

# Dutch collective nouns and conceptual profiling\*

FRANK JOOSTEN, GERT DE SUTTER, DENIS DRIEGHE, STEF GRONDELAERS, ROBERT J. HARTSUIKER, AND DIRK SPEELMAN

*There are groups and groups.*  
(Michael Kinsley)

## Abstract

*Collective nouns such as committee, family, or team are conceptually (and in English also syntactically) complex in the sense that they are both singular (“one”) and plural (“more than one”): they refer to a multiplicity that is conceptualized as a unity. In this article, which focuses on Dutch collective nouns, it is argued that some collective nouns are rather “one”, whereas others are rather “more than one”. Collective nouns are shown to be different from one another in member level accessibility. Whereas all collective nouns have both a conceptual collection level (“one”) and a conceptual member level (“more than one”), the latter is not always conceptually profiled (i.e., focused on) to the same extent. A gradient is sketched in which collective nouns such as bemanning (‘crew’) (member level highly accessible) and vereniging (‘association’) (member level scarcely accessible) form the extremes. Arguments in favor of the conceptual phenomenon of variable member level accessibility derive from an analysis of property distribution, from corpus research on verbal and pronominal singular-plural variation, and from a psycholinguistic eye-tracking experiment.*

## 1. Introduction

Collective nouns such as *committee*, *family*, or *team* are all too often erroneously defined as semantically plural but grammatically singular, as nouns that have “singular form but plural meaning” (Chelaru-Ioniță and Bantaş 1981: 224).<sup>1</sup> In fact, things are more intricate: “The characterization of [...] expressions such as *the team* as semantically plural is actually

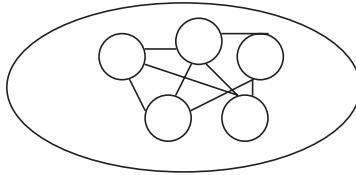


Figure 1. *Conceptualization of a collective noun*<sup>5</sup>

a half-truth: at a higher level, they may also be viewed as semantically singular” (Gil 1996: 64).<sup>2</sup>

It should be intuitively clear that collective nouns such as *team* have a complex conceptual structure. As Jespersen (1924: 195–196) puts it, “a collective [noun] [...] is logically from one point of view ‘one’ and from another point of view ‘more than one’”. A team, for instance, is ‘more than one’ in the sense that it consists of a number of people, but these people, being externally related to each other, at the same time can be thought of as a distinct conceptual unity. More precisely, then, a collection, i.e., the conceptual counterpart of a collective noun, has two conceptual individuation levels that we will name *collection level* (i.e., the collection conceptualized as ‘one’) and *member level* (i.e., the collection conceptualized as ‘more than one’, viewed in terms of the individual members making up the collection) (see Figure 1).<sup>3</sup> This conceptual “double-sidedness” explains the alternations between singular (*a group* vs. *\*many group, the committee is ...*) and plural constructions (*\*a group of student* vs. *a group of students, the committee are ...*) with collective nouns.<sup>4</sup>

Though all collective nouns have such a dual conceptual structure, we argue that they may differ from one another in the conceptual profiling of the individuation levels. (*Conceptual*) *profiling* is the term used in Cognitive Grammar for the elevation of a conceptual substructure to a special level of prominence. The *profile* is defined as the “substructure [...] that is obligatorily accessed, [that] functions as the focal point [...], and achieves a special degree of prominence”. It “stands out in bas-relief” against the *base* (Langacker 1987: 183, 491).

Not all collective nouns profile the members of the collection to the same extent. Profiling is a matter of degree (Langacker 1987: 218) and collective nouns differ in the conceptual accessibility of the member level (or, put differently, in the conceptual permeability of the collection level). An introductory example: in *an old club* the adjective *old* applies to the collection level (the individual members can be young), whereas in *an old audience* it directly indicates a property of the members, which suggests that the member level of *club* is less easily accessible (i.e., profiled to a



Figure 2. Club, audience, and member level accessibility

lesser extent) than the member level of *audience*. Diagrammatically, this difference can be represented as in Figure 2.

The rest of this article is organized as follows. In Section 2 we briefly provide some further thoughts on the definition of collective nouns, comment on the various contextual factors that may influence conceptual profiling and consequently formulate some methodological restrictions that have to be taken into account. Section 3 offers a first approximation of variable member level accessibility in Dutch collective nouns, based on their differences in property distribution. Collective nouns are shown to be different from one another in the way they distribute properties such as *big* or *young* over their collection and member level. In Section 4 these findings are corroborated by corpus data: an analysis of verbal and pronominal singular-plural variation for about twenty Dutch collective nouns demonstrates that high member level accessibility and high plural concord go together. Section 5 adds a last type of evidence: the results of a psycholinguistic eye-tracking experiment suggest that low member level accessibility results in significantly slower reading times for plural pronouns. Section 6 addresses some remaining questions and suggestions for further research.

## 2. Defining collective nouns and limiting the field

As collective nouns have only been introduced by ostension until now, perhaps some further elucidation is in order. Though lack of space prevents us from discussing in detail the numerous definitions of the term *collective noun*,<sup>6</sup> not only crosslinguistically (see Gil 1996), but also within one language (see Benninger 2001), a basic distinction between two general definitional tendencies can be resumed briefly.

