The Thetic/Categorical Distinction and Bare Nominals in Spanish

Author(s): William Byrne


Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via http://linguistics.berkeley.edu/bls/.

The Annual Proceedings of the Berkeley Linguistics Society is published online via eLanguage, the Linguistic Society of America's digital publishing platform.
1. Introduction

In this paper I argue that the restricted distribution of bare nominals in Spanish is properly treated as a semantic/pragmatic problem. To be specific, since bare nominals in Spanish are inherently weak, i.e. they have a cardinal, non-specific reading, they are disallowed in environments that call for strong NPs, i.e. those with specific, partitive, or generic readings. Focusing on one-place predicates, I will show that the distribution of Spanish bare nominals is directly related to the thetic/categorical judgment distinction, which is based on the speaker’s construal of a given situation (Kuroda 1992). The logical subject of a categorical judgment by definition has a strong construal. Therefore, if a Spanish sentence represents a categorical judgment, its logical subject cannot be represented by a bare nominal. I will also discuss the factors that influence whether a given sentence is interpreted as a categorical vs. a thetic judgment. These include the information structure of the sentence and whether the sentence involves a stage-level or individual-level predicate. Remaining restrictions on bare nominals are shown to be a result of the interaction between the lexical semantics of certain verbs and the information structure of the sentences in which they occur.

2. Distributional facts

The distribution of bare nominals in Spanish is sensitive to the stage-level/individual-level distinction. Stage-level predicates are state descriptions associated with particular situations or events, e.g. drunk, in the living room, open, etc. Individual-level predicates are stable properties that generalize over situations, e.g. tall, intelligent, a fireman, etc. (Carlson 1977). I will divide the data accordingly in §2.1 and §2.2.

2.1 It has often been claimed that bare nominals in Spanish are disallowed as preverbal subjects under conditions of normal stress and intonation. The type of example that seems to support this claim are given in (1) and (2).

(# = unacceptable in given or assumed context; * = ungrammatical)

(1) #Soldados llegaron.
    students arrived-PRET

(2) #Abogados trabajaron allí.
    lawyers worked-PRET there

However, very little attention has been paid to the fact that the status of examples like these greatly improves when the context makes it clear that they describe particular situations in progress. For example, the sentences in (3-4) are grammatical despite their use of preverbal bare nominal subjects.

(OM = object marker for specific animate DOs and all IOs; CL = clitic; IMP = imperfect tense; PRET = preterit tense; CF = contrastive focus; T = topic)
(3) La batalla continuaba. **Soldados** llegaban al campamento heridos y agotados. The battle continued-IMP soldiers arrived-IMP to-the camp hurt and exhausted.

(4) Un viento huracanado inundó el recinto de los primates. **Gorilas** a wind hurricane inundated-PRET the enclosure of the primates gorillas abrazaron a sus crias; **orangutanes** gritaron estrepitosamente. hugged-PRET OM their babies orangutanes screamed-PRET noisily. 'A hurricane wind filled the primate enclosure. Gorillas hugged their young; orangutans screamed noisily.'

In addition, it is well known that when preverbal bare subjects carry contrastive focus stress, the sentences are perfectly grammatical, as shown in (5) and (6).

(5) **[SOLDADOS]**CF llegaron, no generales. 'SOLDIERS arrived, not generals.'

(6) **[ABOGADOS]**CF trabajaron allí, no profesores. 'LAWYERS worked there, not professors.'

It seems then, that preverbal bare nominals in Spanish are not ungrammatical. Rather, they are infelicitous in certain contexts. There are also restrictions on the distribution of postverbal bare nominal subjects, as seen in examples such as (7) and (8).


(8) Llegaron estudiantes. arrived-PRET students 'Students arrived.'

The sentence in (7) sounds odd in isolation, yet (8) is fine (see also Torrego 1989). However, we will see that the acceptability of examples like (7) and (8) depends on the lexical semantics of the verb and the discourse involved.

