceding and in the following context. Thus, for a given example, at most four correspondents of SP and four correspondents of S have been considered. The semantic properties will be presented in the following, the formal properties in Section 3.3.3.

We begin with the semantic properties of the correspondents of SP. In the case of SP, three semantic relations are distinguished: *same*, *similar* and *evoked* (cf. Section 3.2 for a description). Table 5 and Figure 2 show the frequency of these relations among the correspondents of SP. As we can see, there is a clear dominance of the relation *evoked* with 74.73% of the correspondents. The remaining relations *similar* (15.38%) and *same* (9.89%) follow with considerable distance. The rank in frequency of the three relations thus negatively correlates with the semantic similarity between the correspondent and the SP. The closer the semantic relation, the less frequently it is attested (cf. (19)).

(19) a. similarity: same > similar > evoked

b. frequency: evoked > similar > same

	abs.	%
same	9	9.89
similar	14	15.38
evoked	68	74.73
total	91	100.00

Table 5: Frequency of relation types between SP and correspondents

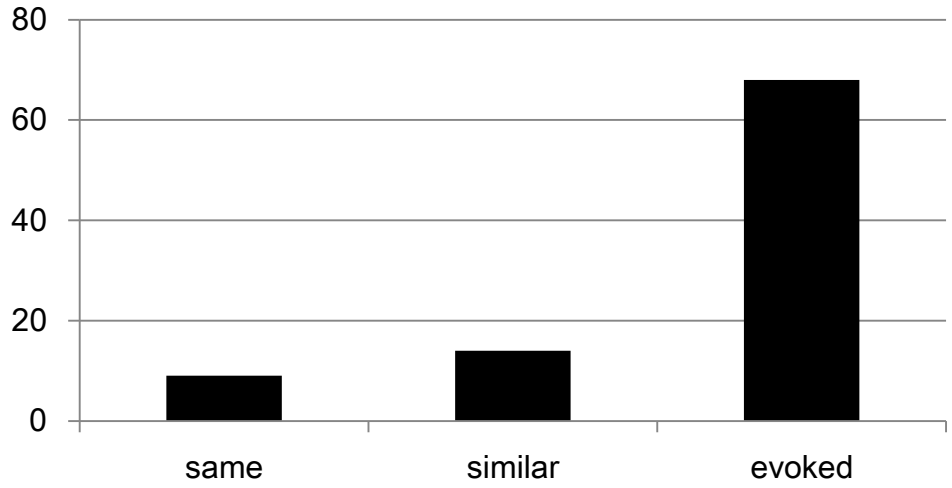


Figure 2: Frequency of relation types between SP and correspondents (absolute)

As concerns the distribution of correspondents between the three semantic relation types in pre-text and post-text, we see differences in the case of the relation *same* and the relation *evoked*. *Same* is more frequent in the post-text than in the pre-text while for *evoked* it's just the opposite (cf. Table 6 and Figure 3 below).

	pre-text	post-text	total
same	2	7	9
similar	7	7	14
evoked	41	27	68
total	50	41	91

Table 6: Relation types in pre-text and post-text

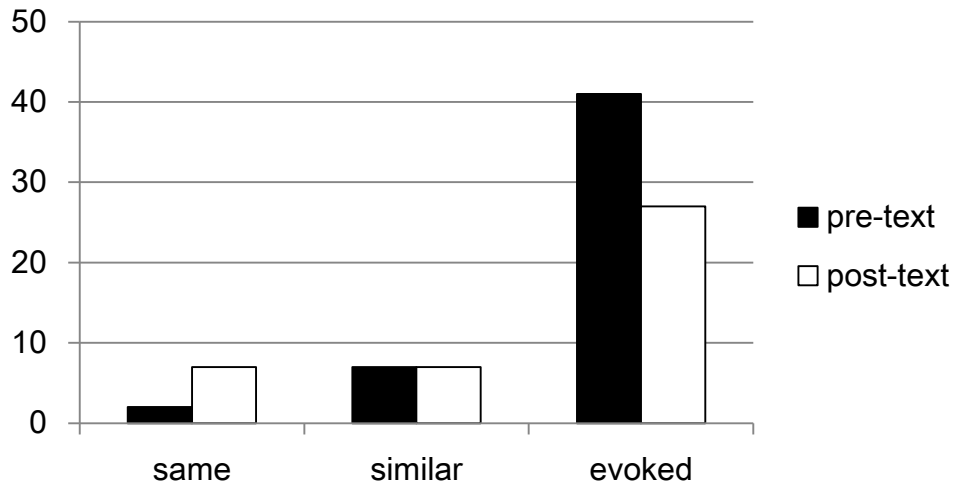


Figure 3: Relation types in pre-text and post-text

Below, examples for the three semantic relations between correspondents and SP are given. In (20a.), an example for the semantic relation *same*, the secondary predicate *sorprendido* 'surprised' has the correspondent *perplejo* 'perplexed' in the post-text. The utterance *¡perpleja!* must be interpreted as an elliptical copula construction such as *Rosa / ella está perpleja* 'Rosa / she is perplexed'. Another example for the relation *same* is (20b.), where the secondary predicate *amontonado* 'piled up' has the correspondent *como ratas en una madriguera* 'like rats in a den' in the post-text.

(20) same

a. Rosa me mira sorprendida$_{SP}$, ¿comprendes?, ¡perpleja!$_{CORR}$

'Rosa looks at me surprised: "Do you understand?" She is perplexed!'

(Azúa, Diario: 185; mod. StH)

b. […] vivían amontonados$_{SP}$ los unos con los otros [...] tanto campo vacío y nosotros aquí, [como ratas en una madriguera]$_{CORR}$.

