



**HAL**  
open science

## Verbal agreement with partitive noun phrases

Benoît Leclercq, Ilse Depraetere

► **To cite this version:**

Benoît Leclercq, Ilse Depraetere. Verbal agreement with partitive noun phrases. *Studia Linguistica*, 2018, 72 (2), pp.340-361. 10.1111/stul.12059 . hal-01614452

**HAL Id: hal-01614452**

**<https://hal.univ-lille.fr/hal-01614452>**

Submitted on 4 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

"This is the peer reviewed version of the following article: [Leclercq, B., & I. Depraetere. (2018). Verbal agreement with partitive noun phrases. In *Studia Linguistica* 72(2): 340-361.], which has been published in final form at [<https://doi.org/10.1111/stul.12059>]. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions."

## Verbal agreement with partitive noun phrases<sup>1</sup>

Benoît Leclercq  
University of Paris 8

Ilse Depraetere  
University of Lille

### 1. Background and aim

This paper offers empirical evidence for the claim that complex NPs that express partitive meaning, as in the sentence in (1), instantiate the structure *QP (Quantifier Phrase) of N* rather than *N of NP*, which is the structure of the non-partitive complex NP illustrated in (2).

- (1) *A majority of these studies* have (has) assessed individual differences. (COCA (Corpus of Contemporary American English), written)  
 (2) *The accuracy of the results* was (were) being questioned. (COCA, written)

Partitives 'have a characteristic interpretation that implies an idea of partition, which can be of two types: they can denote a subset of a set or a part of a whole' (Martí Girbau 2010:21). In other words, a partitive NP does not establish reference to a set (quantity) of individuals or things (example 3) but it takes a quantity out of an already existing group of individuals or things (example 4).

- (3) A dozen of publications.  
 (4) Two-thirds of the publications.

We will use the terms *partitive NP* and *partitive construction* interchangeably.

### 2. The partitive construction: two hypotheses.

Nominal constructions with a prepositional *of*- complement are generally associated with the binominal *N of NP* structure. Given its structural make-up, the partitive is usually, and rather uncritically, analysed as a binominal construction (cf. Jackendoff 1972; Quirk et al. 1985; Cobuild 1990; Swan 1995; Abbott 1996; Hoeksema 1996; Biber et al. 1999; Corbett 2000, 2006; Depraetere and Langford 2012). It is the semantics of the construction and the formal realization of the embedded NP that have received a lot of attention rather than the structural make-up of the complex NP in itself, the binominal analysis being rather taken for granted. However, the binominal analysis of the partitive has been challenged by Martí Girbau (2003, 2010) who argues that partitives might in fact instantiate a different structure, namely *QP of N*, which differs from the binominal analysis both in terms of what the head of the construction is and in terms of the internal structure of the construction. The aim of this paper is to check which of the two structural analyses is borne out by empirical evidence. The binominal analysis of the construction will be referred to as *hypothesis 1*, and Martí Girbau's *QP of N* analysis will be referred to as *hypothesis 2*.

<sup>1</sup> We are grateful to the reviewers for their critical observations and suggestions for improvement.

The evidence for the ‘QP of N’ structure hinges on facts about verbal agreement. Verbal agreement is the relation between a verb and its Subject that is determined by the feature of number (or person, when the verb *be* is used). More specifically, when the Subject is realised by an NP, it is (the number of) the head that determines the number of the verb. In sentences with complex NPs, like that in (5), the question is how to determine whether it is *daughters* or *butcher* that functions as head of the NP.

(5) *The daughters of the butcher* are coming to the party.

Haegeman & Guéron (1999:66) point out that ‘the N *daughters* determines the properties of the full constituent’. This explains why only the plural pronoun *they* (and not *he*) can substitute for the entire NP. (test 1) Moreover, the verb must be in the plural, a feature which it shares with *daughters* rather than *butcher*. (test 2)

(6) *They* are coming to the party.

(7) \**He* are coming to the party.

(8) \**The daughters of the butcher* *is* coming to the party.

These observations lead Haegeman and Guéron to conclude that the head of this complex NP is *daughters* (N<sub>1</sub>) rather than *butcher*. Let us apply Haegeman and Guéron’s tests to some examples with partitives.

(9) A majority of them was in complete agreement over the restrictions on car. (BNC (British National Corpus), written)

(10) A majority of our members wear shoes with Velcro rather than shoelaces. (COCA, spoken)

(11) A number of us was asked to comment. (BNC, spoken)

(12) A number of the regulations have given rise to concern. (BNC, written)

The pronominalization test (test 1) points in the direction of N<sub>2</sub> being the head of the NP. In all the examples given below, plural pronouns have to be used, which seems to suggest that it is N<sub>2</sub> which ‘determines the properties of the full constituent’.

(13) A majority (sg) of them (pl) was (sg) in complete agreement over the restrictions on car. (BNC, written) – They were / \* It was in complete agreement over the restrictions on car.

(14) A majority (sg) of our members (pl) wear shoes with Velcro rather than shoelaces. (COCA, spoken) - They wear / \* It wears shoes with Velcro rather than shoelaces.

(15) A number (sg) of us (pl) was (sg) asked to comment. (BNC, spoken) – \* It was asked / We were asked to comment.

(16) A number (sg) of the regulations (pl) have (pl) given rise to concern. (BNC, written) – \* It has / They have given rise to concern.

However, the results of the second test are inconclusive. Replacing a singular verb by a plural verb (or vice versa) does not result in ungrammaticality. A partitive NP (with conflicting numbers) featuring in Subject position can, in principle, combine with a singular as well as a plural form of the verb, as is also clear from the example in (1). Various proposals have been made to explain agreement, and they have taken into account, apart from the structural make-up of the NP, the impact of semantics (notional agreement) as well as formal factors such as proximity (proximal agreement).

The data from the BNC and COCA presented in this paper show that in the majority of cases, the verb agrees with the number of the second N in the complex NP rather than with the first. We will argue that these facts can be explained in the most elegant and economic way when the partitive construction is analysed in terms of a quantifier phrase followed by a nominal head rather than in terms of a binominal structure with a noun that is complemented by a post-nominal NP. In this way, we will provide solid empirical evidence for the structural analysis of partitive constructions put forward by Martí Girbau (2003, 2010).