In the Anglo-Saxon tradition (e.g., Juul 1975; Quirk et al. 1985; Bache and Davidsen-Nielsen 1997; Levin 2001) it is customary to define collective nouns fairly strictly, on the (primarily) syntactic basis of variable

concord. Collective nouns, then, are nouns such as *committee*, *family*, or *team*, i.e., nouns that in the singular may combine with both singular and plural verbs and pronouns:

- (1) a. The committee has met and it has rejected the proposal.  
 b. The committee have met and they have rejected the proposal.  
 (Quirk et al. 1985: 316)

Consequently, in Anglo-Saxon studies collective nouns are generally considered to be animate; inanimate nouns such as *forest* or *archipelago* do not allow variable concord (e.g., *the forest is* / *\*are* . . . *it* / *\*they*; *the archipelago has* / *\*have* . . . *it* / *\*they*).<sup>7</sup>

Opposed to the Anglo-Saxon syntactic notion of collective nouns, there is a “continental” semantic tradition. In French, German, and Dutch studies (e.g., Michaux 1992; Borillo 1997; Lecolle 1997; Leisi 1975; Kuhn 1982; Mihatsch 2000; Haeseryn et al. 1997), for instance, collective nouns are usually NOT defined on syntactic grounds. Instead, a broad semantic definition prevails: collective nouns are nouns lexically referring to a (denotational) multiplicity that — in some way or the other — is conceptualized as a unity.<sup>8</sup> French collective nouns, then, are nouns such as *club*, *comité*, *archipel*, *groupe*, *troupeau*, *bourgeoisie*, *Pyrénées*, *lingerie*, *meublier*, or *bétail*. Of course, the reason why collective nouns in French, German, and Dutch are not defined on syntactic grounds, is the general absence of variable verbal concord in those languages, e.g., *Le comité a* / *\*ont rejeté la proposition*.<sup>9</sup>

Our position, broadly discussed in Joosten (2003), is more or less a compromise between these two traditions. In an analysis based on form-meaning relationships, Joosten demonstrates that the “continental” semantic category of collective nouns is not only morphosyntactically, but also semantically disparate. He adduces conceptual-syntactic arguments that strongly contradict the view that there is a notion of conceptual unity in noncount nouns such as *furniture*, *clothing*, or *jewelry* (and their noncount equivalents in other languages). The end result is a rigid terminological distinction between count *collective nouns*, that can be combined with the singular determiner *a* and be pluralized (*club*, *team*, *group*, *herd*, *archipelago*), and noncount *aggregate nouns*, that occur with the zero determiner and with *much* and that cannot be pluralized (*furniture*, *clothing*, *jewelry*, *scum*, *underwear*). Fuzzy categories are situated somewhere in between (e.g., nouns such as *nobility*, *clergy*, *press*).<sup>10</sup>

In short, an important idea for the remainder of this article is that syntactic differences can generally be shown to correlate with conceptual ones — a certain degree of syntactic arbitrariness notwithstanding. That means that if we want our discussion of variable member level accessibility

to be credible, we need to select nouns that are syntactically similar. If not, we risk undertaking a meaningless analysis, viz. one that tries to compare the incomparable.

Conceptual profiling (of the collection or member level) is the product of quite a few intertwining factors. It is not only dependent on the (conceptual-syntactic) type of noun, but also on factors such as the predicated property, animacy, and relationality.

When a team is said to be *big*, the collection level is profiled; when it is said to be *young*, the conceptual focus shifts to the individual members. The property<sup>11</sup> in question, therefore, often determines whether the collection level or the member level is conceptually profiled. But other factors may also play a role. Mihatsch (2000: 48), for instance, rightly argues that animacy influences conceptual individuation and profiling: “Eine Reihe von Faktoren begünstigen die Fokussierung der Elemente [...] im Diskurs [...]. So sind menschliche Elemente, danach tierische stärker individualisiert als unbelebte”. As said above, variable concord in English is restricted to animate collective nouns: *archipelago* does not allow plural verbs or pronouns. In Dutch, verbs are generally singular, no matter whether the collective noun is animate or not (see above), but as far as pronominal concord is concerned, animacy definitely plays a role. As in English, animate and — *a fortiori* — human collective nouns are regularly followed by plural pronouns (2a), whereas inanimate ones never are (2b).

- (2) a. Het comité kwam gisteren samen. Ze hebben het voorstel verworpen.  
 ‘The committee gathered yesterday. They rejected the proposal.’  
 b. In de Atlantische Oceaan ligt er een mooie archipel. \*Ze zijn onbewoond.  
 ‘In the Atlantic Ocean there is a beautiful archipelago. \*They are uninhabited.’

In Sections 4 and 5 we will demonstrate that pronouns are highly revealing in a discussion on conceptual profiling.

Another interfering factor is relationality. Some collective nouns are highly relational, i.e., highly dependent on a contextual identification of the members. For instance, collective nouns such as *group*, *herd*, or *swarm* usually require more information about the identity of the members, often in the form of an *of* complement (e.g., *a group of boys*, *a herd of sheep*, *a swarm of wasps*). It is more than likely that if the members are explicitly identified, they are conceptually more salient: “Tauchen die Elemente explizit im Syntagma auf, so treten sie ebenfalls leichter in den Vordergrund” (Mihatsch 2000: 48).<sup>12</sup>

Though conceptual profiling is evidently influenced by a number of contextual factors, the basic tenet of this article is that collective nouns can be LEXICALLY different with regard to member level accessibility. In order to demonstrate this, it is desirable to minimize contextual factors as much as possible. That is why we prefer to limit our discussion in what follows to Dutch human collective nouns, such as *vereniging* ('association'), *team* ('team'), or *echtpaar* ('married couple'). (Count) collective nouns are preferred to (noncount) aggregate nouns, since they are more variable in their distribution of collection and member level readings (Joosten 2003). Furthermore, aggregate nouns generally tend to be inanimate (Joosten 2003: Ch. 5); as inanimate nouns lack a pronominal singular-plural distinction (see above), differences in conceptual profiling can be analyzed less easily. Finally, human collective nouns are also preferred to nonhuman ones because of their low degree of relationality (Joosten 2003: Ch. 6).

