Abstract

This corpus-based study analyses the number mismatches found in English collective noun-based subjects taking *of*-complementation (i.e. determiner + collective + *of* N). The plural *of*-PPs in these constructions have proved to be a determinant factor of verbal number, as evinced by the overrides of canonical (i.e. singular) agreement with the singular collective noun that they favour. This paper aims at measuring the impact on verbal number of both the morphology of the plural noun in the *of*-dependents (i.e. overt vs. non-overt plurality as in *boys* vs. *people*) and the (syntactic and structural) complexity of the *of*-PPs (i.e. structural design of *of*-PPs in terms of constituency and in number of words). The results obtained show a strong correlation between the influence that *of*-dependents exert on the number of the verb and their complexity, thus attesting that the rate of plural verbal agreement diminishes significantly with increasing (syntactic and structural) complexity.

Keywords: agreement, collective, *of*-PP, complexity, corpus.

1. Introduction

Collective noun-based constructions have been the object of study of numerous investigations aimed at exploring the syntactic, semantic, regional and processing factors which condition their patterns of agreement (Dekeyser 1975; Levin 2001; Hundt 2006, 2009; Acuña-Fariña 2009). However, the potential effects of the *of*-dependents which frequently accompany some collective nouns have been overlooked on many occasions. This investigation extends previous studies on the matter by considering, from a syntactic point of view, the consequences which *of*-dependents have for the patterns of verbal agreement of a set of twenty-three collective nouns in British and American English.

The organisation of the paper is as follows: First, in Section 2 I will define some basic theoretical concepts. In Section 3, after describing the corpora and the methodology used, I will present the results obtained. The paper will be rounded off by the discussion of the main conclusions and the questions for future research in Section 4.

---

1 The research reported in this paper was funded by the Spanish Ministry of Science and Innovation and the European Regional Development Fund (grants no. FFI2013-44065-P, FFI2014-51873-REDT and FPU FPU13/01509) and the Autonomous Government of Galicia (grants no. GPC2014/060 and R2014/016).
2. Agreement and collective noun-based constructions

The phenomenon of (canonical) agreement is usually defined as the “systematic covariance between a semantic or formal property of one element and a formal property of another” (Steele 1978, 610). Two elements are therefore involved in this operation: the ‘controller’, the (usually nominal) element which determines the agreement relation, and the ‘target’, the element whose form is conditioned by the features of the controller (Corbett 2006, 5), as illustrated in Figure 1:

![Canonical or syntactic agreement](image)

In the case of collective nouns, however, speakers — especially British speakers — very frequently take their inherently plurality as the ultimate determinant of verbal number. As a matter of fact, the form of the target is determined by the semantic (not the formal) characteristics of the controller, thus conforming to semantic or notional agreement, as in (1):

(1) The committee have decided (Corbett 2006, 155)

The constructions under scrutiny here consist of not only a collective noun but also an of-dependent with a plural nominal element which often constrains the patterns of verbal agreement of these structures. In fact, the plural non-head nominal element within the of-PP tends to affect the agreement operation and thus trigger plural verbal number, the so-called ‘attraction’ (Bock et al. 2001; Acuña-Farías 2009), illustrated in (2):

(2) A group$_{SG}$ of parents$_{PL}$ were$_{PL}$ standing in the corner [BNC: CHR 861]

In this respect, it is worth drawing attention to the fact that in (2) the semantics of the collective noun has most probably an effect on verbal number too (i.e. semantic agreement), but this issue has been left for future research.

---

2 The concept of ‘collective noun’ is taken here in a broad sense, as in Dekeyser (1975, 35fn.1), meaning “all morphologically singular nouns designating a group of animates or inanimates”.

3 In the psycholinguistic literature ‘attraction’ is associated with errors or ungrammatical outcomes of agreement relations. In this paper it will be given a wider scope and, thus, it will denote the (grammatically correct or not) influence exerted by the of-dependent on the number of the verb, as in (2).
3. Case study

3.1. Methodology and data

This section introduces the corpora and the data supporting the conclusions of this paper and is rounded off by the analysis of the complexity of the of-dependents along with its role in verbal agreement.

