Complement Anaphora in Spanish: Proportional References and Discourse Relations

Iker Zulaica-Hernández

Abstract Although still disputed by some authors, complement anaphora interpretations are widely accepted today. However, they are marked and subject to strict interpretation conditions. Most commonly, monotone decreasing quantifiers facilitate a salient complement set interpretation that is suitable for subsequent anaphoric reference, whereas monotone increasing quantifiers appear to block such interpretation. In this paper, I investigate the possibilities of anaphoric reference to complement sets in Spanish quantified expressions of type QUANTIFIER(A)(B) and argue that Spanish aligns with the general pattern observed cross-linguistically according to which anaphoric reference to the reference set (refset) is default regardless of monotonicity type, and that anaphoric reference to the complement set (compset) is only possible with monotone decreasing quantifiers. My claim is that the observed discrepancy can be explained in terms of the different proportions denoted by the quantifiers, and propose an explanation based on discourse relations. My approach provides additional cross-linguistic evidence in favor of the general observed pattern, and a promising path for future investigation into the refset/compset divide.

Keywords complement anaphora, quantifiers, monotonicity, discourse, Spanish

Introduction

Antecedents of anaphoric pronouns can have different forms and denote different semantic objects in natural languages. The amount of cognitive effort needed by speakers to resolve anaphors varies depending, among other factors, on antecedent type and denotation. Under normal conditions, not much effort is needed to process and resolve the pronoun in (1), where the null pro-form in the second sentence unambiguously corefers with the noun phrase in the first sentence (coreference is marked with a subscript).\(^1\) The null pro-form refers to the reference set and its antecedent is highly salient (see Ariel, 1990; Gundel et al., 1993; Almor et al. 2017 for discussion on the accessibility status of pronoun antecedents).

(1) \(\text{Juan, se ha comprado un barco. } \text{pro}_{i} \text{ es muy rico.}\) ‘John just bought a yacht. He is very rich.’

However, some pronoun antecedents are not easily accessible in some specific anaphoric configurations, and they need to be inferred. In complement anaphora (Kibble, 1997; Nouwen, 2003) a pronoun following the expression QUANTIFIER(A)(B) can pick as its reference those A’s which are B (the reference set or refset, henceforth), as in (2), or those A’s which are not B (the complement set or compset, henceforth), as in (3). In these two examples the quantified expression puts the reference set or the complement set in focus, and pronominal anaphora is licensed.

\(^1\) Spanish is a null-subject language. For the sake of clarity, I will keep my examples simple throughout this paper and my subject pronouns will be, for the most part, phonetically null pronouns: pro_{personal}
Complement anaphora is an interesting phenomenon insofar as it seems to dispute general claims about anaphora and reference, and has raised some concerns as regards its true status. Some authors like Corblin (1996) and Percus et al. (1997) have challenged the reference to compset arguing that the alleged cases of reference to complement set are, in fact, cases of generic reference. However, because the experimental evidence in favor of complement set reference is strong (see Moxey & Sanford, 1993, and references therein), the notion of compset reference is widely accepted among semanticists today. To my knowledge, complement anaphora has been little studied in Spanish, and this paper aims to provide a general overview of this particular type of anaphoric reference that may serve as a reference to subsequent studies, of a theoretical or experimental nature, on the possibilities of pronominal reference to complement sets in Spanish quantified sentences. I also present a preliminary view on complement anaphora that combines quantifier focus effects and discourse structure in order to explain the possibility of compset readings with quantified expressions. The majority of studies on complement set anaphora to date have focused on the logical properties of quantifiers and have largely ignored the role that discourse structure may play in facilitating the inference to compset, or lack thereof. In this paper, I argue that, when they interpret proportional quantifiers in context, speakers show a strong tendency to focus on the highest proportion of entities denoted by the quantified expression, regardless of their monotonicity properties. Assuming, as it is reasonable, that the refset is default, appealing to the highest proportion can explain why two possible anaphoric interpretations are available for monotone decreasing quantified expressions: refset is default and the highest proportion corresponds to compset. On the other hand, refset and the highest proportion coincide in monotone increasing quantifiers, being the lowest proportion inaccessible unless it is accommodated. Following the theoretical framework of Segmented Discourse Representation Theory (Asher & Lascarides 2003), I argue that plausible rhetorical connections can be established from refset and highest proportion interpretations among the relevant discourse segments, namely, the quantified proposition and the proposition with the pro-form. World knowledge and specific grammatical, and lexical clues may also help hearers to infer a particular rhetorical relation among the constituent segments and help them resolve the potential ambiguity.

A brief note on natural language quantifiers

Natural language determiners can be represented as two-place relations between sets of entities in terms of shared membership.2 These two sets are: the set denoted by the noun meaning (the restrictor) and the set denoted by the verb meaning (the nuclear scope). For example, the quantifier represented by the determiner ningún (‘no’) in ningún estudiante aprobó ‘no student

2 The logical properties of natural language quantifiers have been extensively studied. See, for example, Barwise and Cooper (1981); Van Benthem (1986); Keenan (1996), and Gutiérrez-Rexach (2003) for an excellent study on quantification in Spanish.
passed’ relates the two sets denoted by the noun and VP denotations, and is true with respect to some model $M$ only if the intersection of the set of students and the set of things that passed is the empty set, as in (4). An existential quantifier such as un in un estudiante aprobó ‘a student passed’ will be true, with respect to some model $M$, if the intersection of the two sets share at least one individual member, that is, the set is non-empty, as in (5).

\[(4) \quad \llbracket \text{estudiante} \rrbracket^M \cap \llbracket \text{aprobó} \rrbracket^M = \emptyset \]

\[(5) \quad \llbracket \text{estudiante} \rrbracket^M \cap \llbracket \text{aprobó} \rrbracket^M \neq \emptyset \]