Strangely enough, the suggestion that not all collective nouns are 'one' and 'more than one' to the same extent has scarcely been made in the collective noun literature. Admittedly, Mihatsch (2000) discusses the most essential individuation differences between collective nouns and aggregate nouns,<sup>13</sup> but apart from that, allusions to the phenomenon of variable member level accessibility are limited to the following two short quotations:

Different lexical items may [...] be associated with different degrees of plurality and singularity. (Gil 1996: 64)

A fact that is often forgotten is that there are differences between the concord patterns of individual nouns. For instance, plural forms are more likely with nouns like *family* and *team* than with nouns like *committee* and *government* [...]. The reason is probably that speakers and writers more often think of families and teams as being made up of different individuals, while committees and governments more often are seen as units. (Levin 1998: 16–17)

Both observations are based exclusively on the variable concord patterns of different collective nouns. That is possibly too rash a conclusion, since *a priori*, differences in verbal or pronominal concord do not have to coincide with differences in conceptual profiling. Our first step will therefore be to look at variation in property distribution.

### 3. Variation in property distribution

As amply demonstrated by Vossen (1995), "conceptual individuation is made explicit by the distribution of predicated properties, reflecting the

entities we have in our minds [...]” (Vossen 1995: 35). A first, rough approximation of the phenomenon of variable member level accessibility can therefore be based on the way in which collective nouns distribute properties over their collection and member level.

Let us start with a fairly straightforward English example. The properties *small* and *young* combined with the collective nouns *association*, *team*, and *couple* clearly yield distinct interpretations.

- (3) a. a small, young association  
 b. a small, young team  
 c. a small, young couple

In (3a) both *small* and *young* normally apply to the collection level (compare e.g., *a small association of giants*, *a young association of elderly people*), whereas in (3c) they both apply to the member level: a small, young couple is a couple that consists of small, young people. In (3b) *small* is a property of the collection, but *young* applies to the members. The variable distribution of collection and member level interpretations for the same group of adjectives suggests a conceptual difference between the three collective nouns: in (3a)–(3c) there is a gradual increase of member level accessibility. *Team* has a member level that is conceptually more accessible than that of *association*, but less accessible than the member level of *couple*.<sup>14</sup>

In the same vein we will make a selection of properties for Dutch. This is by no means an easy or self-evident task, since for the majority of properties collection and member level interpretations are hardly distinguishable. In *a rich family* or *a motivated team*, for instance, *rich* and *motivated* seem to apply to both levels simultaneously: rich families cannot exist without rich family members and a team can only be motivated if (most of) its members are. Properties of interest, therefore, are only those that clearly distinguish between collection and member level interpretations, or those that trigger collection or member level interpretations exclusively. For Dutch one could select the six properties in Table 1.

Table 1. *A selection of properties*

Property	Result <sup>15</sup>	Individuation level
<i>eeuwenoud</i> ‘age-old’	+ or –	+ = collection level
<i>oprichten</i> ‘to found, to start’	+ or –	+ = collection level
<i>groot</i> ‘big’	+	+ = collection level or member level
<i>jong</i> ‘young’	+	+ = collection level or member level
<i>blond</i> ‘blond’	+ or –	+ = member level
<i>dronken</i> ‘drunk(en)’	+ or –	+ = member level

*Eeuwenoud* ('age-old') and *oprichten* ('to found, to start') are two collection level "triggers". Since human beings cannot become age-old and cannot be founded or started, human collective nouns combinable with these properties should be interpreted at the collection level. The opposite holds for *blond* ('blond') and *dronken* ('drunk(en)'): these adjectives typically apply to individuals and can thus be considered member level "triggers".<sup>16</sup> Finally, *groot* ('big') and *jong* ('young') are two properties that may yield both collection and member level interpretations, but those interpretations are fairly distinct. Collective nouns usually prefer either of them. *Groot* in member level interpretations means literally 'big, tall'; in collection level interpretations it is synonymous with 'composed of many members'. *Young* either means 'youthful' (member level) or 'recently formed' (collection level).

Not only selecting adequate properties, but also interpreting their combination with collective nouns is far from evident. A certain degree of subjectivity and personal variation notwithstanding, however, it seems safe to say that Dutch collective nouns can be divided into three main types. Taken together, the types form a gradient of increasing member level accessibility (see Table 2).

### Type 1

*vereniging* 'association', *maatschappij* 'company', *firma* 'firm', *bond* 'union', *club* 'club', *partij* 'party', *organisatie* 'organisation', *comité* 'committee', *koor* 'choir', *leger* 'army', *regering* 'government', *orkest* 'orchestra', *orde* 'order'

The first type of collective noun consists of those that generally trigger collection level interpretations. They can pattern with *eeuwenoud* ('age-old') and *oprichten* ('to found, to start'), whereas combinations with the member level "triggers" *blond* ('blond') and *dronken* ('drunk(en)') turn out to be very unusual (e.g., *??een dronken vereniging* '??a drunken association'). When said to be *groot* ('big') or *jong* ('young'), the properties usually apply to the collection level, though *jong* can sometimes be interpreted as referring to the individual members as well.

Table 2. *Dutch collective nouns and variation in property distribution*<sup>17</sup>

<i>eeuwenoud</i> (‘age-old’)	<i>oprichten</i> (‘to found, to start’)	<i>groot</i> (‘big’)	<i>jong</i> (‘young’)	<i>blond</i> (‘blond’)	<i>dronken</i> (‘drunk(en)’)	
c	c	c	c(/m)	—/?	—/?	Type 1
—	c	c	m	m	m	Type 2
—	—	c/m	m	m	m	Type 3



The fact that associations, committees, parties, orchestras and so on can become age-old suggests low member level accessibility: the collections can live a life on their own, independent of the individual members. As noted by Dölling (1991: 164), some collections “may, within certain limits, gain or lose members without detriments to their identity and continued existence. For example, an orchestra can continue to exist even though in the course of time it undergoes a complete change of membership.” The opposite is also true: an orchestra can cease to exist without the members having died.

Another typical characteristic of associations, committees, or orchestras is that identical membership does not necessarily imply full identity. It is possible, for instance, for two committees, *Committee A* and *Committee B*, to have the same members.<sup>18</sup> That explains why (4) sounds perfectly normal:

(4) The committee is new, but the members are still the same.