3.1.1. Corpora and data retrieval

The main sources of data here consist of samples of the written components of the British National Corpus (BNC) and The Corpus of Contemporary American English (COCA). The data analysed were restricted to a set of twenty-three singular collective nouns usually taking plural of-complements: band, batch, bunch, class, clump, couple, crowd, flock, gang, group, herd, host, majority, minority, number, pack, party, rash, series, set, shoal, swarm and troop⁴ (list retrieved from both Biber et al. 1999, 249 and Huddleston and Pullum et al. 2002, 503).

Since the object of study is verbal agreement with collective nouns taking plural of-PPs, only those instances containing (i) verbs inflected for number and (ii) morphologically marked or unmarked plural nouns within the of-PP were included in the database. Furthermore, this investigation is limited to a maximum of six thousand instances per collective noun, which were manually analysed so as to discard those that were not valid for this study, that is, those that did not show verbal agreement.

3.1.2. Data

The main object of this study concerns the patterns of verbal agreement of collective noun-based subjects containing plural of-dependents. These binominal NP subjects comprise both a singular collective head noun and a plural oblique noun, that is, the plural noun within the of-dependent, which can be either a morphologically-marked plural noun (here also ‘NN2’), as in (3), or the morphologically unmarked plural noun people,⁵ as in (4).

(3) a group of boysNN2 / thingsNN2

(4) a group of people

The presence of plural of-dependents has proved to be a significant determinant of verbal agreement. As a matter of fact, although only singular collective nouns are taken into consideration in this study, the data show a general tendency towards plural agreement when the of-dependent is present. By contrast, in this same set of twenty-three collective nouns, the

---

⁴ Only the collective/quantificational meanings of the collective nouns have been considered here. Instances such as Our band of hearing includes all the sounds which are significant for us [BNC: FEV 929] have been excluded.

⁵ People is the most frequent non-overtly-marked plural noun in my database (almost 80% of the instances of both the BNC and COCA) and, thus, is the only noun in this category that will be considered here.
absence of the prepositional constituent evinces a clear preference for singular verbal forms, as Figure 2 illustrates.

As Figure 2 shows, plural of-PPs have remarkable consequences for verbal agreement. The differences observed for the contrast between the presence and absence of the prepositional dependent turn out to be statistically significant ($\chi^2(1), p<0.0001$), which implies that the plural noun within the of-PP constitutes one of the factors conditioning the mismatches in subject-verb number agreement observed in this study, as illustrated in (5):

(5) a group$_{SG}$ of British skiers$_{PL}$ were$_{PL}$ horrified to see a man [BNC: CCK 737]

Mismatches such as (5) will find an explanation here in formal factors such as the morphology of the plural oblique noun within the of-PP (§3.2.1) or the (syntactic and structural) complexity of the of-PPs under scrutiny (§3.2.2).

3.2. Analysis of the data

In this section I will measure the significance of the plural of-dependent in the structures analysed here in relation to the morphology of the plural oblique noun and the syntactic and structural complexity of the prepositional constituent.

3.2.1. Attraction

As already pointed out, attraction refers to the process whereby the ‘interloper’ (or non-head nominal element) “illegally attracts [number] agreement in the verb” (Acuña-Fariña 2009, 392). Since this phenomenon is more frequently attested in local syntactic domains, this section will concentrate on bare of-PPs, that is, on those constituents which lack both premodification and postmodification of the oblique noun, as in (6).

(6) The crowd [of cockneys]$_{BARE\ OF\ PP}$ were singing along [BNC: BPA 62].
According to the data obtained for both the British and the American varieties, attraction is highly significant in both corpora, as the presence of the plural of-dependent favours a remarkable rate of plural verbal forms (over 60%), as illustrated by Figure 3:

![Figure 3. Agreement (%) with each oblique noun in bare of-PPs in the BNC and COCA](image)

Apart from confirming the determinant role of attraction, Figure 3 evinces incipient differences between the two types of nominal elements considered here, namely overtly-marked plural nouns (or NN2) such as boys, and the morphologically unmarked plural noun people. Hence, contrary to expectations, people exerts a higher influence on verbal number than overtly-marked plural nouns, a finding which seems to suggest that the presence of an overt morphological plural marker does not necessarily reinforce conceptual plurality. The difference between these two tendencies lends statistical support in both varieties ($\chi^2(1)$, $p<0.0001$) and, thus, entails important implications which will deserve further consideration in future research.