### 3. Empirical analysis

In this section, we will describe and analyse the agreement patterns with partitive constructions in Subject position in a large sample from the BNC and COCA. These corpora give us an insight into potential regional variation, and they also allow for a comparison between verbal agreement in the spoken and the written registers.

#### 3.1. Hypothesis 1 (N of NP) and hypothesis 2 (QP of N)

The findings will be assessed in terms of the two hypotheses referred to in the introduction: the binominal analysis of the partitive construction (N of NP), whereby the first noun is the head of the complex NP (*hypothesis 1*); the analysis in terms of ‘QP of N’, whereby the second noun is the head of the complex NP (*hypothesis 2*). To facilitate reference to the nouns that appear in the complex NP, we will use (hypothesis-neutral)  $N_1$  to refer to the noun that is mentioned first and (hypothesis-neutral)  $N_2$  to the noun that occurs after  $N_1$ . This convention is hypothesis-neutral in the sense that it does not assign a specific function to the nouns in question; it does not imply a specific stance concerning the structural make-up of the complex NP. Non-indexed N refers to the head of the complex NP. According to hypothesis 1, the verb should agree with  $N_1$ ; according to hypothesis 2, the verb agrees with  $N_2$ .

Anticipating the discussion somewhat, the following examples illustrate the challenges that the corpus data imply for either hypothesis.

- (17) Two-thirds (pl) of the tale (sg) *was* (sg) devoted to how the pig played mind games with the wolf. (COCA, written)
- (18) Two-thirds (pl) of the island (sg) *is* (sg) national park. (COCA, written)
- (19) Two-thirds (pl) of the entire territory (sg) *are* (pl) grounded on ‘permafrost’ – permanent frozen earth. (BNC, written)
- (20) Three-quarters (pl) of the book (sg) *are* (pl) my pictures. (COCA, spoken)
- (21) A part (sg) of them (pl) *is* (sg) kind of nasty. (COCA, written)
- (22) A majority (sg) of the members (pl) *votes* (sg) for the nominee. (COCA, written)
- (23) A couple (sg) of us (pl) from work *are* (pl) in the 90 Minutes Magazine one. (BNC, written)
- (24) A number (sg) of the regulations (pl) *have* (pl) given rise to concern. (BNC, written)

In sentences (19), (20), (21) and (22), hypothesis 1 makes the right predictions: the verb indeed agrees with the number of  $N_1$  (irrespective of the number of  $N_2$ ). Exceptions to the rule, illustrated by (17) to (18) and (23) to (24) have been described in terms of *proximal agreement* (agreement with the noun closest to the verb) and *notional agreement* (agreement with the notion of number conveyed by the entire NP). According to the view that we adhere to, hypothesis 2, the verb is bound to agree with  $N_2$ : irrespective of the number of  $N_1$ , verbal agreement is always controlled by  $N_2$ . On this approach, it is the examples in (19) to (22) that do not follow the general rule and that need to be accounted for.

### 3.2. Selection of sample: methodology

In order to test the different hypotheses, a dataset was extracted from two corpora: the BNC and COCA (<http://corpus.byu.edu/overview.asp>). The BNC is a corpus of a 100 million words of spoken (10%) and written (90%) British English collected between 1980 and 1993. COCA is composed of 520 million words of spoken and written American English, collected between 1990 and 2015 (Davis 2004, 2008). First, a list of partitive constructions was compiled on the basis of the descriptions in major reference grammars. (Quirk *et al* 1985:257, Cobuild 1990:110, Biber *et al.* 1999:185, Huddleston and Pullum *et al.* 2002:350)<sup>2</sup> The string consisting of the first three words of the examples was extracted from both corpora in order to identify the most frequent ones. The first column in table 1 lists the sixteen strings selected for the analysis. It is important to add that the three-word strings do not always appear in a partitive construction: for the construction to be partitive, N<sub>2</sub> must be semantically definite.<sup>3</sup> The preceding observations imply that not any token of one of the sixteen strings was included; certain examples were excluded for reasons that have to do with referent (semantically definite or not) of N<sub>2</sub>. (column 3) Moreover, given the aims of our study, a further requirement is that the partitive construction has to be used in subject position and it necessarily has to be followed by a finite verb form inflected for number. This requirement resulted in the exclusion of a further set of examples (column 4). This selection also excluded sentences like (25) in which the partitive is conjoined with an additional NP.

(25) Later that year, *he and a number of his associates* were jailed. (COCA, written)

The sorting procedure described resulted in a sample of 10,122 sentences:

<sup>2</sup> The strings arrived at in this way resulted in less frequent N<sub>2</sub>s being excluded from the list, which, with the exception of *bunch*, and possibly *couple*, *group* and *part*, consists of abstract nouns. Further empirical research will be required to study the impact of the abstract/concrete nature of N<sub>1</sub> on the number of the verb.

<sup>3</sup> It is usually argued that not any kind of NP can fill the N<sub>2</sub>P slot in a partitive construction: partitives are subject to the structural restriction that N<sub>2</sub>P is definite. (cf. Jackendoff 1977, Selkirk 1977, Bock & Miller 1991:7). In other words, while *a couple of my friends* expresses partition, *a couple of friends* does not. However, as Abbott (1996) convincingly argues, the partitive constraint 'does not mean that the containing group has to be one with which the addressee is assumed to be already familiar (i.e. definite, in one sense of 'definite')' (1996: 41). Rather, the group referred to by N<sub>2</sub>P 'must be relatable to the context in some way' (ibid.); and this does not necessarily require N<sub>2</sub> to be formally definite. For instance, in *40% of children do not have access to education*, the context establishes reference to the particular group from which partition is possible. Abbott does not believe there is any constraint on the partitive construction. However, she only provides evidence for examples with numerals, quantifiers, fractions and percentages. Accordingly, in the sample discussed here, numerals, quantifiers, fractions and percentages followed by a formally indefinite but semantically definite N<sub>2</sub> were also included.