### Type 2

*team* ‘team’, *bende* ‘gang’, *familie* ‘family’, *ploeg* ‘team’, *staf* ‘staff’, *redactie* ‘editorial staff’, *klas* ‘class’, *jury* ‘jury’, *panel* ‘panel’, *delegatie* ‘delegation’

The second type of collective noun seems to occupy a middle position, in the sense that both collection and member level interpretations are very common. In contrast to the distribution pattern sketched for type 1, *eeuwenoud* does not yield semantically acceptable combinations for this group (e.g., \**een eeuwenoud team* ‘\*an age-old team’).<sup>19</sup> *Dronken* (‘drunk(en)’) and *blond* (‘blond’), on the other hand, trigger member level interpretations and usually the same goes for *jong* (‘young’). Both *oprichten* (‘to found, to start’) and *groot* (‘big’) are generally interpreted as collection level properties.

### Type 3

*duo* ‘duo, pair’, *echtpaar* ‘married couple’, *klied* ‘clique’, *gezin* ‘family, household’, *publiek* ‘public’, *bemannig* ‘crew’, *tweeling* ‘twins’, *trio* ‘trio, threesome’

The third and last type of collective noun is in many ways the opposite of type 1. Put simply: couples cannot become age-old or be founded, but they can be blond or drunk. Type 3 nouns generally trigger member level interpretations, not only when combined with *blond* (‘blond’) or *dronken* (‘drunk(en)'), but also with *jong* (‘young’) and often with *groot* (‘big’): if the number of members in the collection is fixed (*duo*, *echtpaar*, *tweeling*, *trio*), the only available option is a member level interpretation.<sup>20</sup> Almost

anything that is predicated of type 3 nouns, can be applied to the member level.

Type 3 collective nouns exhibit maximal member level accessibility. For collective nouns such as *echtpaar* ('married couple') or *bemannig* ('crew') the collection level and member level are far less distinct than for *vereniging* ('association') or *comité* ('committee'). The collection does not live its own life independent of the individual members: if all the crew members die, there is not a crew any more, a couple in which one of the two partners leaves the other, is no longer a couple, and it seems fairly difficult to distinguish *family A* from *family B* if the members of A and B are fully identical:

(5) ??The family is new, but the members are still the same.

In short, type 1 and type 3 collective nouns are two extremes of the same gradient, a gradient of member level accessibility. Further evidence for this cline is that corpus examples in which collective nouns are accompanied by the reciprocal pronoun *elkaar* ('each other'), by the prepositions *tussen* ('between') or *onder* ('among'), or by the adverb/adjective *onderling* ('mutual, between themselves'), only include type 2 and type 3 nouns. Type 1 nouns are notoriously lacking, which is far from surprising: the constructions mentioned above have more than one conceptual slot to be filled (e.g., *between (X, Y, (Z))*), which means that only collective nouns exhibiting a sufficient degree of 'more than one'-ness (i.e., of member level accessibility) are likely candidates. In the next section other corpus data will provide further arguments in favor of the gradient.

#### **4. Variation in verbal and pronominal singular-plural oppositions: corpus research**

A problem of the "property distribution approach" outlined in the previous section is that it can only assign relative positions on a gradient. The method does not provide a means to quantify member level accessibility. In this section we demonstrate that this problem can largely be solved by corpus research on verbal and pronominal singular-plural oppositions. As already mentioned, it is plausible that differences in member level accessibility correlate with differences in verbal and/or pronominal concord. If that turns out to be true, then two independent methods (property distribution and corpus research) point to one and the same phenomenon. It is obvious that such a combined approach can less easily be falsified than an argumentation that is exclusively based on property distribution or concord patterns (e.g., Levin's 1998 conclusion in Section 2).

Corpus research on verbal and pronominal concord with collective nouns is far from novel. Geerts (1977) for Dutch and Nixon (1972), Levin (2001) and Depraetere (2003) for English, all deal with collective nouns and concord phenomena. What is new, however, is our focus on possible lexically determined differences between collective nouns. The authors mentioned above are primarily concerned with contextual factors affecting concord patterns, and therefore generally fail to discuss individual differences between collective nouns. Furthermore, Geerts' (1977) examples are only meant to be illustrative (as figures and statistics are absent), Nixon's (1972) corpus is fairly limited in size, and Depraetere (2003) restricts her discussion to verbal concord only. Levin's (2001) monograph offers the most elaborate treatment of concord with English collective nouns, but he too is scarcely concerned with possible conceptual-lexical motivations for the concord patterns in his corpus material.<sup>21</sup> Eighteen singular collective nouns, all selected from the three main types presented in Section 3, were analyzed in Dutch language corpora for verbal and pronominal singular-plural oppositions:

- (6) Type 1: *bond* 'union', *club* 'club', *comité* 'committee', *firma* 'firm', *koor* 'choir', *leger* 'army', *maatschappij* 'company', *regering* 'government', *vereniging* 'association' (9)  
 Type 2: *bende* 'gang', *delegatie* 'delegation', *familie* 'family', *team* 'team' (4)  
 Type 3: *bemannings* 'crew', *duo* 'duo, pair', *echtpaar* 'married couple', *gezin* 'family, household', *publiek* 'public' (5)

Four types of singular-plural oppositions were investigated: verbal concord (7a) and three types of pronominal concord, viz. relative (7b), possessive (7c), and personal pronouns (7d):

- (7) a. Het koor zal bekende nummers ten gehore brengen.  
 'The choir will (third person singular) sing well-known songs.'  
 (INL)  
 b. Een Braziliaans gezin dat een pizza had besteld, [...]'  
 'A Brazilian family that (singular neuter) had ordered a pizza, [...]'  
 (Condiv)  
 c. Het leger liet zien wat hun honden in hun mars hebben.  
 'The army demonstrated what their (plural) dogs can do.'  
 (Condiv)  
 d. Vrijdagmiddag vertrok een delegatie van AG'85 per bus naar Venlo. Daar werd ze ontvangen door de familie Veniger.