3.2.2. Syntactic and structural complexity

Authors like Corbett (1979) or Levin (2001, 95) claim that syntactic distance between the controller and the target increases the proportion of plural (i.e. semantic) agreement. This trend is said to be particularly salient with collective nouns, since across syntactic boundaries our short-term memory keeps activated the semantic (not the formal) characteristics of the previous lexical material, which favours a stronger preference for plural verbal forms. In this vein, the likelihood of finding plural agreement can be expected to be higher in the structures under scrutiny here, as they comprise both a semantically plural collective noun and a morphologically and semantically plural oblique noun within the of-PP. As a way of measuring the effects of distance on agreement, the present study will examine both the syntactic and the structural complexity of the main object of study: the of-dependent.
a. Syntactic complexity

Syntactic complexity is taken here as the analysis of the shallow syntactic structure of both the of-PP and the lexical material following it. In this respect, this study will focus on four different syntactic configurations of the prepositional dependent: bare of-PPs (7), premodified of-PPs (8), postmodified of-PPs (9), and both pre- and postmodified of-PPs (10).

(7) Yet a minority [of Senators] has hijacked the process [COCA: ACAD LawPublicPol]

(8) a bunch [of white men] carve up their land. [BNC: CGC 1003]

(9) the vast majority [of Israelis who were happy with the status quo] are now much more unhappy with it. [BNC: A9E 415]

(10) A group [of Chinese seniors living at the Frances Beavis residence] (Figure 2) have transformed [COCA: ACAD GeographRev]

<table>
<thead>
<tr>
<th>Structure of the of-PP</th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. BARE NP</td>
<td>31.95</td>
<td>68.05</td>
</tr>
<tr>
<td>2. PREMOD + NP</td>
<td>34.99</td>
<td>65.01</td>
</tr>
<tr>
<td>3. NP + POSTMOD</td>
<td>34.57</td>
<td>65.43</td>
</tr>
<tr>
<td>4. PREMOD + NP + POSTMOD</td>
<td>40.83</td>
<td>59.17</td>
</tr>
</tbody>
</table>

Table 1. Agreement (%) with NN2 and people in bare, pre- and/or postmodified of-PPs in the BNC and COCA

Table 1 evinces that the four structures favour plural agreement, that is, attraction. However, one can also observe a general decrease in the rate of plural agreement as structures become syntactically more complex. In fact, a closer analysis of the data has lent statistical support to the contrast between the patterns of verbal agreement of bare of-PPs and those of the of-PPs which are both pre- and postmodified ($\chi^2$(1), p=0.0002).

Given this result, which refutes the claims of the literature hitherto, the ensuing sections will provide a more fine-grained analysis of the data which aims at shedding some light on the decreasing tendency in plural agreement observed so far. To this end, both postmodified and pre-/postmodified structures will be considered so as to explore the potential influence of modification on verbal agreement.
Figures 4 and 5 present the typology of the postmodifier according to the data obtained for postmodified of-PPs and pre-/postmodified of-PPs, two structures which show very similar results. The figures demonstrate that the of-dependentst studied here are very frequently postmodified by prepositional phrases and TPs (i.e. non-finite clauses) – over 38% and about 30% respectively in both corpora. Leaving aside the category ‘other’, which comprises a wide range of minor types of postmodifiers in my database, the rest of the categories can be organised according to their syntactic configuration into two different categories: (i) non-clausal, that is, PPs and NPs, and (ii) clausal, both TPs and relative clauses.

<table>
<thead>
<tr>
<th></th>
<th>BARE NP + POSTMOD</th>
<th>PREMOD + NP + POSTMOD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>PL</td>
</tr>
<tr>
<td>Non-clausal</td>
<td>34.66</td>
<td>65.34</td>
</tr>
<tr>
<td>Clausal</td>
<td>49.06</td>
<td>50.94</td>
</tr>
</tbody>
</table>

Table 2. Agreement (%) according to each type of postmodifier of the of-PP in the BNC and COCA

The division of the data reveals a correlation between the syntactic complexity of the postmodifiers and the patterns of verbal agreement. Hence, in line with the results presented so far, non-clausal postmodifiers favour a higher percentage of plural agreement. On the other hand, clausal constituents, despite being syntactically more complex, show in general terms a lower rate of collocation with plural verbs. This remarkable finding corroborates the previous tendencies and, thus, leads to the confirmation that syntactic distance and complexity do not correlate with a higher likelihood of plural agreement in the structures studied here. These figures also corroborate the negligible effect of premodification on agreement, as its presence yields no statistically significant results in the figures above.