'They lived piled up all together […] so much free space and we live here, like rats in a den'

(Mendoza, Prodigios: 167; mod. StH)

Examples for the relation *similar* are given in (21). The relation of similarity holds between the correspondent *dormido* 'asleep' and the secondary predicate *adormilado* 'sleepy', and between the correspondent *arrullado* 'lulled' and the secondary predicate *melancólico* 'melancholic'.

(21) similar

a. […] cabeceo adormilado$_{SP}$, y si no caigo dormido$_{CORR}$ es por la náusea de la resaca.

'I nodded off sleepy, and I only didn't fall asleep because of the nausea of the hangover'

(Azúa, Diario: 150; mod. StH)

b. Podría pasar en este balcón la noche entera, pensaba entonces melancólico$_{SP}$ […]; pasarme aquí el verano entero, arrullado$_{CORR}$ por los sonidos de esta ciudad anónima. ' "I could spend the whole night on this balcony", did he think melancholic […]; "I could spend here the whole summer lulled by the sounds of this anonymous city." '

(Mendoza, Ciudad: 133; mod. StH)

Finally, examples for the relation *evoked* can be found in (22). In the first example, the VP *decir algo en voz risueña* 'say something smilingly' evokes the secondary predicate *encantado* 'delighted'. In the second example, the offer of a better paid and safer job evokes the secondary predicate *contento* 'satisfied, pleased'.

(22) evoked

a. La que estaba de frente dijo en voz muy alta y risueña$_{CORR}$: 'Bueno bueno bueno, es que me asquea.' Lo dijo encantada$_{SP}$, estuvo a punto de palmearse los muslos mulatos.

'The one that was face to face said in a loud and smiling voice: "Well, well, well, that's to say he disgusts me." She said it delighted, she was about to tap her mullato thighs.'

(Marías, Corazón: 137; mod. StH).

b. (Context: guards are offered safer and better paid jobs)$_{CORR}$

Estos guardas no eran más que tres, aceptaron contentos$_{SP}$

'These guards were not more than three and they accepted pleased'

(Sender, Réquiem: 72; mod. StH)

As for the information status of the secondary predicate, we examined how often the secondary predicate has correspondents or not (in order to see whether the secondary predicate is new or non-new) and looked at the semantic relation between SP and the correspondent (in order to see which subtype of non-new information holds).[8] Table 7 shows the frequency of the four information statuses.

	abs.	%
same	2	4.00
similar	5	10.00
evoked	26	52.00
new	17	34.00
total	50	100.00

Table 7: Information status of SP

The quantitative relation between new and non-new information has already been stated in Section 3.3.1: The secondary predicate expresses non-new information (66%) more often than new information (34%). Within the group of non-new SPs, we see that evoked SPs clearly dominate over SPs with the status *same* and *similar*. Although secondary predicates frequently express information that is not completely new to the discourse, this non-new information has only a rather weak link to the preceding context. In most of the cases it is merely evoked by some element of the preceding context, which is, in a certain sense, the newest non-new information possible.

To conclude the description of the semantic properties of correspondents of secondary predicates, it needs to be stressed that in the vast majority of cases, namely 96.70%, the

[8] If a SP had more than one correspondent we counted the one with a closer semantic relation to the SP (e.g., *same* in the case of *same* and *similar*).

predication expressed by the correspondent of SP is over the same participant as the predication expressed by the secondary predicate itself, as illustrated in (23a.). Only in a minor number of cases, namely 3.30%, does the correspondent of SP predicate over a participant other than S, as illustrated in (23b.). These three correspondents of SP are located in the pre-text, two have the semantic property *evoked*, one has the property *similar*. One factor that favors the frequent identity between the predication goal of SP and its correspondents is the high number of correspondents of S (i.e., the SP's goal of predication) in the context.

(23) a. John was sad_{CORR} and he went home sad_{SP}.

b. Paul left sad_{CORR} and John arrived sad_{SP}.

Turning to the semantic properties of the correspondents of S, we distinguish between three relations: *same*, *part-of-group*, *evoked* (cf. Section 3.2 for a description). Table 8 and Figure 4 show the frequency of these relations among the correspondents of S. As we can see, there is a clear dominance of the relation *same*, i.e., coreference, with 96.05% of the correspondents. The relation *part-of-group* (3.95%) follows with considerable distance and the third relation *evoked* is not attested at all. The rank in frequency of the three relations correlates with the similarity (or referential identity) between the correspondent and S. The more similar, the more frequent the correspondents are attested (cf. (24)).

(24) a. similarity: same > part-of-group > evoked

b. frequency: same > part-of-group > evoked

	abs.	%
same	170	96.05
part-of-group	7	3.95
evoked	0	0.00
total	177	100.00

Table 8: Frequency of relation types between S and correspondents

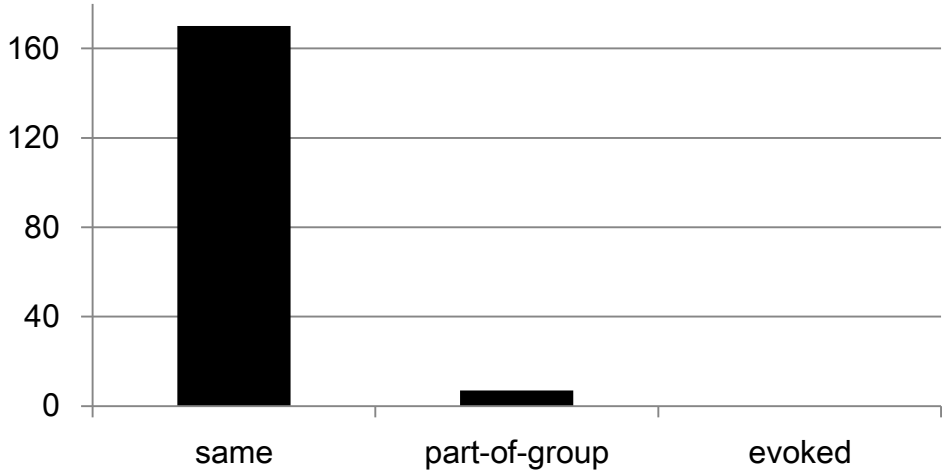


Figure 4: Frequency of relation types between S and correspondents (absolute)