‘Friday at noon an AG’85 delegation left for Venlo by bus. There she (singular feminine) was welcomed by the family Veniger.’

(INL)

The main corpus used was the 47 million words Condiv corpus (see Grondelaers et al. 2000), but for less frequent collective nouns corpus material was added from the 38 million words INL corpus (see Kruyt and Dutilh 1997).<sup>22</sup> All the data are presented in Appendix 1. From the analysis of these data, three basic conclusions can be drawn.

A first observation of interest is that collective nouns can differ significantly in verbal and pronominal singular-plural oppositions. Not all collective nouns pattern with singular or plural verbs and pronouns to the same extent. Some prefer a singular, others a plural, and still others are less straightforward. The difference between *duo* (‘duo, pair’), the collective noun with the highest mean percentage of plural forms (i.e., mean A, see Appendix 1), and *regering* (‘government’), the collective noun with the lowest percentage, is almost 40%. In mean B (possessive and personal pronouns only) the difference is even close to 80%. If the eighteen collective nouns are divided into three groups — high plural concord (mean B higher than 50%), medium plural concord (mean B between 50 and 30%) and low plural concord (mean B lower than 30%) (see Table 3) — then the singular-plural oppositions for possessive and personal pronouns are highly significant ( $p < 0.001$ ).<sup>23</sup>

Table 3. *Verbal and pronominal singular-plural oppositions*

	High plural concord (5) (Mean B >50%)				Medium plural concord (6) (Mean B 50–30%)				Low plural concord (7) (Mean B <30%)			
	Mean A (total): 37.6% Mean B (total): 70.9%				Mean A (total): 18.1% Mean B (total): 35.3%				Mean A (total): 7.9% Mean B (total): 15.2%			
	Singular		Plural		Singular		Plural		Singular		Plural	
	n	%	n	%	n	%	n	%	n	%	n	%
Verbs	1503	98.4	24	1.6	2316	99.7	7	0.3	2943	99.6	11	0.4
Relative pronouns	161	93.1	12	6.9	473	98.5	7	1.5	583	99.0	6	1.0
Possessive pronouns	63	32.3	132	67.7	128	88.3	17	11.7	429	96.0	18	4.0
Personal pronouns	70	25.9	200	74.1	99	41.1	142	58.9	269	73.7	96	26.3

Table 4. Property distribution and verbal and pronominal singular-plural oppositions compared

<i>duo</i> 'duo, pair' (86.1%)	3	<i>firma</i> 'firm' (32.4%)	1
<i>echtpaar</i> 'married couple' (79.3%)	3	<i>koor</i> 'choir' (32.3%)	1
<i>bemannig</i> 'crew' (72.0%)	3	<i>leger</i> 'army' (26.7%)	1
<i>gezin</i> 'family, household' (63.6%)	3	<i>club</i> 'club' (25.9%)	1
<i>familie</i> 'family' (58.2%)	2	<i>bond</i> 'union' (20.2%)	1
<i>bende</i> 'gang' (43.4%)	2	<i>vereniging</i> 'association' (14.3%)	1
<i>delegatie</i> 'delegation' (37.7%)	2	<i>maatschappij</i> 'company' (12.1%)	1
<i>team</i> 'team' (34.6%)	2	<i>comité</i> 'committee' (9.2%)	1
<i>publiek</i> 'public' (33.1%)	3	<i>regering</i> 'government' (7.7%)	1

In the first group (high plural concord) both possessive and personal pronouns are generally plural (respectively 67.7% and 74.1%). For the middle group (medium plural concord) that does not hold true: possessive pronouns are mostly singular (only 11.7% plural), whereas for personal pronouns plural forms prevail (58.9%). Finally, in the third group (low plural concord) both possessive and personal pronouns are predominantly singular (respectively 4.0% and 26.3% plural).

Our second conclusion is even more significant than the first: there are strong parallels between the gradient sketched in Section 3 and the gradient emerging from the corpus data (see Appendix 1). If the two are compared (see Table 4), then it appears that type 3 collective nouns typically have the highest plural concord scores, whereas type 1 collective nouns have the lowest. Type 2 collective nouns are situated in between. This indicates an important tendency: in general, high member level accessibility and high plural concord go together. The two are obviously correlated.

Correspondences between the two gradients are not always one-to-one, though. The position of *publiek* ('public') in Table 4, for instance, is lower than one would expect from its type 3 status in property distribution. Broadly speaking, however, both gradients match each other very well.<sup>24</sup>

Thirdly, our corpus data in Appendix 1 and Table 3 provide ample evidence for a linguistic phenomenon that Corbett (1979) has named the "Agreement Hierarchy". Corbett (1979: 203) claims that "as syntactic distance increases, so does the likelihood of semantic agreement" and outlines the following hierarchy, in which the likelihood of plural concord with a (syntactically) nonplural subject monotonically increases from left to right:

(8) attributive — predicate — relative pronoun — personal pronoun

One of Corbett's examples is the English noun *committee*. Attributively, *committee* can only be combined with the syntactic, singular *this* (*this* /

\**these committee*), but the other three positions in the Agreement Hierarchy allow both “syntactic agreement” and “semantic agreement” (*the committee is / are; the committee that / who; the committee . . . it / they*). The more to the right a position is in the hierarchy, the likelier a plural form becomes.