Given the significance of attraction in the constructions studied here, a final variable in relation to syntactic complexity has also been considered: the presence of plural nominal elements in the postmodifier of the of-PP and their possible interference in verbal agreement.
Table 3. Agreement (%) in relation to the number of NN2 within the postmodifier of the of-PP in the BNC and COCA

Table 3 shows only the figures for morphologically marked plural nouns (NN2), as morphologically unmarked plural nouns yielded no significant data in this respect. As a result, it has been found out that, as the number of nouns inflected for plural number increases, so does the rate of plural agreement. Nonetheless, the limited amount of instances from the database containing more than three plural nominal elements does not tolerate statistical validation.

b. Structural complexity

The structural complexity of the construction is measured by taking into account the number of lexical elements intervening in between the oblique noun within the of-PP and the main verb. In what follows, I will deal with the number of words separating the oblique noun from the verb and its potential influence on verbal agreement.
As Table 4 illustrates, with the increasing number of words the rate of plural agreement diminishes considerably, to such an extent that the contrast between categories ‘0’ and ‘20’ is quite significant ($\chi^2(1)$, $p=0.02$). Similar results are applicable to the analysis of the structural complexity of the postmodifier of the of-PP, although in this case the data retrieved lend no statistical support for the correlation between structural complexity and the decrease in plural verbal forms.

All in all, the data presented confirm that in collective noun-based constructions taking of-dependents syntactic distance and, therefore, (syntactic and structural) complexity does not increase the likelihood of finding plural agreement, not even when the oblique position is occupied by a (overt or non-overt) plural noun. The explanation for such a result seems to lie in the actual complexity of the prepositional phrase, yet more research is needed in this respect as regards the other variables considered here. Notwithstanding, the results obtained allow me to refute Levin’s (2001) argumentation and findings since the study of complexity yielded significant decreasing tendencies of plural agreement in postmodified structures.

In brief, the statistically significant correlation between the increasing complexity of the of-PP and the progressive decline in the frequency of plural verbal forms in the data analysed seems to suggest important processing implications which point to the preference for the use of the singular number as a resource to ease the cognitive processing of complex constructions such as the ones investigated here, a remarkable finding which will be given further consideration in future research.

4. Concluding remarks and further research

This study has explored the extent to which of-dependency and complexity interact in the accomplishment of the agreement operation in collective noun-based constructions. The results obtained confirm that the plural of-dependent is a statistically significant determinant of plural verbal number in the structures analysed here, yet only in local syntactic domains, a trend which is particularly significant with the oblique noun which does not show overt plural marking: people.

The analysis of the syntactic and structural complexity of of-dependents proves that, although the data evince some significant tendencies, they do not support the increase in plural agreement with increasing complexity attested in the literature. By contrast, in this study, the influence of plural of-dependents on verbal number significantly weakens in the most (syntactically and structurally) complex collective noun-based subjects.

Despite the arguments and conclusions reported here, more investigation is needed so as to carry out a more fine-grained analysis of the implications of overt vs. non-overt morphology as well as to consider further variables which have been attested as relevant determinants of subject-verb agreement, as is the case of the possible idiomatisation, or even grammaticalisation, of the constructions under scrutiny here (see Traugott 2008a, 2008b and Traugott and Trousdale 2013 on a lot of and homologous constructions).
References


BNC (Interface Lancaster University): http://bncweb.lancs.ac.uk/cgi-binbncXML/BNCquery.pl?theQuery=search&urlTest=yes

BNC (Interface Brigham Young University): http://corpus.byu.edu/bnc/


COCA (Interface Brigham Young University): http://corpus.byu.edu/coca


——. 2006. Agreement. Cambridge: Cambridge UP.