2.2 Individual-level predicates disallow Spanish bare nominal subjects in all cases. Consider the examples in (9-10).

(9) a. *Gente es inteligente. people is intelligent 'People are intelligent.'

b. *Es inteligente gente. is intelligent people 'People are intelligent.'

(10) *[GENTE]**CF es inteligente, no monos. (adapted from [14c] Contreras 1996) 'PEOPLE are intelligent, not monkeys.'
As shown, bare nominals in this case are ruled out whether preverbal or postverbal and whether or not they carry contrastive focus stress.

2.3 To summarize, sentences with stage-level predicates allow preverbal bare nominal subjects when they carry contrastive focus stress and when the context makes it clear that the sentence describes a situation in progress, as shown in (3-6). There are also restrictions on postverbal subjects, as shown in (7-8). In contrast, sentences involving individual-level predicates disallow bare nominal subjects in all cases.

3. Restrictions on bare nominals in individual-level predicates.

The data above reveals two separate problems: On the one hand, we must account for the fact that individual-level predicates disallow bare NP subjects in all cases (cf. 9-10). On the other hand, we must account for the restrictions on the distribution of bare NP subjects among stage-level predicates (cf. 1-8). We will start with individual-level predicates. Before getting to the details, however, it is necessary to give some background on the thetic/categorical distinction and on the semantics of Spanish bare nominals. This is done in the following subsections.

3.1 As stated in §1, the distribution of bare nominal subjects is directly related to the thetic/categorical distinction. In particular, sentences representing categorical judgments preclude the use of a bare nominals as their logical subject. According to Kuroda (1992), the utterance of a sentence must be viewed on two levels: both as a PROPOSITION and as a JUDGMENT. A sentence represents a proposition inasmuch as it bears a truth value with respect to what the facts are. However, Kuroda has argued in his studies of English and Japanese that certain grammatical distinctions do not yield a difference in truth conditions. Rather, they correspond to functional distinctions based on how the cognitive agent perceives of a given situation. These distinctions correspond to different judgment forms.

The notion of categorical judgment is fairly straightforward because of its association with the traditional subject/predicate distinction of Aristotelian logic. The cognitive act involved in a categorical judgment involves two steps: Initially the speaker designates or recognizes something as the logical subject and then predicates a property of it. For example, consider the sentence in (11). As a categorical judgment, the logical subject Zelda is recognized as separate from the property is a Zapatista. Given this two-step process, categorical judgments are often referred to as DOUBLE JUDGMENTS.

(11) CATEGORICAL JUDGMENT: Recognition of an individual (logical subject) as separate from the property predicated of it.

example: [logical subject Zelda] [predicate is a Zapatista].

Since in a categorical judgment the speaker picks out a certain entity as the target of predication, the NP representing this entity will have a strong construal. That is, it must be identifiable and distinguishable from others, either as a unique object or class, in which case it has a specific reading, or as a unique set or subset, in which case it has a partitive or generic reading.
The thetic judgment, in contrast, is said to be self-contained; that is, it is equated with the description of a single perception in which the speaker merely recognizes the existence of an entity or a situation. Unlike categorical judgments, thetic judgments have no division of logical subject vs. predicate; there is no entity that is recognized as separate from the general perception of the situation. In this way, weather sentences like *It's raining* in (12) are unambiguous examples of thetic judgments given that there is no referent that could be the logical subject. Thereconstructions in English are also prototypical examples of thetic judgments.

(12) **Thetic Judgment**: Perception of an event or situation (eventuality).

(Description)

- examples: [event It's raining].
  [event There are dogs in the yard].

There is an important relation between the stage-level/individual-level distinction and the thetic/categorical distinction. Namely, while sentences with stage-level predicates can potentially represent either thetic or categorical judgments, sentences involving individual-level predicates can only represent categorical judgments. Consider the example in (13), which involves the stage-level predicate *is reading a book*.

(13) John is reading a book.  (Stage-level predicate: Thetic or Categorical)

Mejías-Bikandi 1993 points out that, as a categorical judgment, the speaker who utters (13) is attributing the property of reading a book to the individual denoted by the expression *John*. In contrast, as a thetic judgment, the speaker is not really saying something about the individual denoted by John; rather he or she is just describing a particular event or state of affairs. Under this interpretation, *John*, in spite of being a grammatical subject, is not a logical subject.