Quantifiers can also be represented functionally as denoting expressions of type $(e \to t) \to ((e \to t) \to t)$ that map sets onto sets of sets. Such expressions are called generalized quantifiers. Under the functional view, the negative quantifier ningún is the function that assigns to every noun denotation $\llbracket N \rrbracket$ the family of sets $\{X\mid \llbracket N \rrbracket^M \cap X = \emptyset\}$, and the indefinite determiner un the function that assigns to every noun denotation $\llbracket N \rrbracket$ the family of sets $\{X\mid \llbracket N \rrbracket^M \cap X \neq \emptyset\}$ (Cann et al., 2009, p.107). This is illustrated with examples (6) and (7).

\[(6) \quad \text{Ningún estudiante aprobó.} \quad \text{‘No student passed.’} \quad \llbracket \text{(Ningún (estudiante')) (aprobó')} \rrbracket^M \text{ is true iff the cardinality of the intersection of the sets denoted by } \llbracket \text{estudiante} \rrbracket^M \text{ and } \llbracket \text{aprobó} \rrbracket^M \text{ is or equals 0.} \]

\[(7) \quad \text{Un estudiante aprobó.} \quad \text{‘A student passed.’} \quad \llbracket \text{(Un (estudiante')) (aprobó')} \rrbracket^M \text{ is true iff the cardinality of the intersection of the sets denoted by } \llbracket \text{estudiante} \rrbracket^M \text{ and } \llbracket \text{aprobó} \rrbracket^M \text{ is greater than or equals 1.} \]

Monotonicity is one of the formal properties of generalized quantifiers that are most relevant for our purposes. Monotonicity is relevant as it involves the possibility of inference to superset or subsets of the set under consideration. Informally, this property enables us to check whether adding or subtracting entities from the sets denoted by restrictor and nuclear scope may affect the truth value of a quantified sentence. Formally, a quantifier $Q$ is said to be monotone decreasing (MON↓) if for any sets $A$, $B$, $B' \subseteq E$: if $Q(A,B)$ and $B' \subseteq B$ then $Q(A,B')$, as in (8). On the other hand, a quantifier $Q$ is said to be monotone increasing (MON↑) if for any sets $A$, $B$, $B' \subseteq E$: if $Q(A,B)$ and $B \subseteq B'$ then $Q(A,B')$, as in (9).\(^3\)

\[(8) \quad \text{Pocos estudiantes estudiaron } \models \text{ Pocos estudiantes estudiaron francés.} \quad \text{‘Few students studied entails Few students studied French.’} \]

\[(9) \quad \text{Todos los estudiantes estudian francés } \models \text{ Todos los estudiantes estudian.} \quad \text{‘All students study French entails All students study.’} \]

---

**Cardinality, proportions and context-dependent quantifiers**

\(^3\) For the purposes of this paper, I will only consider monotonicity based on the increase or decrease in the extension of the right argument of the quantifier, that is, its nuclear scope or VP denotation.
As we have seen in the previous section, some quantifiers such as the universal *todos* (‘every’) or *ningún* (‘no’) and existential *un* (‘a’) receive a fairly general but straightforward interpretation in terms of empty or non-empty set intersection. Other quantifiers are cardinal because they refer to a number of elements in a set. Some of them give a precise number of elements, or specify that the cardinality of the set is greater than or equal to \((\geq)\), or less than or equal to \((\leq)\) some number. Adopting the relational perspective, the quantifiers *cinco* (‘five’) and *no más de siete* (‘no more than seven’) in quantified sentences such as *cinco trabajadores recibieron un ascenso* (‘five employees got a raise’) and *no más de siete casas fueron vendidas* (‘no more than seven houses were sold’) are given the following interpretation, where \(|A|\) represents the cardinality of the intersection.

\[
\begin{align*}
\text{(10)} & \quad \llbracket \text{Cinco} \ (A,B) \rrbracket = 1 \iff |A \cap B| = 5 \\
\text{(11)} & \quad \llbracket \text{No más de siete} \ (A,B) \rrbracket = 1 \iff |A \cap B| \leq 7
\end{align*}
\]

However, there is an interesting class of quantifiers for which we can only assume some underspecified value \(n\) that is heavily dependent on context. These quantifiers are vague, and their truth conditions cannot be easily determined by the number of elements found in the relevant sets. Quantifiers of this kind are *pocos* (‘few’), *muchos* (‘many’), *la mayoría/la mayor parte de* (‘most’), *bastantes* (‘several’), among others. Consider (12), where the quantifier would be true if the number of students that passed the course is greater than the number of students who failed. In (12), \(E\) represents the set of entities in the universe of discourse, and \((E - X)\) is the set of all entities in the universe of discourse minus all the entities that passed the course.

\[
\begin{align*}
\text{(12)} & \quad \text{La mayoría de los estudiantes aprobó.} \\
& \quad \text{‘Most (of the) students passed.’} \\
& \quad | \llbracket \text{estudiante} \rrbracket^M \cap \llbracket \text{aprobar} \rrbracket^M | > | \llbracket \text{estudiante} \rrbracket^M \cap (E - \llbracket \text{aprobar} \rrbracket^M) |
\end{align*}
\]

The interpretation for the quantified NP *la mayoría de los estudiantes* (‘most students’) is shown in (13):

\[
\begin{align*}
\text{(13)} & \quad \llbracket \text{La mayoría de los (estudiante)} \rrbracket^M = \\
& \quad \{X \mid |\llbracket \text{estudiante} \rrbracket^M \cap X| > | \llbracket \text{estudiante} \rrbracket^M \cap (E - X) | \}
\end{align*}
\]