As concerns the distribution of the three semantic relation types in pre-text and post-text, hardly any differences exist between the two positions (cf. Table 9 and Figure 5).

	pre-text	post-text	total
same	86	84	170
part-of-group	2	5	7
evoked	0	0	0
total	88	89	177

Table 9: Frequency of relation types between S and correspondents in pre-text and post-text (absolute)

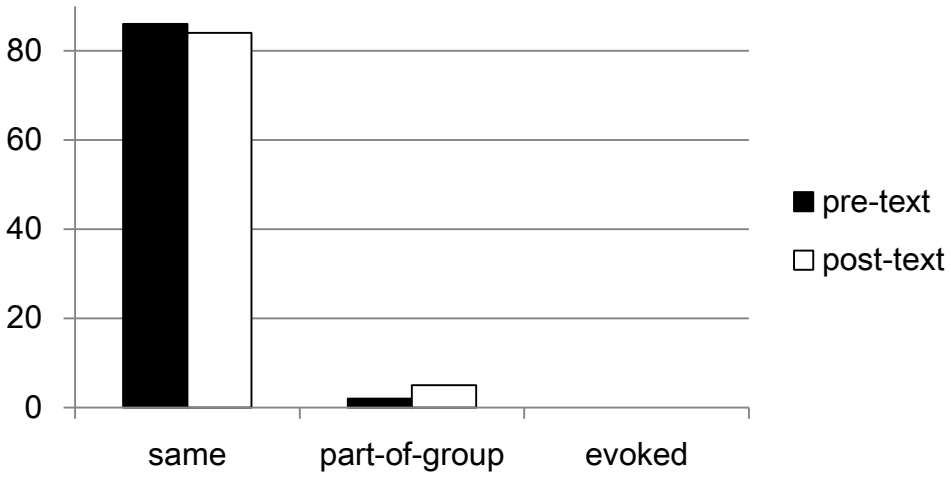


Figure 5: Frequency of relation types between S and correspondents in pre-text and post-text (absolute)

Examples for the two attested relations are given below. The most frequent relation, where the correspondent is referentially identical to S, is illustrated in (25): the secondary predicate *alelado* 'stupefied' predicates over the zero-subject of the sentence and this subject has a referentially identical correspondent in the preceding context, namely *los camareros* 'the waiters'.

(25) Los camareros$_{CORR}$ no osan intervenir; __$_{S}$ miran alelados como madrileños ante un accidente de circulación.
'The waiters didn't dare to intervene; they watched stupefied like *Madrileños* facing a traffic accident.'
(Azúa, Diario: 140; mod. StH)

In (26), the secondary predicate *aturdido* 'shocked' predicates over the zero-subject of the sentence and this subject has a correspondent in the following context, namely the zero-subject in the sentence headed by the verb *salimos* '(we) left'. In this case, the correspondent of S is not referentially identical to S, the correspondent refers to a group that S is part of (1st person plural vs. 1st person singular pronoun).

(26) La __$_{S}$ seguí aturdida por el sueño, sin saber aún qué sucedía; pero __$_{CORR}$ salimos al pasillo […]
'I followed her shocked, without knowing what was going on; but left towards the hallway'
(Montero, Bella: 108; mod. StH)

As for the information status of the S, we looked at how often S has correspondents in the pre-text or not (in order to see whether S is new or non-new) and at the semantic relation between S and the correspondent (in order to see which subtype of non-new information holds).[9] Table 10 shows the frequency of the four information statuses.

[9] If a S had more than one correspondent we proceeded as in the case of SP: we counted the one with a closer semantic relation to S (e.g. *same* in the case of *same* and *part-of-group*).

	abs.	%
same	46	92.00
part-of-group	0	0.00
evoked	0	0.00
new	4	8.00
total	50	100.00

Table 10: Information status of S

The quantitative relation between new and non-new information has already been stated in Section 3.3.1: S nearly always expresses non-new information (92%) rather than new information (8%). Table 10 shows that the whole group of non-new Ss consists of *same* Ss. Hence, S typically has a coreferent antecedent in the preceding context and is thus discourse given information. This result strongly corresponds to Prince's (1981: 243) findings on the information status of nominal constituents in English. She distinguishes subjects from non-subjects and finds that 93.40% of the subjects are given (while only 48.80% of the non-subjects are given). Further, the fact that S is in most cases a zero-pronoun shows that S is a highly accessible referent (in terms of Ariel, 2001, 2014; cf. Section 3.3.1).

Let us now compare SP and S with respect to the semantic relation with their correspondents. The above results clearly show that Ss and their correspondents have a much more direct relation than SPs and their correspondents; while in the case of S, the most direct relation possible clearly dominates (namely *same,* i.e., referential identity), in the case of SP, the most indirect relation possible dominates (namely *evoked*). Recall from Section 3.3.1 that S has many more correspondents in the context than SP and that S is quantitatively more anchored in the discourse than SP. The analysis of the semantic properties of the correspondents shows that the same holds on the semantic level. S is semantically more anchored in the discourse than SP because its semantic relation to its correspondents is much tighter than in the case of SP.