Now compare (13) with (14), which involves the individual-level predicate *is a carpenter*.

(14) John is a carpenter.  (Individual-level predicate: Categorical Only)

The sentence in (14) can only represent a categorical judgment. This is because sentences with individual-level predicates do not make reference to any particular situation, which is a primary characteristic of thetic judgments.

The relations among individual-level vs. stage-level, thetic vs. categorical, and weak vs. strong are summarized in Table 1 below. As shown, stage-level predicates can potentially represent either thetic or categorical judgments. In contrast, sentences with individual-level predicates can only represent categorical judgments. Finally, regardless of the predicate-type, the logical subject of a categorical judgment must be strong.

<table>
<thead>
<tr>
<th>Stage-level Predicate</th>
<th>Weak NP</th>
<th>Strong NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual-level Predicate</td>
<td>THETIC/CATEGORICAL</td>
<td>CATEGORICAL</td>
</tr>
</tbody>
</table>

Table 1.
3.2 I stated earlier that the distribution of Spanish bare nominals is a result of a conflict between their inherently weak interpretation and environments that call for strong NPs. In particular, the fact that Spanish bare nominals cannot be construed as the logical subject of a categorical judgment limits their distribution.

Evidence that Spanish bare nominals can only have a weak interpretation has been given recently in McNally 1995 and Laca 1996, and has also been suggested in earlier work. One way in which the inherent weakness of Spanish bare nominals manifests itself is in the fact that, unlike with English bare nominals, Spanish bare nominals cannot denote kinds. Consider the examples in (15) taken from Laca 1996. The use of the predicate extinguirse ‘to become extinct’, requires reference to kinds. However, as shown in the ungrammatical (15a), the bare nominal tigres ‘tigers’ does not suffice. To get the kind reading, the definite NP los tigres must be used, as shown in (15b).

   b. En la India se están extinguiendo los tigres. in the India CL are becoming-extinct the tigers
      ‘Tigers are becoming extinct in India.’

The inherently weak interpretation of Spanish bare nominals is also reflected in the fact that they must take narrowest possible scope with respect to any other operators in the same clause. Compare (16a) with (16b).

      today Juana has that read some articles
      ‘Today Juana has to read some articles.’
   b. Hoy Juana tiene que leer artículos. (d only)
   c. (∃x: article(x)) [φ [read (j, x)]
   d. φ (∃x: article(x)) [read (j, x)]

As shown in (16a), NPs with indefinite determiners, such as unos artículos ‘some articles’ can take either wide or narrow scope with respect to the modal operator and therefore this sentence can be interpreted as either (16c) or (16d). That is to say, unos artículos can have either a specific or nonspecific reading. By contrast, the bare nominal artículos in (16b) only has the nonspecific reading, i.e. the interpretation in (16d). The same scope effects are seen with negation and intensional predicates.

3.3 With this background on the thetic/categorical distinction and on the semantics of Spanish bare nominals, we can begin to explain the stage-level/individual-level contrast presented in §2. The relevant examples will be repeated below. First of all, recall that sentences with stage-level predicates allow bare nominal subjects, as shown in (17) and (18), while those with individual-level predicates do not, as shown in (19) and (20).
Stage-level Predicates

(17) La batalla continuaba. Soldados llegaban al campamento heridos y agotados. ‘The battle continued on. Soldiers arrived to camp hurt and exhausted.’

(18) [Soldados]CF llegaron, no generales. ‘SOLDIERS arrived, not generals.’

Individual-level Predicates

(19) *Gente es inteligente. ‘People are intelligent.’

(20) *[GENte]CF es inteligente, no monos. ‘PEOPLE are intelligent, not monkeys.’