These quantifiers are not only vague, they are, in addition, ambiguous as they also present a proportional reading. Once again, the number of entities that count as *muchos N* or *pocos N* is extremely variable, and highly dependent on context. For Keenan and Stavi (1986) these determiners are *value judgment determiners* in that they give speakers the possibility to express their judgment about the number of entities under discussion, regardless of their specific number. Also, an implicit comparison is assumed in the analysis of determiners like *muchos* and *pocos*; however, the term of the comparison is never given, which adds yet another difficulty to their interpretation. Thus, a quantified sentence such as *pocos aficionados acudieron al partido* (‘few fans attended the game’) can be true compared to the average number of fans that usually attend
a game on Saturday, but it can be false if we compared that number with the number of fans that usually attend a game on rainy days. We can therefore propose at least two possible interpretations for a determiner like *pocos*: (i) a cardinal interpretation (14a) and (ii) a proportional interpretation (14b), as in Gutiérrez-Rexach (2003, p.161). In the cardinal interpretation, a context-dependent parameter $n$ is introduced, and that parameter can take any number depending on every particular situation. For the proportional interpretation in (14b), the universe of discourse $|E|$ is introduced, and the value of *pocos* is obtained via the proportion of the VP extension in relation with the universe $E$.4

\[ pocos_1(A,B) = 1 \text{ iff } |A \cap B| \leq n \]

\[ pocos_2(A,B) = 1 \text{ iff } |A \cap B| < \left( \frac{|B|}{|E|} \right) \times |A| \]

### Complement anaphora

A quantified expression of type $D(A)(B)$ can allow for subsequent anaphoric plural reference to three sets associated with it: the maximal set $A$, the reference set $A \cap B$ and, under some special conditions that will be discussed, the complement set $A \cap \neg B$. In complement anaphora, the pronoun anaphorically refers to the complement set. In what follows, I adopt Nouwen’s terminology (2003, p.74): given a quantificational structure $D(A)(B)$, where $D$ is a conservative determiner, $A$ is its restrictor, and $B$ its nuclear scope, I call the denotation of $A$ the *maximal set* or *maxset*; the intersection of the denotations of $A$ and $B$ will be referred to as the *reference set* or *refset*; and the theoretical difference between the denotations of restrictor $A$ and nuclear scope $B$ will be called the *complement set* or *compset*.5 In (15), for example, the maxset is the set of all associates (or a contextual relevant set of associates, as in Westerståhl (1985); the reference set is the set of associates who actually attended the meeting; and the complement set is the set of associates who did not attend the meeting.

(15) Muchos socios acudieron a la asamblea.

‘Many associates attended the meeting.’

In a series of experimental studies on complement anaphora, Moxey and Sanford (1993, 1996) and Sanford, Moxey and Paterson (1996) reported that nearly 71% of responses involving negative quantifiers, mostly monotone decreasing quantifiers, resulted in compset continuations, while positive quantifiers resulted almost exclusively in refset continuations. In their experiments, they asked participants to provide a continuation to quantified expressions of the form ‘$Q$ of the $A$’s were $B$’, as in (16) and (17). Subjects were then asked to indicate the reference of the pronoun ‘they’ as they were presented with a set of possibilities: MPs in general,

---

4 There are other possible proportional interpretations for determiners *muchos* and *pocos* based on a variety of semantic, pragmatic or syntactic considerations. See Gutiérrez-Rexach (2003) for a detailed discussion.

5 Conservativity (Barwise & Cooper, 1981; Keenan & Stavi, 1986) is a universal property of determiners that indicates that the first argument is the crucial element for interpreting the determiner. For example, the sentences ‘every boy walks’ and ‘every boy is a boy who walks’ are equivalent. Formally:

**CONSERVATIVITY (CONS):** A determiner $D$ is conservative iff for every $A$, $B \subseteq E$: $D_E(A)(B) \iff D_E(A)(B \cap A)$. 


all MPs, MPs who went to the meeting, MPs who did not go to the meeting, or none of the above. They found that subjects consistently used the plural pronoun to refer to the complement set (the MPs who did not attend the meeting) when Q was replaced with the determiners *hardly any* (‘casi ningún’), *not many* (‘no muchos’), *very few* (‘muy pocos’), and *few* (‘pocos’). Interestingly, the determiner *only a few* (‘sólo unos pocos’) only showed reference to compset in structures with the connective *because*. Another study involving numerical quantifiers (Moxey & Sanford, 1993) showed that the determiner *less than n%* (‘menos del n%’) had a preference for compset continuations.

(16) Q of the MPs attended the meeting. They...
(17) Q of the MPs attended the meeting, because they...

On the other hand, determiners such as *many* (‘muchos’), *a few* (‘unos cuantos’), *almost all* (‘casi todos’), *more than half* (‘más de la mitad’), and numerical determiners such as *n%* (‘un n%’), *only n%* (‘sólo un n%’), *more than n%* (‘más del n%’) showed a strong preference for refset reference (i.e. the set of MPs who attended the meeting), while they showed no preference or hardly any preference for compset reference continuations. Moxey and Sanford’s experiments point towards a strong correlation with downward monotonicity in English. However, as Kibble (1997) notes, the correlation is not absolute insofar as the non-monotone quantifier *only a few* (‘sólo unos pocos’) also allows for compset continuations. More importantly, regardless of the preference for a particular type of quantifiers, what these experiments show is that reference to compset is not only possible but also common in natural language, and that compset reference seems to strongly correlate with monotonicity in English.

As regards Spanish, it seems that reference to compset is possible with the same type of monotone decreasing quantifiers. Consider the examples in Tables 1 and 2 showing the available interpretations for two sets of monotone decreasing and increasing quantifiers. Two possible continuations are provided for every quantified sentence. Monotone decreasing quantifiers such as *pocos* ‘few’, *muy pocos* ‘very few’, *no muchos* ‘not many’, and *casi ningûn* ‘almost no’ license subsequent anaphoric reference to both refset and compset.