	SP	S
same	9.89	96.05
similar / part-of-group	15.38	3.95
evoked	74.73	0.00
total	100.00	100.00

Table 11: Percentage of relation types for SP and S

As concerns information status, we observe that both SP and S express non-new information in the majority of cases (cf. Table 12).[10] This is, however, where the similarities end. Crucially, the correspondents of S are much more similar to S than the correspondents of SP are to SP. While S typically has a coreferent correspondent in the pre-text, SP predominantly has a correspondent in the pre-text which merely evokes the state denoted by the secondary predicate.

		SP	S
non-new	same	4.00	92.00
	similar / part-of-group	10.00	0.00
	evoked	52.00	0.00
new		34.00	8.00
	total	100.00	100.00

Table 12: Percentage of information status for SP and S

3.3.3 Formal properties of correspondents

Although the focus of our analysis lies on the semantic properties of the correspondents, i.e., their semantic relation to SP or S, we shall consider also the formal properties of the correspondents. The formal properties are relevant in order to verify whether the degree of semantic relatedness between SP / S on the one hand and their correspondents on the other hand has a correlate on the level of form. We have seen in the previous sections that Ss maintain a much closer semantic relationship with their correspondents than SPs. The question is whether this difference between S and SP also holds on the level of form.

As concerns the form of the correspondents of SP, we need to bear in mind that SPs express states and that states are typically expressed by adjectives and in constructions such as secondary predication (e.g., *He went home sad)*, copulas (e.g., *He is sad*), NP internal adjectives (e.g., *the sad man*) or nouns describing states (e.g., *sadness*). Thus, one might expect that the correspondents of SPs also take one of these forms. However, as Table 13 shows, the correspondents of SPs in the corpus rarely have the typical form of state descriptions. Instead, most correspondents have a form other than these typical adjectival constructions.

[10] The chi-square test shows that the difference between SP and S with respect to *non-new* vs. *new* is statistically significant ($p<0.005$).

	abs.	%
SP (lexical)	7	7.69
SP (pronominal)	0	0.00
copula (lexical)	1	1.10
copula (pronominal)	0	0.00
NP internal adjective	0	0.00
state noun	4	4.40
other	79	86.81
total	91	100.00

Table 13: Form of SP-correspondents

It might seem odd to use an inventory of categories where the resulting categorization lumps together the vast majority of cases in a trash can category such as *other*. The point to be made here is, however, that the correspondents of the secondary predicates do not take any of the forms that state descriptions usually take. These other forms can be VPs such as in (27) or utterances as in (28).

(27) correspondent = VP

a. sigue bebiendo Calisay → torcido (Azúa, Diario: 171)
'keep drinking Calisay' 'crooked'

b. tener pánico → despavorido (Ferrero, Opium: 94)
'to panic' 'terrified'

(28) correspondent = utterance

a. "¡Ah!" → sorprendido (Azúa, Diario: 185)
'ah' 'surprised'

b. "¿He oido bien?" → perplejo (Ferrero, Opium: 73)
'Did hear right?' 'perplexed'

c. "¡No es cierto!" → furioso (Ferrero, Opium: 81)[11]
'It not certain!' 'furious'

The formal properties of the correspondents, and especially the minor role of adjectival constructions, need to be interpreted against the foil of their semantic properties. Recall that the dominant semantic relation between the correspondent and the secondary predicate is *evoked*. To evoke a state of a discourse referent it is not necessary to directly express the

[11] *"¡No es cierto!"*, which combines both exclamation and protest, can be considered a correspondent of the SP *furioso* because both express the speaker's discontent.

state by means of an adjectival construction. Indeed, it seems to be the case that adjectival constructions are not ideal to evoke a state because they directly express a state (which is the same as the one denoted by the secondary predicate, which is similar to it or which is different from it; the last case being irrelevant for our study). The indirect semantic relation between SPs and their correspondents discussed in Section 3.3.2 thus has a correlate on the formal level: correspondents of SPs do not have the form of adjectival constructions.

Turning to the form of the correspondents of S, we see a completely different picture, and not only with respect to the inventory of categories, but more importantly with respect to the formal similarity between S and its correspondents. Recall from Section 3.3.1 that Ss typically have the form of a zero-pronoun (54%). Table 15 shows that zero-pronouns are also the most frequent category for the correspondents of S (with 44.07%), and that further, zero- and overt pronouns taken together amount to 70.62% of all correspondents of S. Hence, the close semantic relation between S and its correspondents has a correlate on the level of form in that zero-pronouns are the predominant form of both S and their correspondents.

	abs.	%
zero-pronoun	78	44.07
overt pronoun	47	26.55
proper name	23	12.99
definite or quantized NP	27	15.25
indefinite NP	1	0.56
bare NP	1	0.56
total	177	100.00

Table 14: Form of correspondents of S

3.3.4 Interim conclusion: the discourse profile of secondary predicates

We have seen that both SP and S are anchored in the discourse in that they have correspondents in the pre-text and the post-text. However, we have found considerable differences between SP and S as concerns the frequency of these correspondents and their formal and semantic properties. Correspondents of S are eight times more frequent than correspondents of SP. Taking semantic and formal properties together the typical correspondent of SP belongs to the semantic type *evoked* and the formal type *other* (i.e., non-adjectival) (66 of 91 cases). The typical correspondent of S belongs to the semantic type *same* and the formal type *pronoun* (97 of 177 cases). We may thus conclude that SPs do have correspondents in the discourse, but these correspondents are relatively infrequent

(compared to that of S) and the correspondents have only a rather lose semantic relation to the secondary predicate.