strings (QP of)	in the corpora overall	in a partitive construction	relevant data
A number of *	50,332	3,231	834
A couple of *	45,997	2,282	298
A portion of *	2,962	2,289	273
A majority of *	4,308	1,970	423
A part of *	15,435	10,986	287
A group of *	20,024	624	59
A bunch of *	9,271	596	86
A dozen of *	646	599	83
A percentage of *	1,608	797	28
A fraction of *	2,732	1,684	87
A quarter of *	5,197	5,197	738
A third of *	5,517	5,517	1,364
80 % of *	4,623	4,623	1,512
90 % of *	4,003	4,003	1,467
Three-quarters of *	3,012	3,012	701
Two-thirds of *	5,871	5,871	1,882
<b>TOTAL</b>	<b>180,538</b>	<b>53,281</b>	<b>10,122</b>

Table 1: Frequency of the selected strings

All the examples were annotated in terms of 6 different criteria: register (spoken or written), regional variety (American or British English), lexical realization of  $N_1$  (*a couple of*, *a fraction of*, *three-quarters of*, ...), number of  $N_1$  (singular as in *a group of* or plural as in *three-quarters of*)<sup>4</sup>, nature and number of  $N_2$  (singular count noun, plural count noun, mass noun or collective noun) and the number of the verb (singular or plural).

In order to identify significant patterns in the dataset, the distribution of forms (singular/plural  $N_1$ , singular/plural  $N_2$ , singular/plural verb) was submitted to chi-square tests, the results of which will be discussed in the next section.

### 3.3. Results

#### 3.3.1. Interpretation of the data in terms of N of NP (hypothesis 1).

Hypothesis 1 predicts that verbs agree in number with  $N_1$ . Figure 1 (percentages) and table 2 (absolute figures) give an overview of the frequency of singular and plural verbs with a singular and a plural  $N_1$ . A singular  $N_1$  combines more than twice as often with a plural verb than with a singular verb: out of 4,560 examples with a singular  $N_1$ , 3,313 (72.66%) combine with a plural verb and 1,247 (27.34%) combine with a singular verb. When  $N_1$  is plural (2,583), there are more plural verbs (2,022 - 78.3%) than singular verbs (561 - 21.7%).

<sup>4</sup> Percentages (here 80% and 90%) were not annotated *singular* or *plural* but *number neutral* in the sense that they are inherently not associated with a distinct number. The data analysis reveals that 96.6 % of partitives that include percentages agree in number with  $N_2$ . When collectives (cf. p.5) in  $N_2$  position are not taken into account, then the figure is even higher, that is, 99.7%.

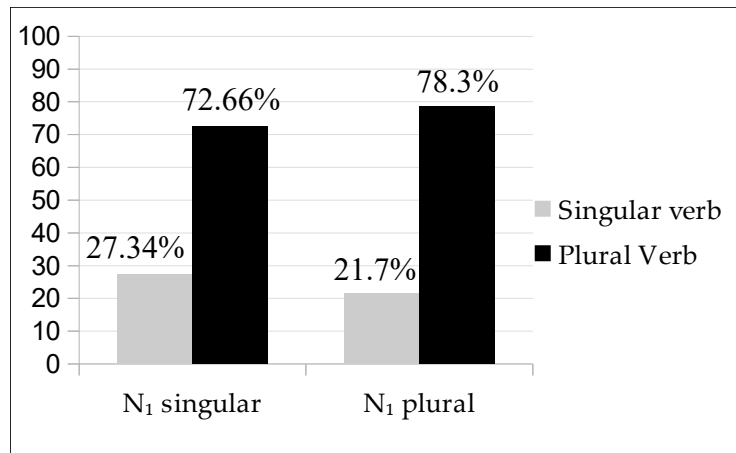


Figure 1: Verbal agreement and number – N of NP

Table 2 also shows that overall, only 3,269 (45.76%) sentences confirm the predictions made by hypothesis 1 (when N<sub>1</sub> is singular, the verb is in the singular and when N<sub>1</sub> is plural, the verb is in the plural) while 3,874 sentences (54.24%) do not.<sup>5</sup> That is, more than half of the sentences show unexpected patterns of agreement.

	Predicted	¬ Predicted	Total
<b>N<sub>1</sub> singular</b>	1,247	3,313	4,560
<b>N<sub>1</sub> plural</b>	2,022	561	2,583
<b>Total</b>	3,269	3,874	7,143

Table 2: Frequency of verbs that meet the predictions of hypothesis 1.

We first checked whether the high number of exceptions to hypothesis 1 could be explained in terms of regional differences or differences in register. Table 3 shows that there is no significant difference between American English and British English or between spoken and written data in terms of the impact of the variety on the number of verb.<sup>6</sup>

	Regional difference			Register		
	Coca	BNC	Total	Spoken	Written	Total
<b>Predicted</b>	2,717 (exp: 2,715.24)	552 (exp: 553.76)	3,269	825 (exp: 847.57)	2,444 (exp: 2,421.43)	3,269
<b>¬ Predicted</b>	3,216 (exp: 3,217.76)	658 (exp: 656.24)	3,874	1,027 (exp: 1,004.43)	2,847 (exp: 2,869.57)	3,874
<b>Total</b>	5,933	1,210	7,143	1,852	5,291	7,143
	X <sup>2</sup> = 0.0124      P value = 0.911397			X <sup>2</sup> = 1.496      P value = 0.22129		

Table 3: Regional difference and register (hypothesis 1).

The differences in table 3 are not significant since the p-value is higher than 0.05. That is, the probability that the observed difference (between American English and British English, or between spoken and written English) is due to chance is higher than 5%.<sup>7</sup>

<sup>5</sup> Since percentages have been analysed as number neutral (cf. footnote 3), the figures concerning the number of N<sub>1</sub> do not include percentages. In other words, the total number of sentences in Table 2 is not 10,122 but 7,143 because of the exclusion of the percentages (80% and 90%).

<sup>6</sup> The observed difference is statistically significant when  $p < 0.05$ .