<table>
<thead>
<tr>
<th>Table 1 Available reference with monotone decreasing quantifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monotone decreasing (mon↓)</td>
</tr>
<tr>
<td><em>Muy pocos socios acudieron a la asamblea.</em></td>
</tr>
<tr>
<td>1. pro votaron subir el sueldo al presidente.</td>
</tr>
<tr>
<td>2. pro estaban demasiado ocupados.</td>
</tr>
<tr>
<td>‘Very few associates attended the meeting.*</td>
</tr>
<tr>
<td>1. They voted to raise the president’s salary.</td>
</tr>
<tr>
<td>2. They were very busy.’</td>
</tr>
<tr>
<td><em>Pocos estudiantes han venido hoy a clase.</em></td>
</tr>
<tr>
<td>1. pro querían repasar la lección con el profesor.</td>
</tr>
<tr>
<td>2. pro tenían que preparar un ensayo.</td>
</tr>
<tr>
<td>‘Few students came to class today.*</td>
</tr>
<tr>
<td>1. They wanted to review the lesson with the professor.</td>
</tr>
<tr>
<td>2. They had to write an essay.’</td>
</tr>
</tbody>
</table>
No muchos aficionados acudieron al estadio.
1. pro se sentaron en tribuna principal. ✓
2. pro vieron el partido por televisión. ✓

‘Not many fans went to the stadium.
1. They sat on the main stand.
2. They watched the game on TV.’

Casi ningún atleta participó en el maratón.
1. pro se habían preparado bien para la carrera. ✓
2. pro no se habían preparado bien para la carrera. ✓

‘Almost no athlete took part in the marathon.
1. They prepared well for the race.
2. They didn’t prepare well for the race.’

These observations are compatible with general observations about English quantifiers and reference to the complement set. They are also compatible with Nouwen’s view of complement anaphora as a special type of inference that requires the hearer to infer the existence of a complement set, an inference that is only possible with monotone decreasing (i.e. downwards entailing) quantifiers. As shown in Table 2, reference to the refset is default with monotone increasing quantifiers. In addition, monotone increasing quantifiers (upwards entailing quantifiers) license the inference that a property of a subset is also the property of the maxset. It would thus be possible to say Muchos socios acudieron a la asamblea. De hecho, acudieron todos (‘Many associates attended the meeting. In fact, they all attended.’), however, a reading where the null pro-from refers to the set of associates that did not attend the meeting (the compset) is not available.

Table 2 Available reference with monotone increasing quantifiers

<table>
<thead>
<tr>
<th>Monotone increasing (mon↑)</th>
<th>refset</th>
<th>compset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muchos socios acudieron a la asamblea.</td>
<td>✓</td>
<td>#</td>
</tr>
<tr>
<td>1. pro votaron subir el sueldo al presidente.</td>
<td>✓</td>
<td>#</td>
</tr>
<tr>
<td>2. pro estaban demasiado ocupados.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| ‘Many associates attended the meeting.
1. They voted to raise the president’s salary.
2. They were very busy.’ | |
| Unos cuantos estudiantes vinieron a clase. | ✓ | # |
| 1. pro querían repasar la lección con el profesor. | ✓ | # |
| 2. pro tenían que preparar un ensayo. | | |

6 All anaphoric interpretations for quantifiers in this paper rely on my own judgments as a native speaker of Spanish. As I will point out later on in this paper, experimental evidence is needed, ideally in the form of a questionnaire aimed at a representative sample of native respondents, before these interpretations can be fully confirmed. However, all the native speakers informally consulted agreed with my own judgments regarding examples in Tables 1 and 2.
Although the set of negative and positive quantifiers analyzed in this paper is not exhaustive, my sample seems to indicate that Spanish quantified expressions generally conform with the pattern observed in English regarding the availability of refset/compset interpretations. However, there appear to be certain differences among Spanish quantifiers that need to be addressed. For example, not all negative quantifiers allow for compset interpretations to the same degree. In other words, refset and compset present varying degrees of salience: monotone decreasing quantifiers denoting refsets of minimal cardinality such as ningún, casi ningún or muy pocos seem to favor the compset interpretation more favorably than monotone decreasing quantifiers with more balanced refset/compset cardinalities such as pocos and no muchos. I will return to this question in Sections 5 and 6, where I discuss the relevance of proportions in refset/compset interpretations.

### Plurals, pseudo-reference, and missing antecedents

Some pronouns have plural quantificational structures as antecedents, as in the following example from Evans (1980, p.7), where the plural pronoun is interpreted as the set of all the congressmen who admire Kennedy; i.e. the set of elements that satisfies both restrictor and nuclear scope of the quantifier.

(18) Pocos congresistas admiran a Kennedy, y pro son muy jóvenes.
    ‘Few congressmen admire Kennedy, and they are very young.’

In Discourse Representation Theory (Kamp & Reyle, 1993), an abstraction rule and an abstraction operator $\Sigma$ is proposed that creates plural discourse referents from quantificational expressions. Upon processing discourse (18), a plural discourse referent $X$ is created by abstracting over the maximal set of individuals that satisfy the conditions in the sub-DRS: $X = \Sigma y$. This is shown in Figure 1.
However, it is commonly assumed that not all sets are included in a discourse representation. As Kamp and Reyle argue (1993, p.307): “apparently, subtracting one set from another is not a permissible operation for the formation of pronominal antecedents.” Based on DRT assumptions, only the maximal set -maxset- and maybe other explicitly introduced subsets, would be included in the discourse representation, hence no default set-subtraction operation (i.e. maxset minus refset) can be represented. Under this assumption, given the specific continuation in (19), reference to the most salient interpretation for the quantified expression (i.e. the set of the congressmen not admiring Kennedy) would not be possible.7

(19) Pocos congresistas admiran a Kennedy. pro creen que es un incompetente.
    ‘Few congressmen admire Kennedy. They think he’s incompetent.’