Given the relation between predicate-types and judgment forms outlined in the last section, the explanation for the restriction on Spanish bare nominals in individual-level predicates is straightforward. The two-part derivation is as follows (cf. Ladusaw 1994):

1. The logical subject of a categorical judgment must be strong. Spanish bare nominals are inherently weak. Therefore, Spanish bare nominals cannot serve as the subject of a categorical judgment.

2. Sentences with individual-level predicates represent only categorical judgments. It follows that bare nominals cannot serve as subjects of individual-level predicates.

4. Restrictions on bare nominals in stage-level predicates

So far I have accounted for the restriction on bare nominals in sentences with individual-level predicates. However, the data presents two additional puzzles, which involve restrictions on Spanish bare nominals among sentences with stage-level predicates.

First, recall that speakers find sentences like (21) and (22) unacceptable, at least in isolation. It is only when context is added so that the sentence describes a situation in progress, as in (17) above, or when the preverbal subject carries contrastive focus, as in (18), that they become acceptable.

(21) *Soldados llegaron.
    soldiers arrived-PRET

(22) *Soldados trabajaron.
    soldiers worked-PRET

I will argue that the categorical/thetic distinction accounts for these effects as well. That is, in isolation, sentences such as (21-22) are interpreted as categorical judgments as a result of the default TOPIC-COMMENT information structure which is assigned to them. Since sentences that represent categorical judgments preclude the
use of bare subjects (as we saw in §3), (21-22) are ruled out. In contrast, the context and information structure in (17) and (18) allow them to be interpreted as thetlic judgments, in which case the bare nominal subject is licensed.

Second, we must account for the distinction between (23) and (24) below. Recall that, in isolation, postverbal bare subjects are ruled out with estornudar ‘sneeze’, yet they are fine with llegar ‘arrive’.

(23) *Estornudaron estudiantes.
    sneezed-PRET students

(24) Llegaron estudiantes.
    arrived-PRET students
    ‘Students arrived.’

Unlike the restriction on preverbal bare subjects shown in (21-22) above, the discrepancy in (25-26) is not due to the thetlic/categorical distinction. In fact, it turns out that both examples represent thetlic judgments. Rather, I will argue that the unacceptability of (23) is a result of a conflict between the lexical semantics of estornudar and the PRESENTATIONAL information structure associated with this sentence. I begin in §4.1 with a discussion of the relationship that pragmatic and contextual factors have with word order.

4.1 In every language, speakers use different grammatical structures under different communicative circumstances (Lambrecht 1994). This grammatical encoding of discourse and contextual information in natural language is referred to here as INFORMATION STRUCTURE, which I will take to be a separate component of grammar.

For example, consider the discourse in (25). The sentence in (25a) conforms to the well-known TOPIC-COMMENT relation, which is a relation of aboutness between the pragmatically presupposed subject referent and the proposition expressed by the sentence. In other words, this sentence takes a familiar or active discourse referent as its subject, i.e. un hombre ‘one of the men’, and predicates a property of it. Although different word orders are possible in Spanish, a topic-comment interpretation is only viable in sentences with preverbal subjects. The comparison of (25a) with (25b) illustrates this point. If un hombre is postverbal, as shown in (25b), the topic-comment interpretation is not available.

(25) Había varios hombres en el bar.
    there-were-IMP several men in the bar
    ‘There were several men in the bar.’

a. [Un hombre] estaba fumando.
    one man was-IMP smoking
    ‘One of the men was smoking.’

    [...]

b. #Estaba fumando un hombre.
    ‘A man was smoking.’

Following Lambrecht 1994 among others, I will assume that topic-comment sentences are the unmarked or default sentence-type with respect to the discourse. As Lambrecht puts it ‘The topic-comment articulation is [...] communicatively speaking the most useful pragmatic articulation. It is therefore the one to which speakers will most naturally resort for the pragmatic construal of isolated sentences’
(1994:132). On this view, speakers will assign the topic-comment information structure to sentences with the SV(O) word order in neutral contexts.