<sup>7</sup> Whether a difference is significant can be determined by comparing the observed frequency of a pattern with the expected frequency. The expected frequency of a pattern is calculated on the basis of the sum of the observed

In the literature, the concept of *notional agreement* is sometimes used to explain examples in which a plural verb is used when the form of the noun predicts a singular verb. (cf. Biber *et al*, 1999:187, Reid, 1991:197) Quirk *et al* provide the following definition: ‘Notional concord is agreement of verb with subject according to the notion of number rather than with the actual presence of the grammatical marker for that notion’. (Quirk *et al*, 1985:757)

The line of reasoning for partitives could be as follows: even though the form of  $N_1$  is singular, from a grammatical point of view, there is reference to a set of things or a group of people and this is what determines the number of the verb. An explanation along these lines raises several problems though: first, it is not clear why ‘plural semantics’ determine the choice of verb in some cases but not in all partitives with a singular  $N_1$  or why they do not result in the use of a plural verb with all partitives altogether. Secondly, notional agreement does not account for all exceptions. The following table gives a detailed overview of the distribution of singular and plural verbs:

	Singular verb	Plural verb
$N_1$ sg of $N_2$ sg	90% (1208)	10% (133)
$N_1$ sg of $N_2$ pl	1.2% (39)	98.8% (3180)
$N_1$ pl of $N_2$ pl	0.11% (2)	99.89% (1867)
$N_1$ pl of $N_2$ sg	78.29% (559)	21.71% (155)

Table 4: detail structural make-up of partitives and distribution of verb forms.

Notional agreement can explain the cases that have been shaded grey (a subsample that constitutes the majority of cases when  $N_1$  is singular). However, it does not solve all problems as it does not account for the unexpected cases with a plural  $N_1$  (in bold). Thirdly, the operationalization of the concept of notional agreement is not straightforward. For instance, what evidence is there to say that notional agreement applies in (26)? On such an approach *my pieces* might be argued to influence the conceptualization of the number of *a couple*. However, could it not be the case, rather, that the plural form, *my pieces*, is conclusive for the choice of number? Moreover, if notional agreement applies to (26), why does it not have a similar effect in (27)?

(26) A couple (sg) of my pieces (pl) have (pl) made it to dress rehearsal. (COCA, spoken)

(27) A couple (sg) of my children (pl) was (sg) there. (COCA, spoken)

Proximal agreement offers a potential alternative explanation of the exceptional cases. Unlike notional concord, it is easy to operationalize: in this case, the choice of a plural (singular) verb follows from the proximity of plural (singular) noun. (Quirk *et al* 1985:757, Depraetere & Langford 2012:133) As there is an intervening  $N_2$  between  $N_1$  and the verb, agreement with the distant  $N_1$  is cognitively too demanding, and therefore, the number of the verb is adapted to that of the noun that is closest to the verb ( $N_2$ ): ‘Such clear deviations from grammatical concord are mostly found in speech, where they are explicable from the psycholinguistic constraints of a limited short-term memory and the pressure of online construction of linguistic input.’ (Biber *et al* 1999:189) Table 5 shows that when a singular  $N_1$  is followed by plural  $N_2$ , there are significantly more plural verbs than expected. Proximal agreement makes similarly strong predictions when  $N_1$  is plural. When a plural  $N_1$  is followed by singular  $N_2$ , it takes significantly more often a singular verb than expected. In other words,

---

frequencies. For instance, in table 3, 656.24 sentences were expected not to follow hypothesis 1 based on the cross calculation that, from the 7,143 sentences observed, 1,210 sentences come from the BNC and 3,874 sentences do not follow Hypothesis 1.



it seems that *proximal agreement* is more powerful as an explanatory principle than main hypothesis 1 as such. If we assume that proximal agreement is the major principle of agreement, this implies that the exception has become the general rule.

	N <sub>1</sub> singular			N <sub>1</sub> plural		
	Predicted	¬ Predicted	Total	Predicted	¬ Predicted	Total
<b>SG N<sub>2</sub> P</b>	1,208 (exp: 366.72)	133 (exp: 974.28)	1,341	155 (exp: 558.93)	559 (exp: 155.07)	714
<b>PL N<sub>2</sub> P</b>	39 (exp: 880.28)	3,180 (exp: 2,338.72)	3,219	1,867 (exp: 1,463.07)	2 (exp: 405.93)	1,869
<b>Total</b>	1,247	3,313	4,560	2,022	561	2,583
	X <sup>2</sup> = 3,763.0647      P value = 0			X <sup>2</sup> = 1,857.4926      P value = 0		

Table 5: Proximal agreement.

### 3.3.2. Interpretation of the data in terms of QP of N (Hypothesis 2)

The observations about proximal agreement constitute a nice transition to hypothesis 2, which predicts that the verb agrees with the only noun in the NP, that is, N<sub>2</sub>. Table 5 shows a distinct correlation: plural count N<sub>2</sub> mostly combine with plural verbs. Singular count N<sub>2</sub> and mass N<sub>2</sub> mostly combine with singular verbs. The category of collective nouns is the only one that does not show such a distinct pattern, even though, again, more singular verbs have been found. However, as will be shown below, the use of a plural verb with a collective can be explained in terms of the semantics of the VP or in terms of the lexical realization of the NP.

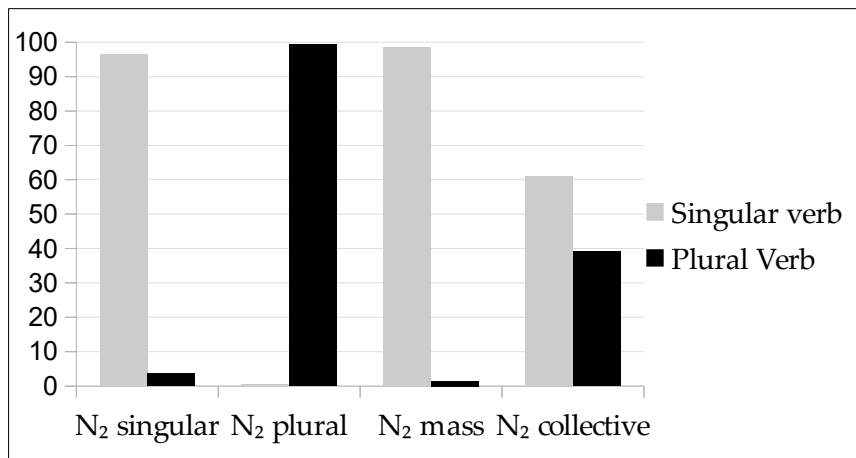


Figure 2: Verbal agreement and number - QP of part N

Figure 2 and table 6 show that hypothesis 2 is borne out by 95.76% of the sentences (9,693); only 4.24% of the sentences (429) exhibit agreement patterns that are not predicted by hypothesis 2. These figures are crucial when it comes to assessing the explanatory power of both hypotheses: the proportion of sentences that is not in line with the predictions made by hypothesis 2 is small compared to the set of examples that deviate from the patterns predicted by hypothesis 1 (54.24%).