DRT’s limitation in the number and type of available antecedents in quantified structures explains why it is not possible to refer to the missing marble with the pronoun it in (20), that is, there is no available antecedent for the pronoun unless the antecedent is accommodated (Roberts, 1989).

(20) Nine of the ten marbles are in the bag.
    # It’s under the couch.

Roberts (1989), mainly within the context of modal subordination, argues that accommodation is necessary to explain these and similar cases, where the antecedent is ‘missing’. She noticed that the continuation with the pronoun is more felicitous after a long pause, especially if the speaker

7 This example is from Nouwen (2003, p.74).
noticed that the hearer is looking for something. That, for example, would be a clear indication for the speaker that there is a salient entity in the context, and that the hearer has performed a mathematical calculation of set subtraction (i.e. the hearer subtracts nine from ten). Only then, a discourse referent for the missing marble is accommodated. Roberts also argues that if we assume that examples like (21) conversationally implicate that one marble is not in the bag, and (21) implicates that nine marbles are in the bag, then both examples would be true in exactly the same worlds. However, the sentence in (21), unlike that in (20), provides a suitable antecedent for the pronoun.

(21) A: One of the ten marbles is out of the bag.
     B: It’s under the couch.

Corblin (1996) and Percus et al. (1997) argue that reference to compset is not possible, and that what resembles reference to the complement set is actually reference to the maximal set, or \textit{pseudo-reference} to compset. This view is mainly motivated by Kamp and Reyle’s example (22). Under the pseudo-reference account, the pronoun ‘they’ could have two possible interpretations, namely, it may pick up its antecedent from either the maximal set ‘women from the village in general’ (i.e., the generic interpretation), or from the refset ‘women from the village who came to the feminist rally’. The two possible interpretations for the pronoun are shown in (23a-b) where, according to Corblin, the generic interpretation (23b) (i.e., the majority of elements in the restrictor set) is the preferred interpretation, and the one that gets commonly confused with the compset interpretation.

(22) A: Few women from this village came to the feminist rally.
     B: No wonder. They don’t like political rallies very much.

(23) a. $X_1 = \sum x [x | \Phi \land \Psi]$ 
    b. $X_2 = \sum x [x | \Phi]$

However, there are objections to the generic account, too. In (24), the continuation with a definite determiner replacing the pronoun is not acceptable, although one should expect the generic reading to be possible.

(24) Few MPs attended the meeting. The MPs were at the pub or with their secretaries.

Furthermore, Moxey and Sanford (1993) provide sufficient empirical evidence in favor of compset focus, and indicate some problems with the idea that compset references are generalizations. In their experiments, their subjects and two independent judges had the opportunity to choose a generic interpretation for a pronoun ‘they’ after a quantified statement, but they chose the compset reference only. Also, the authors indicate that the expression ‘instead’ was frequent in the subjects’ continuations, a clear indication that reference to compset was intended to be made in most cases (p. 64).

---

8 Or, in other words, (20) cannot be true and (21) be false because they describe exactly the same set of affairs.
9 This example is from Kibble (1997).
Proportions and quantifier focus effects

For Moxey and Sanford (1987), negative quantifiers such as *pocos* (‘few’), *no todos* (‘not all’), and *no muchos* (‘not many’) have the property of focusing on the compset, while positive quantifiers such as *algunos* (‘some’), *unos cuantos* (‘a few’), and *muchos* (‘many’) focus on the refset. They explained these differences in terms of the different focus effects created by quantifiers. For all the examples considered so far in this paper, the refset interpretation is acceptable in all cases, and with all quantifiers regardless of monotonicity type. However, the compset interpretation is only acceptable with negative, monotone decreasing, quantifiers. However, although negative quantifiers allow for a refset interpretation, some of them appear to show a preference for compset continuations or, in other words, the compset appears to be more salient than the refset. This is the case of monotone decreasing quantifiers such as *muy pocos* (‘very few’), *casi ningún* (‘almost no’), and *pocos* (‘few’) in specific discourse configurations that will be discussed in the next sections.

(25) *Todos/algunos/unos cuantos/muchos* (los) estudiantes acudieron a la presentación. ‘All/some/a few/many students attended the presentation.’

a. *pro* escucharon al ponente con interés. ‘They listened to the presenter very attentively.’

b. *# pro* prefirieron quedarse en la cafetería. ‘They preferred to stay in the cafeteria.’

(26) *Ningún/no todos/pocos/no muchos* (los) estudiantes acudieron a la presentación. ‘No/not all/few/not many students attended the presentation.’

a. *¿pro* escucharon al ponente con interés. ‘They listened to the presenter very attentively.’

b. *pro* prefirieron quedarse en la cafetería. ‘They preferred to stay in the cafeteria.’

The inferences related to the different monotonicity properties of quantifiers appear to have a direct effect on the available anaphoric interpretations. Thus, positive quantifiers license upwards inferences and focus the hearer’s attention on the refset. Negative quantifiers, on the other hand, license downwards inferences and focus the hearer’s attention on the compset. This observation finds additional support with discourse continuations without pro-forms coreferring with an antecedent in the previous clause. Consider (27a). In this example, the second segment provides a natural explanation for why the highest proportion of students failed the exam (refset); however, it does not provide a plausible reason for why just a few of them passed (compset). Conversely, the second segment in (27b) provides a natural explanation for why only a small proportion of athletes finished the marathon (refset) and the highest proportion of athletes (compset) were not able to complete it. These observations provide additional evidence in favor of the different available interpretations due to focal differences according to quantifier type.
Positive quantifiers such as ‘muchos’ appear to focus on the refset, being the compset very hardly accessible, or simply unavailable. On the other hand, negative quantifiers like ‘pocos’ make reference to both refset and compset possible, and to the same degree.