Not surprisingly, we also find differences between SP and S as concerns information status. Although both SP and S express non-new information in the majority of cases, they differ in that the pre-text correspondents of S are much more similar to S than the correspondents of SP are to SP. While S typically has a correspondent in the pre-text which is coreferent, SP typically has a correspondent in the pre-text which merely evokes the state denoted by SP. The fact that the most frequent form of S itself is a zero-pronoun further highlights this contrast between SP and S. Zero-pronouns are expressions which are used for highly accessible discourse entities (i.e. they are at the center of attention and very salient in memory etc. (cf. Ariel, 2001, 2014; Gundel et al., 1993)).

The above description results in a bundle of prototypical information structural features of secondary predicates (i.e. involving the secondary predicate itself and its subject). Following Ariel (2002), we can refer to this bundle of features as the *prototypical discourse profile* of secondary predicates. Such discourse profiles are based on recurring patterns in the use of a given linguistic expression, but do not correspond to the discoursal appropriateness conditions imposed on the expression (for the latter aspect Ariel (2002: 15) uses the term *discourse function*). For example, the information structural asymmetry between SP and S in terms of information status (SP: rather new vs. S: given and highly accessible) constitutes a recurring pattern in our data. Nonetheless this asymmetry is not an appropriateness condition on the use of secondary predicates since one can easily construe examples with a given SP and a new S (cf. also Section 3.4.2).

3.4 Discussion

3.4.1 On the newness of secondary predicates

The finding that the majority of SPs (66%) in our corpus is somehow anchored in the preceding context, and thus expresses non-new information, seems to conflict with the existing literature on the information structural properties of SPs. Recall from Section 1 that secondary predicates are described as constituents which typically express new and focal information (cf. Guemann, 1990; Porroche Ballesteros, 1990; Rodríguez Espiñeira, 1992). Taking this discrepancy as a starting point, we will reconsider in this section the significance of our empirical finding.

First, we need to consider what it means in the case of SP to express non-new information. Based on the relation between the correspondent and the secondary predicate, we have distinguished three types of non-new information, namely *same*, *similar* and *evoked*. Among these three types we have found considerable differences in frequency: evoked (78.79%) >

similar (15.15%) > same (6.06%) (cf. Table 5, Section 3.3.2). The frequency of these statuses negatively correlates with the semantic similarity between correspondent and SP. The vast majority of non-new uses of SP are cases where the secondary predicate is evoked by some element from the preceding context. Within the cases of *evoked,* further distinctions can be made which concern the relation between the secondary predicate and the correspondent. Let us consider the relation between the pre-text correspondents and the states in (29).

(29)	*pre-text correspondent*	*state*	
	a. to swim	wet	entailed state
	b. to cry	sad	probable state

While the pre-text correspondent *to swim* entails the state *wet*, the correspondent *to cry* only suggests the state *sad*. Hence we can distinguish between evoked states which are entailed and evoked states which are only probable.[12] The relation between correspondent and the state denoted by the secondary predicate is undoubtedly closer in the case of entailment.
As concerns our own data, practically all evoked secondary predicates are probable and not entailed states. The only case which might be interpreted as an entailed state is shown in (30). The verb *cabecear* 'to nod off' seems to entail the state *adormilado* 'sleepy'.

(30) cabecear 'to nod off' + adormilado 'sleepy'

[…] cabeceo adormilado, y si no caigo dormido es por la náusea de la resaca.

'I nodded off sleepy, and I only didn't fall asleep because of the nausea of the hangover'

(Azúa, Diario: 150)

In all other cases of *evoked*, the correspondent in the pre-text only makes the state probable, but does not entail it. Examples for this relation are given in (31).

[12] This distinction has a correlate in the nominal domain. Clark (1975: 171) distinguishes various bridging relations, amongst others between necessary and probable parts (cf. also Irmer's (2011: §6) comprehensive overview on types of bridging phenomena). While the definite subject-NP in (ia.) is a necessary part of *the room*, the subject-NP in (ib.) is only a probable part of *the room.*

(i) a. I looked into the room. The ceiling was very high.

b. I walked into the room. The windows looked out to the bay.

(31) a. darse de bruces contra la pilastra ('fall on the face against the pilaster') + dolorido 'in pain'

[…] se dio de bruces contra la pilastra$_{CORR}$. ¡Menstruación de la Virgen!, gritó dolorido$_{SP}$ […]

'he fell on the face against the pilaster. "Menstruation of the virgin!", did he scream in pain.'

(Azúa, Diario: 125; mod. StH)

b. (prospect of being executed) + enloquecido 'crazed'

(Context: subject of SP, together with other persons, is lined up for execution)

Mosén Millán, usted me conoce - gritaba enloquecido$_{SP}$.

'Mosén Millán, you know me – he screamed crazed'

(Sender, Réquiem: 102; mod. StH)

Hence, if we say that among all non-new information statuses the one with the weakest relation in the pre-text dominates, then this statement is true in a double sense. First, *evoked* dominates among the three non-new relations, and second, among the instances of *evoked*, cases where the evoked state is probable clearly dominate over cases where the evoked state is entailed.

The conclusion from the above observations is that our findings require a cautious interpretation. Although secondary predicates have a relation to the pre-text in the majority of cases, this relation is rather weak and indirect. Moreover, there is not a single example in our corpus where the pre-text correspondent has the same lexical basis as the secondary predicate (e.g., a copula construction formed with the same adjective as the secondary predicate).

Finally, we shall apply the following two binary distinctions to our data: new vs. non-new and given vs. non-given (the latter roughly corresponds to Roberts' (2003) distinction between weak vs. strong familiarity in the domain of definite descriptions). If we apply the second distinction, our findings on the dominance of non-new uses of secondary predicates are further relativized. Table 15 shows the information status of SP and of S using the two binary distinctions.[13]

[13] The chi-square test shows that the differences between SP and S with respect to (i) *non-new* vs. *new* and (ii) *given* vs. *non-given* are statistically significant ($p<0.005$).