	Predicted	¬ Predicted	Total
<b>N2 singular</b>	2,544	384	2,928
<b>N2 plural</b>	7,149	45	7,194
<b>Total</b>	9,693	429	10,122

Table 6: Frequency of verbs that meet the predictions of hypothesis 2 – QP of N

This observation is confirmed by the evidence in table 7. Significantly more sentences meet the predictions made by hypothesis 2.

	Hypothesis 1	Hypothesis 2	Total
<b>Meet the predictions</b>	3,269 (exp: 5,362.73)	9,693 (exp: 7,599.27)	12,962
<b>¬ Meet the predictions</b>	3,874 (exp: 1,780.27)	429 (exp: 2,522.73)	4,305
<b>Total</b>	7,143	10,122	17,269
$X^2 = 5,594.3745$		P value = 0	

Table 7: Opposition in the frequency of verbs that meet the predictions of the two hypotheses.

So far, it has been established that the explanatory power of hypothesis 2 is more convincing than that of hypothesis 1. In the case of hypothesis 1, more than half of the examples with a singular N<sub>1</sub> cannot be explained in terms of the principal claim. In other words, 54.24 % of the instances have to be accounted for in terms of proximal agreement and/or notional agreement.<sup>8</sup> It will be clear though, that it is a more elegant solution to have one hypothesis which makes correct predictions in the majority of cases, rather than one which explains less than half of the data, an exceptional rule (proximal agreement and/or notional agreement) being needed to account for majority of the data.<sup>9</sup> Moreover, as pointed out before, the concept of ‘notional agreement’ appears a somewhat ‘after the fact’ solution that is hard to operationalize, and that potentially implies a danger of circularity.

Having shown the limits of hypothesis 1, it remains a fact that 429 sentences do not meet the predictions made by hypothesis 2. We will now check whether these can be accounted for in a straightforward way. First, as table 8 shows, there are significantly more unexpected patterns of agreement in British English than in American English and more in the spoken register than in the written register.

	Regional difference			Register							
	COCA	BNC	Total	Spoken	Written	Total					
<b>Predicted</b>	8,373 (exp: 8,344.68)	1,320 (exp: 1,348.32)	9,693	2,355 (exp: 2,439.05)	7,338 (exp: 7,253.95)	9,693					
<b>¬ Predicted</b>	341 (exp: 369.32)	88 (exp: 59.68)	429	192 (exp: 107.95)	237 (exp: 321.05)	429					
<b>Total</b>	8,714	1,408	10,122	2,547	7,575	10,122					
$X^2 = 16.3079$			P value = 0.000054			$X^2 = 91.3176$			P value = 0		

Table 8: Regional difference and register (hypothesis 2).

<sup>8</sup> Biber *et al* (1999:190) write that *notional concord* and the *principle of proximity* often work together.

<sup>9</sup> Notional and proximal agreement can account the majority of the exceptions, and leave only 0.11% of the examples (N<sub>1</sub>pl of N<sub>2</sub>pl) unexplained. (cf. table 4)

Register differences seem to be entirely due to the presence of collective nouns. When the collectives are taken out of the analysis, spoken and written registers do not show significant differences in terms of predicted and non-predicted forms (table 9).

	<b>Predicted</b>	<b>¬ Predicted</b>	<b>Total</b>
<b>Spoken</b>	2,187 (exp: 2,190.37)	27 (exp: 23.63)	2,214
<b>Written</b>	6,989 (exp: 6,985.63)	72 (exp: 75.37)	7,061
<b>Total</b>	9,176	99	9,275
$\chi^2 = 0.6373$		P value = 0.424673	

Table 9: Difference between spoken and written data without collectives (hypothesis 2)

A similar observation applies to regional differences: when collectives (in N<sub>2</sub> position) are taken out of the analysis, the difference between AmE and BrE in terms of the number of predicted forms and non-predicted forms is not significant. (p value > 0.05)

	<b>Predicted</b>	<b>¬ Predicted</b>	<b>Total</b>
<b>COCA</b>	7889 (exp: 7,882.95)	79 (exp: 85.05)	7,968
<b>BNC</b>	1287 (exp: 1,293.05)	20 (exp: 13.95)	1,307
<b>Total</b>	9,176	99	9,275
$\chi^2 = 3.0863$		P value = 0.078955	

Table 10: Difference between AmE and BrE data without collectives (hypothesis 2)

Table 11, which looks at collectives *only*, reveals the following difference: AmE uses significantly more singular verb forms with collectives than BrE (p-value < 0.05). Therefore, it confirms the previously observed tendency that collectives in AmE combine more often with a singular than a plural verb. (cf. Levin 2001: 60-70, Depraetere 2003: 103)

	<b>Predicted (singular)</b>	<b>¬ Predicted (plural)</b>	<b>Total</b>
<b>COCA</b>	484 (exp: 455.35)	262 (exp: 290.65)	746
<b>BNC</b>	33 (exp: 61.65)	68 (exp: 39.35)	101
<b>Total</b>	517	330	847
$\chi^2 = 38.7985$		P value = 0	

Table 11: Difference between AmE and BrE data collectives only (hypothesis 2)

In 330 (76.9%) out of the 429 sentences that do not meet the predictions of hypothesis 2, N<sub>2</sub> is a collective noun. Even though in principle collective nouns can agree either to the singular or the plural, given that it is formal number that is the determining principle in hypothesis 2, a singular verb is predicted to be used.<sup>10</sup> A more detailed analysis of the subset