   ‘Many students failed. It was a really difficult exam.’

   b. Pocos atletas terminaron el maratón. Era un circuito realmente difícil.
   ‘Few athletes finished the marathon. It was a really difficult course.’

Based on these observations, I argue that the specific proportion denoted by a particular quantified expression is key to explain why only the so-called negative quantifiers allow for pronominal reference to the compset, while positive quantifiers disallow it. The assumption is that the refset is always available by virtue of being the refset (and default), an assumption widely acknowledged in the literature on (complement) set anaphora. However, with monotone decreasing quantifiers the refset and the highest proportion of A’s which are B’s do not coincide, hence compset and refset interpretations are on a par in terms of salience, with some possible preference for compset. This is the case of the quantified expressions in examples (28) through (30), where refset and compset do not coincide in terms of their proportions. In (28), for example, the refset may well consist of a set of 5 associates out of the 25 that should have attended the meeting, and the compset the set of 20 associates that did not attend, that is, the highest proportion out of the maximal set. The two interpretations for the quantified expression casi ningún(A)(B) in (29) can be interpreted in very similar terms.

(28) Muy pocos socios acudieron a la asamblea.
   ‘Very few associates attended the meeting.’

   **REFSET**: The proportion of associates that attended the meeting is small in relation to the proportion of associates that did not attend.
   Available as: **REFSET**

   **COMPSET**: The proportion of associates that did not attend the meeting is the highest-proportional reading.
   Available as: **HIGHEST PROPORTION**

(29) Casi ningún atleta participó en la maratón.
   ‘Almost no athletes took part in the marathon.’

   **REFSET**: The proportion of athletes that took part in the marathon is small in relation to the proportion of athletes that did not take part.
   Available as: **REFSET**

   **COMPSET**: The proportion of athletes that did not take part in the marathon is the highest-proportional reading.
   Available as: **HIGHEST PROPORTION**
Regarding the quantifier *ningún* in (30), I am giving here the empty set a proportional reading even though it may be argued that quantified expressions such as *ningún*(A)(B) only focus on the compset since the extension of refset is zero.

(30) Ningún estudiante acudió a la presentación.
   ‘No student attended the presentation.’

**REFSET**: The proportion of students that attended the presentation is zero.  
Available as: **REFSET**

**COMPSET**: The proportion of students that did not attend the presentation is the highest-proportional reading.  
Available as: **HIGHEST PROPORTION**

Unlike monotone decreasing quantifiers, in monotone increasing quantifiers both the refset and highest-proportional readings coincide, which makes only one reading salient and suitable to subsequent anaphoric reference. In (31), the refset for the quantified expression *todos los*(A)(B) is also the maximal set, and no compset interpretation is available (i.e. there is not a single student who did not attend the presentation).

(31) Todos los estudiantes acudieron a la presentación.
   ‘All students attended the presentation.’

**REFSET**: The proportion of students that attended the presentation is the highest-proportional reading.  
Available as: **REFSET**
Available as: **HIGHEST PROPORTION**

In (32), the refset could be interpreted as a subset of 60 athletes out of the 100 athletes that were supposed to run the marathon, which is the highest proportion, making it the only available interpretation suitable to be subsequently referred to anaphorically.

(32) Muchos atletas participaron en el maratón.
   ‘Many athletes took part in the marathon.’

**REFSET**: The proportion of athletes that took part in the marathon is higher in relation to the proportion of athletes that did not participate.  
Available as: **REFSET**
Available as: **HIGHEST PROPORTION**

As it was explained in Section 2, the quantifier *muchos* (like *pocos*) is an ambiguous proportional quantifier, highly context-dependent, whose denotation may vary depending on contextual circumstances. Under some specific contextual circumstances, it may be the case that the set denoted by a quantified expression with *muchos* is smaller than the complement set. For example, *muchos atletas terminaron la carrera* (‘many athletes completed the race’) can be true in a situation where only a small proportion of all the participants completed the race due to
extreme weather conditions. In that particular context, the set of athletes not completing the race would be the highest proportion. However, it seems that even in such a special context the quantified expression is to be understood in relation to some contextual parameter. In our marathon example, under extreme weather conditions, it is assumed that only a very small number of athletes would complete the race, hence ‘muchos’ is interpreted relative to this smaller set or, put it differently: under extreme weather conditions we would only expect that 10% of all participants would be able to complete the race, hence the most salient interpretation for ‘many’ is that the highest proportion of that 10% (say 8% or even a higher proportion) completed it. Even in contexts like this the compset interpretation is blocked.

In summary, my main assumption is that there is a default preference for the highest proportion with both monotone increasing and decreasing quantifiers. Assuming that refset interpretations are default regardless of monotonicity, the preference for the highest proportion (i.e., the largest set) denoted by the quantified expression explains why only monotone decreasing quantifiers yield subsequent anaphoric plural reference to the reference set \(A \cap B\) and to the complement set \(A \cap -B\), that is, the refset and the highest proportion have different extensions, and both sets are salient and available for anaphoric reference. The highest proportional set is the compset in monotone decreasing quantifiers. On the other hand, both the reference set and the highest proportional set are identical with monotone increasing quantifiers, hence only one set is salient and available for subsequent anaphoric plural reference. The cognitive reason why speakers tend to favor a highest proportional interpretation with monotone decreasing quantifiers is unclear to me, but this is a claim that should be tested experimentally in the future.

In the remainder of this paper, I suggest that discourse relations can provide some additional evidence in favor of or against compset interpretations. In terms of discourse coherence, two available reference set interpretations for a single quantified expression will result in two possible fully coherent continuations for the quantified segment. I will adopt here the approach to discourse developed in Segmented Discourse Representation Theory whereby rhetorical connections are inferred among utterances. My proposal is that discourse segments with monotone decreasing quantified expressions should allow for, at least, two possible rhetorical connections with an immediate discourse segment, namely, one possible rhetorical continuation for the refset and a different one for the compset interpretation. There should be no need for speakers to accommodate any of these rhetorical relations as both interpretations are salient. On the contrary, discourse segments with monotone increasing quantifiers should only allow for a plausible and salient discourse continuation because the quantified utterance provides a single refset interpretation. A second rhetorical connection, if possible at all, would need to be accommodated with great effort by the speaker.