Our corpus data provide evidence for a very similar hierarchy:

- (9) verb — relative pronoun — possessive pronoun — personal pronoun<sup>25</sup>

Indeed, in Appendix 1 and Table 3 plural percentages monotonically rise from left to right.<sup>26</sup> If the eighteen collective nouns examined are taken together, there is a gradual increase of plural concord, from verbs (0.6% plural) and relative pronouns (2.0%) over possessive pronouns (21.2%) to personal pronouns (50.0%) (see Table 5):<sup>27</sup>

Table 5. *Verbal and pronominal singular-plural oppositions (total)*

	<i>Singular</i>		<i>Plural</i>	
	n	%	n	%
Verbs	6762	99.4	42	0.6
Relative pronouns	1217	98.0	25	2.0
Possessive pronouns	620	78.8	167	21.2
Personal pronouns	438	50.0	438	50.0

It appears that some linguistic phenomena are more straightforward indicators of member level accessibility than others. As previously mentioned, differences are particularly pronounced for possessive and personal pronouns. Verbs and relative pronouns show little variation: they are almost exclusively singular, irrespective of the degree of member level accessibility.<sup>28</sup> The following plural examples are therefore rare exceptions, exceptions that many language users consider to be “ungrammatical” and that usually occur in fairly informal subcorpora:

- (10) a. Toen bleek dat het duo het slot van de tweewieler probeerden door te knippen verwittigde hij de rijkswacht.  
‘the duo/pair tried (past third-person plural)’  
(Condiv)
- b. In elk geval zullen de bemanning een tussentijds advies over zulke manoeuvres krijgen.  
‘the crew will (third person plural)’<sup>29</sup>  
(INL)

- c. [...] pas was er een echtpaar die 72 jaar getrouwt [sic] waren dus.  
'a married couple who (plural)'  
(Condiv)
- d. Windows95 was bedoeld voor het huis tuin en keuken publiek, die hooguit enkele uren per dag spelletjes spelen en nog wat kleinschalige andere zaken.  
'the public who (plural)'  
(Condiv)

Differences between the two ends of the gradient, i.e., high and low plural concord (and high and low member level accessibility), are most obvious for possessive pronouns ( $\chi^2 = 307.4$  vs  $\chi^2 = 142.3$  for personal pronouns). If the middle of the gradient is included in the analysis (see Table 3), then a general picture emerges in which possessive pronouns are the most distinctive ones in the "high area" (mean B >50%), whereas personal pronouns are more distinguishing in the "low area" (mean B <30%).<sup>30</sup> As said before, it is possible to discern three main groups in the gradient of member level accessibility: from predominantly plural possessive and personal pronouns (e.g., *echtpaar* 'married couple') over singular possessive, but plural personal pronouns (e.g., *team* 'team') to singular possessive and personal pronouns (e.g., *regering* 'government').

The distinction between (singular) verbs and relative pronouns, on the one hand, and (singular or plural) possessive and personal pronouns, on the other hand, can be argued to correlate with a distinction between syntactic and conceptual individuation. Whether the member level of a Dutch collective noun is profiled (e.g., *het gezin slaapt* 'the family sleeps') or not (e.g., *het gezin is voltallig* 'the family is complete'), verbs and relative pronouns are nearly always singular. Their form is governed not so much by conceptualization as by syntactic principles. Possessive and personal pronouns, on the other hand, appear to be indicative of conceptual individuation, as the gradients sketched in Sections 3 and 4 exhibit a high degree of similarity (see Table 4).<sup>31</sup>

A discrimination between syntactic and conceptual individuation is supported by experimental evidence. By means of a sentence completion task, Bock et al. (1999: 330) show that "number features of pronouns may be retrieved under control from the speaker's meaning, while the number features of verbs are more likely to be retrieved under control from the utterance's form". That means that accounts of concord phenomena that are either exclusively conceptual (Pollard and Sag 1988) or exclusively syntactic (Perlmutter 1972) in nature, should be seriously questioned.<sup>32</sup>

Whereas, at first sight, our distinction between syntactic individuation (verbs and relative pronouns) and conceptual individuation (possessive and personal pronouns) resembles Corbett's (1979) distinction between what he calls "syntactic agreement" and "semantic agreement", we think that the way in which he interprets the latter pair is not entirely satisfying. As far as collective nouns are concerned, Corbett puts "syntactic agreement" on a par with *singular*, and "semantic agreement" with *plural*. In doing so, he seems to subscribe to a view in which collective nouns are singular in form but plural in meaning (see Section 1). As demonstrated above, such a view can obviously be contradicted: collective nouns have a complex conceptual structure in which both unity (singularity) and multiplicity (plurality) are involved. Consequently, a singular personal pronoun should not be analyzed as "syntactic agreement", but as (singular) "semantic agreement" (i.e., profiling of the collection level). In Table 6 the two different views are schematically represented for Dutch collective nouns:

Table 6. Corbett (1979) vs. Joosten (2003)<sup>33</sup>

	Verb	Relative pronoun	Possessive pronoun	Personal pronoun
Corbett (1979)	syntactic (singular)	syntactic (singular)	syntactic (singular)	syntactic (singular)
	*conceptual (plural)	*conceptual (plural)	conceptual (plural)	conceptual (plural)
Joosten (2003)	syntactic (singular)	syntactic (singular)	conceptual (singular or plural)	conceptual (singular or plural)

There is, however, one important caveat in this discussion. Much in the same way that a verb or relative pronoun (combined with a singular collective noun) can occasionally be plural (see [10a]–[10d]), it is marginally possible that the form of a possessive or personal pronoun is at variance with that which would be expected on the basis of conceptual individuation:

- (11) De bemanning kon niet helpen, omdat ze te nat en uitgeput was.  
Literally: 'The crew could not help, because she was too wet and exhausted.'<sup>34</sup>  
(INL)

In (11) the singular personal pronoun *ze* clashes with the properties *nat* ('wet') and *uitgeput* ('exhausted'), that — like *dronken* 'drunk(en)' or *blond* 'blond' — generally cause profiling of the member level. The same



goes for the singular possessive pronoun *zijn* ('its') in a noun phrase such as *het echtpaar zijn dronkenschap* ('the married couple its drunkenness').

It is advisable, therefore, to regard the difference between syntactic and conceptual individuation as an overall distinction. Sometimes conceptual factors DO play a role in the use of verbs or relative pronouns (plural verbs or relative pronouns for a singular collective noun with a highly profiled member level), and in fact, it DOES happen that possessive and personal pronouns are determined syntactically (singular possessive and personal pronouns in a context of obvious member profiling). Without indications to the contrary, however — i.e., in the great majority of contexts — one may safely assume that verbs and relative pronouns individuate syntactically, whereas possessive and personal pronouns individuate conceptually. The fact that the singular-plural proportions for the former pair are close to 100–0, whereas those of the latter pair are not, confirms that syntactic individuation is a black-or-white issue when compared to conceptual individuation, and that collective nouns are more than just “notionally plural”.