<sup>10</sup> It seems useful to add the following observation about the nature of partitives at this stage: in most accounts of partitives, agreement issues are illustrated with examples in which N<sub>1</sub> is singular and N<sub>2</sub> plural. From one point of view, plurality always seems to be part of the partitive construction as one needs a group of more than one individual in order for partition to be possible. However, as de Hoop (1997:154) nicely explains, partitives call for a group reading of N<sub>2</sub> ('the embedded NP must have a group reading' (de Hoop, 1997:154)) rather than a distributive reading. It is from the entire group that partition is possible, not from individuals themselves. That is why example (i) is grammatical and (ii) is not (de Hoop, 1997:154).

of exceptions to hypothesis 2 in which  $N_2$  is a collective noun shows that there are no genuine exceptions, as the agreement observed is in line with the agreement patterns typical of collective nouns. The following cases can be distinguished:

First, it has been established in previous research (Levin 2001, Depraetere 2003, 2004) that a number of collective nouns, among which *people*, *staff*, *crew*, *cast* preferably combine with a plural verb. Our sample provides further evidence for this observation. This category also includes examples with de-adjectival nouns and nouns that do not have a regular plural form (such as *aircraft*, *youth* or *data*) and that are followed by a plural verb form.<sup>11</sup> 137 examples out of the subset with collectives can be explained along these lines.

Moreover, the semantics of the verb phrase are sometimes such that the use of a plural verb is the only alternative or the most plausible alternative. It is indeed possible to classify the exceptions as semantic groups, as the overview below shows. These patterns constitute hypotheses that it will be interesting to check in further data sets:

(a) The VP contains a reflexive pronoun or a plural NP which signals a one-to-one relationship between the referents of the individuals that make up the subject NP and the referents of the complement NP; there is a distributional reading:

- (28) Two-thirds of the sample *describe themselves* as head of their households. (BNC, written) (*describe oneself*)
- (29) That is, less than a third of the clergy *are members of* the DUP. (BNC, written) (*be a member*)
- (30) I said a majority of the black community *have disciplined their children*. (COCA, spoken) (*discipline one's children*)
- (31) Because the solid waste infrastructure is inadequate, over 80 percent of the population *do not have home collection services*. (COCA, written) (every member of the population does not have access to home collection services in his or her neighbourhood).

(b) The VP assigns a property that is typically predicated of individuals:

- (32) Two-thirds of this group are minorities, and two-thirds of their alleged victims *were white*. (COCA, written) (*be white*)
- (33) On the other hand, since nearly 80 per cent of the nation *were illiterate* in early NEP. (BNC, written) (*be illiterate*)
- (34) The problem is, is that two-thirds of the country *are opposed to him* on this. (COCA, spoken) (*be opposed to*)

(c) The VP contains a verb that refers to actions carried out by individuals or situations that foreground individual action; the use of a plural form falls out of the presence of such a verb.

- (35) A number of our Methodist Women's Fellowship *were gathered* in our church for a weekly prayer. (BNC, written)

- 
- (i) One of the two cats
  - (ii) \*One of both cats

If partitives indeed favour a group reading, then collectives are expected to agree to the singular when used in the partitive construction, since the focus will be on the unit (or 'shell' (Depraetere 2003)) rather than on the individuals that constitute the unit.

<sup>11</sup> Note that agreement with *staff* is variable across regions. While the plural is the preferred option in BrE, in AmE *staff* readily combines with a singular too. (cf. Depraetere 2004)

- (36) Three quarters of the group *were killed*. (BNC, written)  
 (37) Two-thirds of the laity *live scattered across* the diocese. The remaining third live in Poole. (BNC, written)

(d) The VP contains a verb of cognition, and as a result the individuals that make up the collective are in the foreground; the use of a plural verb form follows naturally.

- (38) Okay, two-thirds of my class *don't know* how to write a comparative essay. (COCA, written)  
 (39) So it's not surprising that, down to the present day, fundamentalist Christians have been suspicious of Darwin and his works -- or that in the United States, where 80 percent of the population *believe* God created the universe. (COCA, written)

The analysis of the class of exceptions with collectives (a plural verb is used in spite of the fact that  $N_2$  is singular) reveals that all of the examples can be explained in terms of one of the six principles just listed.

Lexical realization of $N_2$ ( <i>people, staff, etc.</i> )	137
VP contains activity verb typically carried out by an individual	78
Formal element in VP requires use of plural; there is a distributive reading	61
VP predicates individual property of referents of Subject NP	32
VP contains a verb of cognition	20
Other cases	2
<b>TOTAL</b>	<b>330</b>

Table 12: 330 'exceptions' with collective  $N_2$  followed by a plural VP: underlying principles

The category 'other cases' contains the following two examples:

- (40) Two billion people are expected to watch coverage of the wedding. So how come two-thirds of *Britain say* they won't be tuning in. (COCA)  
 (41) Whereas the female faces usually have idealized features, however, some of the male faces are distorted. In fact, a number of the *latter depict* enemies and bear foreign facial markings. (COCA)

'Britain' is understood as 'the British people'; the plural verb is triggered by the collective noun *people*, which always combines with a plural (Declerck 1991: 247). We analysed 'latter' in the example in (41) as a pro-form for 'faces', which explains why a plural verb is used.

A similar semantic analysis accounts for the majority of the remaining 99 sentences that do not meet the predictions of hypothesis 2 (the subset of exceptions in which  $N_2$  is not a collective noun).

First, the semantics of the VP sometimes trigger a specific number:

(a) The VP refers to a situation in which all the entities or individuals are implied or engaged in one and the same process, or work towards a mutual goal; accordingly, a singular verb form is used. Applied to the cases listed below, the process of melting is seen as affecting

a mass (of glaciers), that of recycling as applying to a volume (of items). The recording of three records is the result of a joint effort of a team and the proposal that is made is the result of a joint decision. In others words, in cases like these, the situation is conceived of as applying to a unit rather than to individuals.

- (42) As much as 90 percent of some glaciers has melted in the past 85 years. (COCA, written)
- (43) Nearly a third of all plastic bottles in California now gets recycled. (COCA, written)
- (44) A group of his friends has recorded three CDs. (COCA, spoken)
- (45) A group of the largest commercial hospital chains plans to propose today that individuals be required to have basic health coverage. (COCA, written)

(b) The VP sometimes establishes a one-to-one relationship with the referent of N<sub>2</sub>; in this case, the plural verb form is necessarily used in order to mark the distributional reading.