**Discourse continuations and coherence**

Discourse structure and, more specifically, the rhetorical structure of discourse has been successfully applied to diverse anaphoric phenomena to date (see Alves, 2006; Zulaica-Hernández, 2009, 2012, *inter alia*). Based on this approach to discourse, this section provides additional evidence to complement anaphora interpretation in Spanish arising from the (im)possibility of inferring rhetorical relations in discourses with quantified discourse segments. This additional evidence does not rely much on the logical properties of quantifiers, but on how discourse contributes to facilitating specific referential readings for quantified segments.
In Segmented Discourse Representation Theory (Asher & Lascarides, 2003), discourse interpretation involves pragmatics, word meaning, and compositional semantics. The key hypothesis about the propositional content of discourse is that it crucially depends on a discourse structure consisting of rhetorical relations that link utterances together or, more accurately, the meanings or ‘contents’ that these utterances convey. A variety of rhetorical relations are proposed (i.e. Narration, Continuation, Explanation, Consequence, etc.) that describe the meaningful roles that propositions play in the global discourse content. A proposition can be the argument of several rhetorical relations since a given utterance can make more than one illocutionary contribution. In (33), for example, two relations (Narration and Contrast) can be inferred between the two clauses. The contrast relation is overtly indicated by the particle ‘but’, whereas the temporal information (Narration) must be inferred via world-knowledge and common-sense reasoning. The interpretation of discourse relations is encoded in rules of the form R(α, β) → Condition(α, β)

(33) John bought an apartment but he rented it out.

If we assume that some quantified expressions enable two possible anaphoric interpretations (refset and compset), then two different discourse continuations for a particular rhetorical relation should be available for that particular utterance. Let us consider (34). The utterances in this discourse are rhetorically connected via Elaboration (i.e., the second utterance [a or b] elaborates on the first quantified utterance: Elaboration(α, β).10 As expected, only one elaboration for the proposition with muchos is possible, that is, an elaboration for the salient refset interpretation as in (34a). However, an elaboration involving a compset interpretation is blocked in (34b). Furthermore, Elaboration is a subordinating relation in SDRT, which means that the contents associated with β presuppose and defeasible entail the contents associated with α (Asher & Vieu, 2005).11 Thus, it is reasonable to assume that getting passing marks in your courses presuppose having a good academic year, as in (34a), whereas a similar inference is blocked in (34b) because compset is not available.

(34) Muchos estudiantes de la clase han tenido un buen año académico.
   ‘Many students in the class had a great academic year.’

   a. pro han aprobado casi todos los cursos.  Elaboration(α, β), and β ⇒α
      ‘They passed almost every course.’

   b. # pro han reprobado casi todos los cursos.  Elaboration(α, β), and β ⇒α
      ‘They failed almost every course.’

---

10 There are different meaning postulates in SDRT for the Elaboration relation. For example, the events involved in an elaboration are connected via temporal inclusion, as shown in (a). Also, Elaboration is transitive, as shown in (b):
   a. φElaboration(α, β) ⇒ Part-of (φβ, eα)
   b. (Elaboration(π₁, π₂) ∧ Elaboration(π₂, π₃)) → Elaboration(π₁, π₃)

11 I want to thank an anonymous reviewer for pointing this out to me.
On the other hand, two elaborations are possible for the negatively quantified propositions in (35) and (36), each of these corresponding to the refset and compset, respectively. However, the degree of salience for the compset interpretation with monotone decreasing quantifiers varies across quantifiers, with pocos in (35b) being less felicitous than muy pocos, casi ningún, and ningún. Interestingly, the compset interpretation seems more salient with the latter group of quantifiers, those with minimal (or zero) refset proportions such as muy pocos, casi ningún and ningún, than with pocos. This may be an indication that the degree of salience is directly related with the size/proportion of the complement set as it was pointed out above. Therefore, there seems to be a directly proportional relationship between cardinality and salience in compset interpretations or, in other words, the larger the cardinality of the compset the more salient its interpretation.

(35) Pocos estudiantes de la clase han tenido un buen año académico.
‘Few students in the class had a great academic year.’

a. pro han aprobado casi todos los cursos.  
‘They passed almost every course.’

Elaboration(α, β), and β ⇒ α

b. ? pro han reprobado casi todos los cursos.  
‘They failed almost every course.’

Elaboration(α, β), and β ⇒ α

(36) Muy pocos/casi ningún/ningún estudiante(s) de la clase ha(n) tenido un buen año académico.
‘Very few/almost no/no student(s) in the class had a great academic year.’

a. pro han aprobado casi todos los cursos.  
‘They passed almost every course.’

Elaboration(α, β), and β ⇒ α

b. pro han reprobado casi todos los cursos.  
‘They failed almost every course.’

Elaboration(α, β), and β ⇒ α

The utterances in (37) are connected with the Explanation rhetorical relation, and the second segment(s) explains why the first proposition is true. The conjunction porque ‘because’ could be introduced to overtly mark such relation. Note that the compset interpretation (37a) is at least as salient as the refset interpretation in (37b). As regards the quantifiers pocos and muy pocos, further experimental scrutiny is needed in order to provide a plausible explanation for their ability to focus on compset. As with the relation Elaboration, the positive quantifier la mayoría ‘most’ allows a refset interpretation in (37b), whereas the connection via Explanation is blocked in (38a). Explanation, like Elaboration, is a subordinating relation in SDRT, and the continuation (37a) enables the inference that considering a candidate inexperienced non-monotonically entails not voting for him (i.e., the compset reading). Compare with (34b), where such inference is blocked with monotone increasing muchos, or with la mayoría de in (38a).