	Newness	SP	S	Givenness	SP	S
same				given	4.00%	92.00%
similar / part-of-group	non-new	66.00%	92.00%			
evoked				non-given	96.00%	8.00%
new	new	34.00%	8.00%			
total		100.00%	100.00%		100.00%	100.00%

Table 15: Newness vs. givenness

For SP, the table shows that most secondary predicates are non-new (66%) but only a very small share of the secondary predicates are also given (4%). The results for the binary distinction between given and non-given highlight the weak anchorage of secondary predicates in the preceding context. In contrast to this, we observe for S that both the non-new and the given uses dominate (92%).

3.4.2 On the lack of given secondary predicates

Secondary predicates hardly ever express given information (cf. Table 15). The aim of this section is a better understanding of this finding on the information status of SP. To this end, we will change our perspective and see what it would mean if secondary predicates would frequently express given information. The consideration of this alternative scenario will help us in a certain way to understand the frequently attested one, i.e., the lack of given secondary predicates.

Let us assume three discourses, which differ with respect to the predicate–argument combinations they contain. We assume two predicates, P_1 and P_2, and two arguments, x and y. Discourse A consists of two different predications over one and the same discourse entity (cf. (32a.)). In discourse B, the same predicate occurs twice and it predicates over the same discourse entity (cf. (36b.)). In discourse C, the same predicate occurs twice, but it predicates over two different discourse entities (cf. (32c.)).

(32) a. discourse A: $P_1(x)$, $P_2(x)$
b. discourse B: $P_1(x)$, $P_1(x)$
c. discourse C: $P_1(x)$, $P_1(y)$

As concerns the information status of the second predicate in the discourse, B and C differ from A. In A the second predicate (P_2) is new while in B and C the second predicate (the second mention of P_1) is given. Discourse B and C thus describe the alternatives to the dominant case in our corpus. The alternatives are (i) to predicate the same over the same

discourse entity or (ii) to predicate the same over different discourse entities. The following examples show that one can easily come up with discourses of type B and C.

(33) $P_1(x)$, $P_1(x)$

a. John$_x$ came home drunk$_{P1}$ yesterday and he$_x$ will come home drunk$_{P1}$ tomorrow.

b. John$_x$ arrived drunk$_{P1}$ and he$_x$ left drunk$_{P1}$.

c. A: John$_x$ arrived drunk$_{P1}$, right?

B: Yes, he$_x$ arrived drunk$_{P1}$.

(34) $P_1(x)$, $P_1(y)$

a. John$_x$ came home drunk$_{P1}$ yesterday and Mary$_y$ will come home drunk$_{P1}$ tomorrow.

b. John$_x$ arrived drunk$_{P1}$ and Mary$_y$ left drunk$_{P1}$.

c. A: Mary$_x$ arrived drunk$_{P1}$, right?

B: No, John$_y$ arrived drunk$_{P1}$.

The low frequency of given secondary predicates thus may not result from the fact that such discourses are hard to construct.[14] Instead, we must assume that such discourses are relatively infrequent because, simply speaking, we have more important things to say. There must be other things which have higher relevance than the repetition of a predicate. In this respect, we must stress the role of subjective informational value. The low frequency is not a result of a low frequency of such scenarios in the extra-linguistic world. For the type *$P_1(x)$, $P_1(x)$* it is sufficient to observe a person for some minutes in order to get to a sequence such as (35a.); the same holds for the type *$P_1(x)$, $P_1(y)$*, once we focus on several persons (cf. (35b.)).

(35) a. $P_1(x)$, $P_1(x)$

John$_x$ left the house well-dressed$_{P1}$, he$_x$ walked down the road well-dressed$_{P1}$, and he$_x$ even jumped into the pool well-dressed$_{P1}$.

b. $P_1(x)$, $P_1(y)$

John$_x$ left the house well-dressed$_{P1}$, [his neighbor]$_y$ left the house well-dressed$_{P1}$, and today even Paul$_z$ left the house well-dressed$_{P1}$.

The following examples, which stem from CREA and not from the corpus of our own study, show that the repetition of a predicate occurs outside of constructed examples and in real language data.

[14] Dimroth et al. (2010) discuss a variety of formal means which are used in Romance and Germanic languages to maintain a predication in a discourse.

(36) $P_1(x)$, $P_1(x)$

Desnudo$_{P1}$ vine al mundo y desnudo$_{P1}$ me hundo ahora en su naufragio.

'Naked did I come into this world and naked will I throw myself into the sinking'

(Nieva, 1990; CREA; mod. StH)

(37) $P_1(x)$, $P_1(y)$

Triste$_{P1}$ caminó Pepita$_x$ hacia la puerta del penal. Triste$_{P1}$ caminó [el abuelo de Elvira]$_y$.

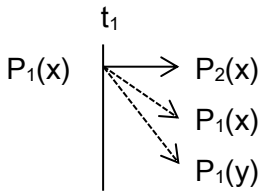
Triste$_{P1}$ caminó [el marido de Reme, pobre Benjamín]$_z$.

'Sad walked Pepita towards the door of the prison. Sad walked Elvira's grandfather. Sad walked the Reme's husband, poor Benjamín.'

(Chacón, 2002; CREA; mod. StH)

It follows that the low frequency of repeated predicates and predications does not result from their infrequency in the extra-linguistic world. Instead, they result from what the speaker chooses to talk about. At a certain point in discourse a speaker thus has several options of how to continue the discourse, all of which assure a communicative progression. Continuations which involve repeated predicates and predications are possible, but infrequent compared to continuations in which a new predicate predicates over a given discourse entity (cf. (38).