In brief, the corpus data in this section provide additional evidence for a gradient of member level accessibility. And there is the added advantage that research on verbal and pronominal singular-plural oppositions offers a way to measure or quantify those conceptual differences. The question remains, however, to what extent the singularity or plurality of verbs and pronouns is affected by the context in which the collective noun is used. As noted in Section 2, conceptual profiling is the end result of quite a few intertwining factors. In Section 3, which dealt with variation in property distribution, contextual influences were minimal: the selected properties were held constant for all collective nouns, so that interpretational differences could be lexically and conceptually motivated as differences in member level accessibility. In this section, however, contexts were far from identical. This problem is addressed in the following section.<sup>35</sup>

## **5. Variation in pronominal singular-plural oppositions: experimental research**

When compared to our intuitive method used in Section 3 (property distribution), corpus research has a number of advantages. First, corpus sentences are not invented; they are instances of real, actual language use. Second, they have been produced by various language users (in different sorts of registers), which considerably increases the objectivity and the general validity of the language data. And thirdly, they can be easily

collected from large electronic text collections, which facilitates computational and statistical processing.

Yet, corpus analyzes have one important drawback: it is fairly difficult, if not impossible, to control contextual factors. In (12a) and (12b), for instance, the plural verbs are not so much caused by conceptual or lexical factors (i.e., a collective noun with a highly accessible member level) as by a plural noun phrase in apposition (*goede vrienden van ons* ‘good friends of ours’, 12a) and a plural complement (*goede mensen* ‘good people’, 12b):

- (12) a. Een computerloos [...] echtpaar, goede vrienden van ons, zitten in de volgende situatie [...].  
 ‘A computerless married couple, good friends of ours, are in the following situation.’  
 (Condiv)
- b. Je gaat er hier maar gewoon vanuit dat de familie Frank geen goede mensen waren.  
 ‘You just assume that the Frank family were’t good people.’  
 (Condiv)

In (13a) and (13b), on the other hand, the context blocks a plural possessive and a plural personal pronoun:

- (13) a. De voetbalclub was in zijn voortbestaan bedreigd.  
 Literally: ‘The football club was threatened in his existence.’  
 (Condiv)
- b. Slechts wanneer een dergelijke vereniging over een overheids-toelating beschikt, mag ze opgericht worden.  
 Literally: ‘Only if such an association has a governmental permission, she may be founded.’  
 (Condiv)

The existence and foundation in question both apply to the collection level, not to the individual members, so that a plural pronoun becomes impossible.<sup>36</sup>

In (14a) and (14b) distance seems to be an interfering factor. The small distance between collective noun and pronoun in (14a) yields a singular pronoun, the greater distance in (14b) a plural pronoun:<sup>37</sup>

- (14) a. Eenmaal een gezin weet hoeveel het kan besteden, kan het veel gericht op zoek naar een huis, bouwgrond of architect.  
 ‘As soon as a family knows how much it can spend, it can search for a house, a building plot or an architect in a more direct way.’  
 (Condiv)

- b. Daarnaast heeft het gezin een oppas die twee halve dagen per week op de kinderen let. De rest van de tijd zijn ze er zelf: Leonoor heeft vier werkdagen en Paul flexibele werktijden.  
'Besides that the family has a babysitter who takes care of the children two half days a week. The rest of the time they are at home: Leonoor works four days a week and Paul has flexible hours.'  
(Condiv)

In the corpus data analyzed in Section 4, these and other contextual factors — see Geerts (1977), Levin (2001), and Depraetere (2003) for a more elaborate discussion — have not been taken into account. In other words, the examples given above are included in the data set and possibly distort the results. Since the contexts in which the collective nouns have been analyzed are not identical, it is in principle possible that differences in concord are partly due to contextual phenomena. Possibly, some collective nouns occur more readily in member profiling contexts, whereas others prefer contexts in which the collection level is profiled. Apart from the fact that such an observation can still be extremely interesting, we think that the combined approach of Section 3 (property distribution) and 4 (corpus research) provides sufficient guarantee for a plausible conceptual notion of member level accessibility.

Nevertheless, a method yielding similar results while excluding contextual influence as far as possible would provide even stronger evidence. If it can be proven that contextual factors do not play a major role (in contrast to what Vossen 1995 seems to assume), then the gradient of member level accessibility becomes irrefutable.

In principle, there are two ways to eliminate contextual effects. A first option is to limit oneself to neutral contexts by excluding from the corpus data all the instances in which the context necessitates collection or member profiling. A second option is, starting from the results obtained, to develop a new, additional method in which contextual influences are controlled. Depraetere (2003) chooses the first option, we prefer the second.

To begin with, we believe that the effect of filtering or not filtering the corpus in Section 4 is minor: the examples in (12) and (13) have been well selected, but they seem to constitute a fairly small minority in the data set. Furthermore, defining and selecting neutral contexts is a “delicate matter, [...] open to a certain amount of argument” (Depraetere 2003: 106). It is far from obvious, for instance, what should be analyzed as a context that triggers collection level interpretations or as a context that yields the profiling of a member level. Nor is it clear how a factor such as distance in (14) could be controlled adequately. As most of the contextual

factors are relative rather than absolute, we prefer a method in which the contexts are identical.

A psycholinguistic eye-tracking experiment in which the context is held constant will demonstrate that reading behavior can provide cues about member level accessibility in Dutch collective nouns. A first subsection (5.1) discusses the hypotheses and the design of the experiment, a second subsection (5.2) analyzes the results.