---

12 The temporal consequence of Explanation involves reversed causation of the two events [Explanation(α, β) → cause(eβ, eα)], which imposes reversed temporal precedence or simultaneity; i.e. φExplanation(α, β) ⇒ (∼ eα < eφ).
Also, in SDRT the discourse relation Result is the reversal of Explanation. If this idea is correct, we would expect that Result(β, α) is not possible in (38a), but fine in (38b). Interestingly, these observations are confirmed since we cannot infer that voting for a candidate is, under normal circumstances, the result of considering such candidate inappropriate or inexperienced (38a). However, the same inference is fine in (38b).13

\[(37)\] (Muy) pocos congresistas votaron al nuevo candidato.  
‘(Very) few congressmen voted for the new candidate.’

\[a. \quad pro\ \text{consideraron que era demasiado inexperto.}\] \hspace{1cm} Explanation(α, β), and β ⇒ α
‘They considered he was too inexperienced.’

\[b. \quad pro\ \text{consideraron que era el candidato ideal.}\] \hspace{1cm} Explanation(α, β), and β ⇒ α
‘They considered he was the ideal candidate.’

\[(38)\] La mayoría de los congresistas votaron al nuevo candidato.  
‘Most congressmen voted for the new candidate.’

\[a. \quad pro\ \text{consideraron que era demasiado inexperto.}\] \hspace{1cm} Explanation(α, β), and β ≠ α
‘They considered he was too inexperienced.’

\[b. \quad pro\ \text{consideraron que era el candidato ideal.}\] \hspace{1cm} Explanation(α, β), and β ⇒ α
‘They considered he was the ideal candidate.’

Finally, the Contrast relation is tested with negative and positive quantifiers in (39)-(40). These examples have been simplified to facilitate processing: past tenses have been replaced with present tenses, and the particle pero ‘but’ has been introduced in order to force an explicit contrast between the utterances. The role of this particle is that of indicating a violation/denial of an expectation. Though difficult to process in every case, it seems that both refset and compset are equally salient with monotone decreasing quantifiers when a contrast is involved. After processing (39), our refset expectation is that only a minimal proportion of congressmen should think of Kennedy as the ideal candidate, hence the continuation in (39a) does not make sense as a contrast to (39) as it does not contradict such expectation. On the contrary, our compset expectation is that most congressmen should not think of him as the ideal candidate, and (39a) does precisely contradict such expectation. Interestingly, the inference pattern is reversed with (39b), and the contrast can only deny an expectation associated with refset.

\[(39)\] (Muy) pocos congresistas admiran a Kennedy...  
‘(Very) few congressmen admire Kennedy...’

\[a. \quad \text{pero} \quad pro\ \text{creen que es el candidato ideal.}\] \hspace{1cm} Contrast(α, β) ok with Compset

\[\text{but} \quad \text{they think he is the ideal candidate.’}\]

13 The relation Contrast entails the truth of the contents associated with their terms, but not a presupposition relation (Asher & Vieu, 2005, p. 598).
b. *pero pro no* creen que sea el candidato ideal. *Contrast(α, β) ok with Refset*

but they don’t think he is the ideal candidate.’

Again, as we would expect, compset is not available with the monotone increasing quantifier *muchos* ‘many’, hence only one contrast relation is available. The negated version (40b) is necessary so that the discourse is felicitous, and the contrast can be inferred among the two propositions.

(40) Muchos congresistas admiran a Kennedy...
‘Many congressmen admire Kennedy...

a. *pero pro creen que es el candidato ideal.  
...but they think he is the ideal candidate.’

b. *pero pro no creen que sea el candidato ideal.  
...but they don’t think he is the ideal candidate.’

Conclusions

This paper addresses an understudied topic within Spanish semantics. The main goal was to establish a first approach to complement set anaphora by investigating how similar or dissimilarly Spanish quantified expressions behave referentially compared to English. We have seen that Spanish appears to conform with the general pattern observed according to which only monotone decreasing quantifiers allow for a complement set interpretation, and that anaphoric reference to the complement set is, therefore, only possible within the context of monotone decreasing quantifiers. However, there are many issues in connection with this topic that need to be further investigated before any definitive conclusions can be drawn. First and foremost, I believe experimental evidence is needed before the availability of compset readings with negative quantifiers can be fully confirmed. Thus, it is necessary to know whether the compset is naturally available to readers/listeners for subsequent pronominal reference, and whether the compset is less, more, or as salient as the refset in negatively quantified utterances. Assuming that compset is available, it would also be interesting to test whether different monotone decreasing quantifiers allow for compset interpretations to the same degree. In this paper, I argued that proportions may be key in explaining why only negative quantifiers, but not monotone increasing ones, allow for complement anaphora interpretations. I also argued that as the size/proportion of compset increases the anaphoric reading becomes more salient and felicitous. But still, more evidence is needed before this claim can be confirmed. I think that the relationship between salience and size of the complement set can provide interesting insights to our current understanding of compset anaphora and to our general understanding of referent salience and accessibility.

Finally, I think it is worth exploring the role of discourse in relation with the refset/compset anaphora distinction. In my view, the logical properties of quantifiers alone will not suffice to provide a plausible explanation as to how language speakers may process different referential interpretations (refset, compset, maxset) with some quantified expressions, but not with others, or to how the alleged ambiguity is resolved. In other words, how proforms are
resolved, and why refsets are preferred over compsets, or vice versa. The role of lexical items such as particles, certain types of verbs (i.e., causatives), and, generally speaking, inferred discourse relations among discourse utterances could eventually be the primary source of knowledge needed for hearers to resolve the potential ambiguity.

The author declares that he has no conflict of interest.

References