(38)



Klein and von Stutterheim (2002: §3.1) argue that explicit and implicit questions are at the basis of these choices. In this respect, the results from our study strongly suggest that questions about a participant, such as "What happened next to *x*?", are more important to speakers than questions about a predicate, such as "Who else did/was *predicate*?"

4 Summary and conclusions

The goal of this paper was to analyze the discourse profile, i.e. the information status and discourse anchorage of Spanish secondary predicates. By defining this goal, we took a clear stance with respect to the question of whether information status is a relevant descriptive category for non-nominal constituents such as secondary predicates (for the opposite view cf. Prince, 1981; Lambrecht, 1994; Nissim et al., 2004; Dipper et al., 2007).

On the basis of 50 Spanish examples with secondary predicates (plus context), we have shown that secondary predicates are anchored both in the preceding and the following context, but to a much lesser extent than their subjects. While authors such as Guemann (1990), Porroche Ballesteros (1990) and Rodríguez Espiñeira (1992) suggest that secondary

predicates have a strong affinity to focus (and thus to expressing new information), our own results show a more differentiated picture. The most frequent information status of secondary predicates is neither *new* nor *given*. Crucially, the most frequent case is that the state denoted by the secondary predicate is evoked by some element of the preceding context. Hence, an intermediate information status between *new* and *given* is the dominant one. The lack of given secondary predicates is not surprising. As concerns new secondary predicates, however, the existing literature would predict this to be the dominant type. Taking into account also the information status of the subject, the most frequent configuration is that an evoked secondary predicate predicates over a given and highly accessible subject.

The dominance of evoked secondary predicates, i.e., of an intermediate status between *new* and *given*, shows that the analysis of the information status of non-nominal constituents requires a more elaborate inventory than a binary given–new distinction (just as in the case of nominal constituents). One important question for further research is, therefore, whether intermediate information statuses of secondary predicates, and non-nominal constituents more generally, have formal reflexes similar to those we find in the nominal domain (e.g. morphosyntactic differences, syntactic position, prosodic realization).

References

Ariel, Mira, 1985. The discourse functions of given information. Theoretical Linguistics 12 (2-3), 99–114.

Ariel, Mira, 2001. Accessibility theory: an overview, In: Sanders, Ted, Schilperoord, Joost, Spooren, Wilbert (Eds.), Text Representation. Linguistic and psycholinguistic aspects. John Benjamins Publishing Company, Amsterdam, pp. 29–87.

Ariel, Mira, 2002. The possessive NP construction: Discourse function and discourse profile. Berkeley Linguistics Society 28, 15–26.

Ariel, Mira, 2014 [1990]. Accessing Noun-Phrase Antecedents. Routledge, New York.

Arnold, Jennifer E., Wasow, Thomas, Losongco, Anthony, Ginstrom, Ryan, 2000. Heaviness vs. Newness: The Effects of Structural Complexity and Discourse Status on Constituent Ordering. Language 76, 28-55.

Asher, Nicholas, 1993. Reference to Abstract Objects in Discourse. Kluwer, Dordrecht.

Baumann, Stefan, 2006. The Intonation of Givenness: Evidence from German. de Gruyter, Berlin.

Chafe, Wallace, 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View, In: Li, Charles N. (Ed.), Subject and topic. Acad. Press, New York, NY, pp. 25–55.

Clark, Herbert H., 1975. Bridging, In: R. C. Schank & B. L. Nash-Webber (Ed.), Theoretical issues in natural language processing. Association for Computing Machinery, New York, pp. 169–174.
CREA = Real Academia Española. Banco de datos (CREA). Corpus de referencia del español actual. <http://www.rae.es>.
Demonte, Violeta, 1991. Detrás de la palabra: Estudios de gramática del español. Alianza Ed., Madrid.
Demonte, Violeta, Masullo, Pascual J., 1999. La predicación: los complementos predicativos, In: Bosque, Ignacio, Demonte, Violeta (Eds.), Gramática descriptiva de la lengua española: Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales. Espasa-Calpe, Madrid, pp. 2461–2523.
Dimroth, Christine, Andorno, Cecilia, Benazzo, Sandra, Verhagen, Josje, 2010. Given claims about new topics. How Romance and Germanic speakers link changed and maintained information in narrative discourse. Journal of Pragmatics 42 (12), 3328–3344.
Dipper, S., Götze, M., Skopeteas, S. (Eds.), 2007. Information Structure in Cross-Linguistic Corpora: Annotation Guidelines for Phonology, Morphology, Syntax, Semantics and Information Structure. Universitätsverlag Potsdam, Potsdam.
Guemann, Steven A., 1990. Secondary Predication in English and Spanish. Phd Thesis, Berkeley.
Gumiel Molina, Silvia, 2005. Los complementos predicativos. Arco Libros, Madrid.
Gumiel Molina, Silvia, 2008. Estructura argumental y predicación secundaria. Univ. de Alcalá, Servicio de Publ., Alcalá de Henares.
Gundel, Jeanette K., Hedberg, Nancy, Ron Zacharski, 1993. Cognitive Status and the Form of Referring Expressions in Discourse. Language 69 (2), 274–307.
Haviland, Susan E., Clark, Herbert H., 1974. What's new? Acquiring New information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior 13 (5), 512–521.
Hegarty, Michael, 2003. Semantic types of abstract entities. Lingua 113 (9), 891–927.
Hernanz Carbó, María L., 1988. En torno a la sintaxis y la semántica de los complementos predicativos en español. Estudi General 8, 7–31.
Himmelmann, Nikolaus P., Schultze-Berndt, Eva, 2005. Issues in the syntax and semantics of participant-oriented adjuncts: an introduction, In: Himmelmann, Nikolaus P., Schultze-Berndt, Eva (Eds.), Secondary predication and adverbial modification. The typology of depictives. Oxford Univ. Press, Oxford, pp. 1–67.
Horn, Laurence R., 1991. Given as new: When redundant affirmation isn't. Journal of Pragmatics 15 (4), 313–336.