- (46) A third of the Recovery Act were tax cuts. (COCA, written)
- (47) Three quarters of the book are my pictures. (COCA, spoken)

It appears, however, that the majority of cases can be explained in terms of the semantics of the noun phrase:

(c) When the QP is lexically realised by *a part* the verb agrees to the singular (96.9%). In ten out of eleven sentences with 'a part of X', the NP is not used in the literal sense, but rather metaphorically, which explains the use of the singular verb ((examples (48) and (49)). The example in (50) is the only case in which a plural verb is used; there is predicative construction and *be* agrees with the plural NP *the consequences*.

- (48) Meh'Lindi gripped his hand.' No, Jaq, that is not the way to think about this. One does not invite death. That is the way of fools and failures who plunge to their own destruction because a part of them has despaired and wishes to die. Thus doom accepts their invitation. (BNC, written)
- (49) A part of us, a part of our hearts, has been taken. (COCA, spoken)
- (50) A part of the hidden context, it is suggested here, are the consequences afforded by the situation. (COCA, written)

(d) When N<sub>2</sub> is realised by a temporal or spatial unit which has an internal structure, a plural verb is used instead of the singular.

- (51) Two-thirds of our planet are covered in water. (COCA, written)
- (52) Three-quarters of an hour were taken up with wondering why it was. (BNC, written)

(e) In other cases, the noun head is followed by another, embedded, noun phrase with which the verb agrees. These are clear cases of proximal agreement.

- (53) In the end, two thirds of the value of the tools destroyed were paid. (BNC, written)
- (54) About two-thirds of the book's total of one thousand pages are concerned with detailed investigations. (COCA, written)

(f) Finally, there is sometimes proximal agreement with an elided post-head NP:

- (55) The Soviet Union's newly relaxed immigration laws have allowed tens of thousands of Soviet Jews to stream into Israel. More than three quarters of a million [Soviet Jews] are expected over the next five years. (COCA, spoken)
- (56) She said this week traditionally is one of the busiest in the tourist season, when 100,000 visitors usually come here. She said only about a third of that number [of visitors] remain in Dare County, which extends from Cape Hatteras to Nags Head. (COCA, written)

Table 13 provides an overview of the distribution of these factors, which can explain a majority of the subset of examples with non-collective nouns which do not meet the predictions of hypothesis 2.

Similar situation, or mutual goal	28
Distributive reading (one-to-one relationship)	5
Lexical realisation of QP: <i>a part</i>	11
Temporal or locational nominal head	14
Complex NP	6
Ellipsis	8
Other cases	27
<b>TOTAL</b>	<b>99</b>

Table 13: 99 'exceptions' with non-collective  $N_2$ : underlying principles.

The category 'other cases' subsumes what is left of this subset of exceptional sentences, as illustrated in sentences (57) to (59).

- (57) A portion of these differences is reflective of the religious and political cultures of these nations. (COCA, written)
- (58) Two-thirds of [the first Gazette] were filled with Carrie's words and Willie's pictures. (BNC, written)
- (59) When you die, a part of me die. (COCA, spoken)

For these examples (27 out of the 429 exceptions to hypothesis 2), it is not possible to explain the number of the verb.

### 3.4. Predictive power of hypotheses

According to hypothesis 1, when  $N_1$  is singular, the verb should be singular and when  $N_1$  is plural, the verb should be plural. As is clear from the overview in figure 1, this prediction is reasonably correct when  $N_1$  is plural (78.3%), but when  $N_1$  is singular, it no longer is: we find more plural verbs (72.66%) than singular verbs (27.34%). Altogether, in more than half of the cases (54.24%), the predictions are not met (cf. table 2), that is, more than half of the time, the number of the verb does not agree with the number of the head of the construction when it is analysed as 'N of NP'. This means that for more than half of the cases, an alternative explanation has to be found. Regional variation or register differences do not appear to be at stake. (cf. table 3) Even though proximal agreement and notional agreement can be called upon, such a solution is less economical and less elegant than one in terms of one single major principle determining agreement in the majority of cases.

The alternative to the binominal analysis, according to which  $N_2$  is the head of the NP (hypothesis 2) which determines verb number, is more in line with the empirical data. Figure 2 shows that the corpus sample substantiates this claim. When  $N$  is a singular count noun or a mass noun, almost all verbs agree to the singular. When  $N$  is a plural count noun, almost all verbs agree to the plural. Table 7 shows that 95.76% of agreement patterns are predicted by hypothesis 2; only 4.24% of the sentences are not. First, compared to the rate of exception (54.24%) of the binominal analysis, 4.24% of unpredicted cases is a rather sensible rate of exception. Moreover, it is mainly collectives (330 examples) that challenge the predictions made by hypothesis 2. While they are expected to take singular verbs, many verbs are plural. The detailed analysis of the examples with collectives shows that agreement can be explained in terms of the lexical realization of  $N_2$  or in terms of the semantics of the VP. Similar observations can account for agreement in 72 out of 99 exceptional examples with a non-collective  $N_2$ . This leaves only 27 examples unexplained under hypothesis 2. In other words, it may be concluded that the empirical data clearly tilt the scales in favour of a structural analysis whereby  $N_2$  is the head of the complex NP.

Two further semantic observations support a QP of  $N$  analysis of the partitive construction. The first one relates to the meaning of *of*. *Of* instantiates the ‘part-of relation’ (Hoeksema 1996:1) in partitive constructions. The meaning of *of* is similar to that of OUT OF (Martí Girbau 2010:27).<sup>12</sup> In other words, the relation established by *of* here cannot be assimilated to its other uses as a preposition, as in, for instance, *the destruction of the city* or *a lot of paintings*. Mohanty *et al* (2004) reserve the label OFpart (OFpartitive) to refer to the specific meaning at stake. It might be argued that unlike OFprep, OFpart turns  $N_2$  into the head of the partitive construction, without which partition is impossible.