Two additional sentence types involve non-topical subjects: PRESENTATIONAL and EVENT-REPORTING sentences. As pointed out by Sufier 1982, among others, the function of PRESENTATIONAL sentences is to introduce (or reintroduce) a ‘new’ or inactive referent usually for the purpose of making it the topic of predication in subsequent discourse. This is shown in (26a). In this way, the verb involved expresses existence, appearance, emergence, etc. Furthermore, sentences with the presentational function have postverbal subjects. As shown in (26b), the preverbal subject in this context is infelicitous.

(26) a. De repente apareció una mujer. Era alta, guapa, y misteriosa.
    of sudden appeared-PRET a woman was-IMP tall attractive and mysterious
    ‘Suddenly a woman appeared. She was tall, attractive and mysterious.’

   b. #Una mujer apareció de repente. [...]‘A woman appeared suddenly.’

In contrast to (26a), the function of the EVENT-REPORTING sentence in (27) is simply to describe a situation or event without focussing on any particular entity (Lambrecht 1994:144). Furthermore, in this case the grammatical subject is preverbal, although non-topical.

(27) El sitio estaba como siempre. Un hombre estaba fumando, una canción
    the place was-IMP like always a man was-IMP smoking, a song
    triste sonaba en la radio, y yo trabajaba de camarero.
    sad sounded-IMP on the radio and I-worked-IMP of waiter
    ‘The place was like always. A man was smoking, a sad song was playing on
    the radio, and I was tending bar.’

Thus, although the second sentence in (27) is identical to (25a) on the surface, the two differ with respect to information structure.

The three-way contrast among topic-comment, presentational, and event-reporting sentences is also found in Mejías-Bikandi 1993, as shown in (28a-c).³

    a tren arrived-IMP to-the platform
    ‘Trains (kind) used to arrive at the platform.’

   b. Al andén llegó un tren.
      to-the platform arrived-PRET a train
      ‘A train arrived at the platform.’

   c. Un tren llegaba al andén.
      a tren arrived-IMP to-the platform
      ‘A train was arriving at the platform.’

4.2 There is a clear correlation between the topic-comment sentences and the categorical judgment on the one hand, and between the presentational and event-reporting sentences and the thetic judgment on the other hand. To be specific, the
topic NP in a topic-comment structure bears a pragmatic relation of aboutness to the proposition expressed by the sentence. Thus, the topic NP is recognized as 'separate' from the proposition predicated of it in the same way the logical subject of a categorical judgment is recognized as a separate entity. In addition, the pragmatically presupposed nature of topics make them necessarily strong, i.e., they have either specific, partitive, or kind readings. In the case of presentational and event-reporting sentences, both represent thetic judgments since neither type adheres to the traditional subject-predicate distinction. That is to say, presentational and event-reporting sentences describe an eventuality (28c). In this way, following Sasse 1987 and Lambrecht 1994, I will claim that the categorical and thetic judgments are superordinate categories for these information structures, as illustrated in (29a-b) below.

(29)  a.  Thetic  
      
      Presentational  
      
      Event-reporting  

b.  Categorical  

4.3 We can now account for the discrepancy between examples (21) and (17), repeated below as (30a) and (30b) respectively.

(30)  a.  #Soldados llegaron.  
       'Soldiers arrived.'

b.  La batalla continuaba. Soldados llegaban al campamento heridos y agotados.  
       'The battle continued on. Soldiers were arriving to the camp hurt and exhausted.'

The sentence in (30a) has a preverbal subject, which on the face of it means it should be compatible with either the topic-comment or the event-reporting function. However, given the lack of context, speakers will assign the default topic-comment information structure to the sentence, which represents a categorical judgment. Since the inherent weakness of Spanish bare nominals precludes their use as subjects of categorical judgments, (30a) is unacceptable. Yet once the appropriate context is supplied for an event-reporting interpretation, a type of thetic judgment, the preverbal bare subject becomes acceptable, as is shown in (30b). This is because thetic judgments do not require strong subjects.