Hummel, Martin, 2000. Adverbale und adverbialisierte Adjektive im Spanischen: Konstruktionen des Typs "Los niños duermen tranquilos" und "María corre rápido". Narr, Tübingen.
Irimia, Monica-Alexandrina, 2005. Types of secondary predication. Toronto Working Papers in Linguistics 25, 20–29.
Irmer, Matthias, 2011. Bridging inferences: Constraining and resolving underspecification in discourse interpretation. Walter de Gruyter, Berlin.
Klein, Wolfgang, Stutterheim, Christine von, 2002. Quaestio und I-perspectivation, In: Graumann, Carl F., Kallmeyer, Werner (Eds.), Perspective and Perspectivation in Discourse. John Benjamins Publishing Company, Amsterdam, pp. 59-88.
Krifka, Manfred, 2007. Basic Notions of Information Structure, In: Féry, Caroline, Fanselow, Gisbert, Krifka, Manfred (Eds.), The Notions of Information Structure. Universitätsverlag Potsdam, Potsdam, pp. 13–55.
Lambrecht, Knud, 1994. Information structure and sentence form: Topic, focus, and the mental representations of discourse referents, transf. to digital print. ed. Cambridge Univ. Press, Cambridge.
Maienborn, Claudia, 2005. On the limits of the Davidsonian approach: The case of copula sentences. Theoretical Linguistics 31 (3), 275–316.
Nissim, Malvina, Dingare, Shipra, Carletta, Jean, Steedman, Mark, 2004. An annotation scheme for information status in dialogue, in: Proceedings of the 4th Conference on Language Resources and Evaluation (LREC2004), Lisbon, Portugal, 2004, pp. 1023–1026.
Porroche Ballesteros, Margarita, 1990. Aspectos de la atribución en español. Pórtico, Zaragoza.
Prince, Ellen F., 1981. Toward a Taxonomy of Given-New Information, In: Cole, Peter (Ed.), Radical Pragmatics. Academic Press, New York, pp. 223–255.
Riester, Arndt, 2008. The Components of Focus and their Use in Annotating Information Structure. Universität Stuttgart, Stuttgart.
Roberts, Craige, 2003. Uniqueness in Definite Noun Phrases. Linguistics and Philosophy 26 (3), 287-350.
Rodríguez Espiñeira, María J., 1992. Sobre la codificación informativa de las cláusulas con predicativo en español. Revue Romane 27 (1), 30–54.
Rothstein, Susan, 2006. Secondary Predication, In: Everaert, Martin, van Riemsdijk, Henk (Eds.), The Blackwell companion to syntax. Volume IV. Blackwell, Malden, Mass., pp. 209–233.
Rothstein, Susan, 2011. Secondary predicates, In: Heusinger Klaus von, Maienborn Claudia, Portner Paul (Eds.), Semantics. Volume 2. de Gruyter, Berlin, Boston, pp. 1442–1462.

Schultze-Berndt, Eva, Himmelmann, Nikolaus P., 2004. Depictive secondary predicates in crosslinguistic perspective. Linguistic Typology 8, 59–131.
Suñer Gratacòs, Avel·lina, 1990. La predicación secundaria en español, Barcelona.
Winkler, Susanne, 1997. Focus and Secondary Predication. Mouton De Gruyter, Berlin.

Novels

Azúa, Diario = Azúa, Félix de, 1991 [1987]. Diario de un hombre humillado. Edit. Anagrama, Barcelona.
Ferrero, Opium = Ferrero, Jesús, 1991 [1986]. Opium. Plaza y Janés Editores, Barcelona.
Marías, Corazón = Marías, Javier, 1992. Corazón tan blanco. Editorial Anagrama, Barcelona.
Mendoza, Ciudad = Mendoza, Eduardo. 1992 [1986], La ciudad de los prodigios. Seix Barral, Barcelona.
Montero, Bella = Montero, Rosa, 1993. Bella y oscura. Seix Barral, Barcelona.
Sender, Réquiem = Sender, Ramón J., 1992 [1950]. Réquiem por un campesino espanol. Ediciones Destino, Barcelona.

Appendix

1	number of correspondents of SP in pre-text
2	semantic properties of closest correspondent of SP in pre-text
3	semantic properties of second closest correspondent of SP in pre-text
4	formal properties of closest correspondent of SP in pre-text
5	formal properties of second closest correspondent of SP in pre-text
6	predication goal of closest correspondent of SP in pre-text
7	predication goal of second closest correspondent of SP in pre-text
8	number of correspondents of SP in post-text
9	semantic properties of closest correspondent of SP in post-text
10	semantic properties of second closest correspondent of SP in post-text
11	formal properties of closest correspondent of SP in post-text
12	formal properties of second closest correspondent of SP in post-text
13	predication goal of closest correspondent of SP in post-text
14	predication goal of second closest correspondent of SP in post-text
15	number of correspondents of S in pre-text
16	semantic properties of closest correspondent of S in pre-text
17	semantic properties of second closest correspondent of S in pre-text
18	formal properties of closest correspondent of S in pre-text
19	formal properties of second closest correspondent of S in pre-text
20	predication goal of closest correspondent of S in pre-text
21	predication goal of second closest correspondent of S in pre-text
22	number of correspondents of S in post-text
23	semantic properties of closest correspondent of S in post-text
24	semantic properties of second closest correspondent of S in post-text
25	formal properties of closest correspondent of S in post-text
26	formal properties of second closest correspondent of S in post-text
27	predication goal of closest correspondent of S in post-text
28	predication goal of second closest correspondent of S in post-text

Table A: Annotation categories