A second observation concerns the cognitive processing of complex NPs. As Lapaire & Rotgé (1993) point out, in most cases, complex NPs instantiate the structure whereby  $N_1$  is complemented by  $NP_2$ . When processing a complex NP,  $N_1$  is the starting point and a relation is expressed whereby  $NP_2$  qualifies  $N_1$ . Cognitively then, speakers first think about  $N_1$  of which they have a mental representation. Only then comes the specification of  $NP_2$ .

(60) The museum of art

(61) Two-thirds of the population

In example (60), *museum*, the starting point, is what people first think about. Only then *art* specifies what kind of *museum* it is. Partitive constructions like that in (61) stand out and do not conform to type of linear processing just described:

Bien que l’unité de dénombrement figure avant la chose dénombrée, elle lui est logiquement postérieure: l’énonciateur pose d’abord un domaine d’évaluation quantitative et procède ensuite au dénombrement. Ce qui vient à droite de OF est donc pensé antérieurement. [Even though the quantifying phrase precedes the quantified entity, the former is logically posterior to the latter: the speaker first posits a domain that needs to be quantified and only then (s)he proceeds to quantification. In other words, what precedes OF is ‘thought’ at a point later in time.] (Lapaire & Rotgé 1993:107)

In other words, in (61) the root for partition, *the population*, first comes to mind and allows for QP to take a specific quantity (*two-thirds*) out of the referent of  $N_2$ .

<sup>12</sup> The semantics of *of* potentially offer an explanation of the partitive constraint, as it requires  $N_2$  to refer to a specific group of individuals.



While semantic considerations of this type constitute further evidence for analysis whereby  $N_2$  is the head of the partitive construction, the quantitative analysis of an extensive data set offers the most conclusive evidence for this approach.

#### 4. Conclusion

In this paper, we have offered a detailed analysis of verbal agreement with partitive constructions in a large data set extracted from the BNC and COCA. Our empirical study offers solid evidence for an analysis of partitive constructions in terms of *QP of part N*. This approach is more elegant and more economical than an analysis in terms of a binominal construction, which can only account for the empirical data by drawing extensively on the additional principles of proximal agreement (and, possibly but arguably,) notional agreement. On a binominal construction approach, the latter principles are needed to account for more than half of the data.

We have also investigated whether regional variation or register impact on the distribution of singular and plural verbs. The significant patterns brought to light were all due to the presence of collective nouns. In other words, the empirical analysis has also revealed some facts about collective nouns, which turn out to be in line with former investigations into non-complex NPs with collective nouns.

## REFERENCES

- ABBOTT, B. 1996. Doing Without a Partitive Constraint. *Partitives: Studies on the Syntax and Semantics of the Partitive and Related Constructions*, ed. J. Hoeksema, 25 - 56. Berlin and New York, NY: Mouton de Gruyter.
- BIBER, D., S. JOHANSSON, G. LEECH, S. CONRAD & E. FINEGAN. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.
- BOCK, J.K. & C.A. MILLER. 1991. Broken agreement. *Cognitive Psychology* 23, 45 - 93.
- COBUILD, C. (eds). 1990. *Collins Cobuild English Grammar*. London: Collins Publishers.
- CORBETT, G.G. 2000. *Number*. Cambridge: Cambridge University Press.
- CORBETT, G.G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- DAVIS, M. 2004-. *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). Available online at: <http://corpus.byu.edu/bnc/>.
- DAVIS, M. 2008-. *The Corpus of Contemporary American English: 450 million words, 1990 present*. Available online at: <http://corpus.byu.edu/coca/>.
- DEPRAETERE, I. 2003. On verbal concord with collective nouns in British English. *English Language and Linguistics* 7, 85 - 127.
- DEPRAETERE, I. 2004. A comparative analysis of verbal concord with *staff*, *crew* and *cast*. *PALC (Practical Applications of Linguistic Corpora) 2003*, ed. B. Lewandowska-Tomaszczyk, 219 - 228. Frankfurt-am-Main: Peter Lang.
- DEPRAETERE, I. & C. LANGFORD. 2012. *Advanced English grammar: a linguistic approach*. London and New York, NY: Continuum.
- GOLDBERG, A.E. 2003. Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences* 7:5, 219 - 224.
- HAEGEMAN, L. & J. GUÉRON. 1999. *English grammar: a generative perspective*. Oxford: Blackwell.
- HOEKSEMA, J. (ed.). 1996. *Partitives: Studies on the Syntax and Semantics of the Partitive and Related Constructions*. Berlin: Mouton de Gruyter.
- DE HOOP, H. 1997. A semantic reanalysis of the partitive constraint. *Lingua* 103, 151 - 174.
- HUDDLESTON, R. & G.K. PULLUM et al. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- JACKENDOFF, R. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA.
- JACKENDOFF, R. 1977. *X-bar Syntax: A Study of Phrase Structure*. Cambridge, Massachusetts: MIT Press.

- LAPAIRE, J. and W. ROTGÉ. 1993. *Linguistique et grammaire de l'anglais*. Toulouse: Presses Universitaires du Mirail.
- LEVIN, M. 2001. *Agreement with collective nouns in English*. Stockholm: Almqvist and Wiksell .
- MARTÍ GIRBAU, N. 2003. Partitives: one or two nouns? *Proceedings of the XXIX Incontro di Grammatica Generativa*, Urbino.
- MARTÍ GIRBAU, N. 2010. *The syntax of partitives*. University of Barcelona: Doctoral thesis.
- MOHANTY, R.K., A.F. ALMEIDA & P. BHATTACHARYYA. 2004. The Complexity of OF in English. *Proceedings of International Conference on Natural Language Processing (ICON- 2004)*, Hyderabad, India.
- QUIRK, R., S. GREENBAUM, G. LEECH & J. SVARTVIK. 1985. *A comprehensive grammar of the English language*. New York: Longman.
- REID, W. 1991. *Verb and noun number in English: a functional explanation*. London and New York: Longman.
- SELKIRK, E. 1977. Some Remarks on Noun Phrase Structure. *Formal Syntax*, eds. P. Culicover, T. Wasow & A. Akmajian, 285 - 316. New York: Academic Press.
- SWAN, M. 1995. *Practical English usage*. Oxford University Press (2nd edition).