4.4 As mentioned earlier, the discrepancy between (23) and (24), repeated below as (31a-b) respectively, is a result of the interaction between the presentational sentence function and the lexical semantics of the verbs involved. Recall that without context (31a-b) will be interpreted as presentationalals given their postverbal subjects. This creates a problem in the case of (31a) because the verb estornudar 'sneeze' does not connote existence, appearance, or emergence, as is required in a presentational sentence. I propose that it is for this reason that sentence in (31a) sounds odd. In contrast, the ideas of appearance and emergence are central to the lexical semantics of llegar 'arrive', which is why there is no such conflict in (31b).
(31) a. #Estornudaron estudiantes.
    sneezed students
    ‘Students sneezed.’

b. Llegaron estudiantes.
    arrived students
    ‘Students arrived.’

5. A Syntactic Approach

Before concluding, I will discuss the drawbacks of trying to account for these data using a purely syntactic approach.

5.1 Contreras (1996) argues that the bulk of the restrictions on the distribution of bare nominals are a result of head government violations. To be specific, he proposes that bare nominals are quantifier phrases with an empty head position, as shown in (32).

(32) [QP [Q e] NP]

On this view, bare nominals in object positions are predicted to be grammatical since they are licensed by the verb or preposition that governs them. By the same token, bare nominals in ungoverned positions such as preverbal subjects as well as postverbal subjects in adjoined positions are predicted to be ungrammatical.

5.2 There are three problems with this analysis. First, it does not account for examples such as (30b) Soldados llegaban al campamento heridos y agotados and (33) below, which are grammatical despite their use of bare nominals in ungoverned positions (cf. McNally & Fontana 1994).

(33) La cosecha; [S [VP [V’ la(1)-destruyeron $t_i$ ] langostas]] (= [22a] Sufiñer
    the harvest   obj-cl destroyed   locusts 1982: 213)
    ‘As for the harvest, locusts destroyed it.’

Secondly, with this approach, we would predict sentences such as (34) and (35) to be grammatical given that the bare nominal objects are governed by the verb. Yet they are ungrammatical.

(34) *Pedro [V’ detesta salmón].
    (Pedro hates salmon.)
    (cf. Pedro detesta el salmón.)

(35) *Pedro [VP [V’ come salmón] [AP crudo]]. (cf. Pedro come el salmón crudo.)
    (Pedro eats salmon raw.)

Finally, this syntactic approach cannot explain the discrepancy between the acceptability of bare nominal subjects in stage-level as opposed to individual-level predicates (§2).
6. Conclusion

I have shown that the distribution of bare nominals in Spanish is due to a clash between their inherent weak interpretation and environments that require strong NPs. Sentences that represent categorical judgments disallow bare nominals because their logical subjects must be strong. Thetic judgments, in contrast, have no logical subject and do not place restrictions on bare nominals. Since sentences with individual-level predicates always represent categorical judgments, it follows that bare nominals do not occur as the subjects of individual-level predicates. Sentences with stage-level predicates only disallow Spanish bare nominals when they represent categorical judgments. Whether or not a sentence represents a categorical vs. thetic judgment depends on the information structure involved. If the sentence has a topic-comment structure, it represents a categorical judgment. Presentational and event-reporting sentences, in contrast, represent thetic judgments. Finally, restrictions on postverbal bare nominals in isolated examples sometimes depends on the extent to which the lexical semantics of the verb involved are compatible with the discourse function of the sentence, as we saw with estornudar 'sneeze' and llegar 'arrive' in (33a-b).

Endnotes

* I would like to thank John Moore, Farrell Ackerman, Chris Barker, and Raúl Aranovich for valuable comments and discussion. All errors are my responsibility.


2. See references in footnote 1 above as well as Krifka et. al. 1995 for discussion of the stage-level/individual-level distinction.

3. Mejías-Bikandi 1993 does not discuss these differences in terms of information structure, however.

4. This does not apply to predicate nominals, e.g. Juan es abogado.
References


LADUSAW, WILLIAM. 1994. Thetic and categorical, stage and individual, weak and strong. SALT 4.


MCNALLY, LOUISE. 1995. Bare plurals in Spanish are interpreted as properties. ms. Universitat Pompeu Fabra.


(byrne@ling.ucsd.edu)