### 5.1. Eye-tracking experiment: hypotheses and design

5.1.1. *Eye tracking.* Eye tracking is a sophisticated psycholinguistic experimental technique in which participants are asked to read short texts on a computer screen, while their eye movements are precisely recorded by an eye camera. The technique enables the researcher to investigate how much time participants need to read a certain passage, where and when there are fixations, and whether or not certain parts of the texts are reread. Many studies have shown that eye movements, such as fixations or regressions, are highly related to cognitive processing difficulties (see Rayner 1998 for an overview). For instance, a low frequency word in a language will generally yield longer fixation times than a high frequency word.

5.1.2. *Hypotheses.* The eye-tracking method is used to examine the extent to which the postulated differences in member level accessibility correlate with reading time differences for plural (possessive) pronouns. Our hypothesis is that a plural pronoun, when it is combined with a collective noun of low member level accessibility (*bond* ‘union’, *vereniging* ‘association’, *club* ‘club’), is more “problematic” — and therefore yields slower reading times — than a singular pronoun. As the individual members for *bond* type collective nouns (type 1) are far from easily accessible, it is likely that their profiling requires a lot of cognitive effort and therefore causes slow reading times. For collective nouns with a highly accessible member level (*echtpaar* ‘married couple’, *bemannig* ‘crew’, *gezin* ‘family, household’), on the other hand, our hypothesis is that reading time differences between singular and plural pronouns are far less pronounced. In short, the plural pronoun *hun* (‘their’) should be more problematic than singular *zijn* (‘his/its’) for *bond* (‘union’), but not for *echtpaar* (‘married couple’), or at least not to the same extent. Even more, reading times for highly accessible collective nouns are possibly faster for plural than for singular pronouns. As appears from the data in Appendix 1, type 3 collective nouns are rather ‘more than one’ than ‘one’: mean B for type 3 collective nouns is always higher than 50%.<sup>38,39</sup>

The dependent variable in this experiment is the reading time in milliseconds (ms) needed for a collective noun and a plural pronoun. Our assumption is that reading times will reflect the conceptual accessibility of the member level: the faster the reading times are, the more accessible the member level is. What is still unclear, however, is where exactly and in which guise (longer or shorter fixation times? more or fewer regressions?) differences in processing ease will appear.

5.1.3. *Selection of collective nouns and pronouns.* From the gradients sketched in Sections 3 and 4 twelve collective nouns were retained for the experiment: six type 1 collective nouns (*bond* ‘union’, *club* ‘club’, *firma* ‘firm’, *organisatie* ‘organization’, *partij* ‘party’, *vereniging* ‘association’) and six type 3 collective nouns (*bemannig* ‘crew’, *echtpaar* ‘married couple’, *gezin* ‘family, household’, *kliek* ‘clique’, *trio* ‘trio, threesome’, *tweeling* ‘twins’). As five of the twelve collective nouns (*organisatie*, *partij*; *kliek*, *trio*, *tweeling*) have not been analyzed in corpora, the selection provided an extra check on the validity of the (property-based) gradient sketched in Section 3.

For all twelve collective nouns two different text fragments were written, each in two variants: one with a singular possessive pronoun (*zijn* ‘his/its’ or *haar* ‘her’) and one with a plural possessive pronoun (*hun* ‘their’). That gave a total of ( $12 \times 2 \times 2 =$ ) 48 texts. There were two independent variables — member level accessibility and possessive pronouns — with two conditions each, respectively high vs. low and singular vs. plural. Possessive pronouns were preferred to personal pronouns because they are more distinctive overall, as shown in Section 4.

5.1.4. *Composition of text fragments.* The 24 different text fragments (in two variants each) were all similar in length (five lines of not more than 75 characters) and, what is more important, they were neutral with regard to profiling of the collection or member level. In other words, they did not contain expressions such as *de vereniging haar tiende verjaardag* (literally, ‘the association her tenth anniversary’) or *het echtpaar hun dronkenschap* (literally, ‘the couple their drunkenness’), i.e., expressions that trigger either collection level or member level interpretations. Furthermore, the crucial sentence of the text fragments, i.e., the one that contained collective noun and possessive pronoun, was in all the cases structurally similar. A more detailed description of the composition principles and the text fragments themselves can be found in Appendix 2. By way of introduction, this is a random example with *bond* (‘union’, low member level accessibility) and a singular possessive pronoun:

(15)

Jarenlang leek het een kat-en-muisspel waarbij geld, macht en corruptie de overhand zouden krijgen. Tot de uitspraak van vandaag. Een topadvocate bezorgde de bond vanmiddag zijn eerste nederlaag. Maar de zaak is nog niet beslecht, aangezien de bond van plan is in beroep te gaan. Volgens de meeste gerechtsdeskundigen blijven de kansen in dat geval fifty-fifty.<sup>40</sup>

5.1.5. *Participants.* Participants were 40 Flemish first-year students at Ghent University, who participated for course credits. They all had normal, uncorrected vision and were native speakers of Dutch.

5.1.6. *Apparatus.* Eye movements were recorded by a Senso-Motoric Instruments (SMI Eyelink) video-based pupil tracking system. Viewing was binocular but eye movements were recorded from the right eye only. A high speed video camera was used for recording. It was positioned underneath the monitored eye and held in place by head-mounted gear. The system has a visual resolution of 20 seconds of arc. Fixation locations were sampled every 4 ms and these raw data were used to determine the different measures of oculomotor activity during reading. The display was 69 cm from the subject's eye and three characters equalled 1° of visual angle. A chin rest was used to reduce head movements during the experiment.

5.1.7. *Procedure.* Before the experiment started, participants were informed that the study was about reading comprehension of short texts that would be displayed on a screen. Text administration was self-paced. The passages of the text were presented as a whole. Participants indicated when they had finished reading the text passage by pressing a button. They were told to read at their normal rate and that periodically they would be asked to answer a comprehension question about the passages. This was done on one-fourth of the trials. The participants had no difficulty answering the questions; the questions were simple true-or-false statements, and the participants were correct 87% of the time. The initial calibration of the eye-tracking system generally required approximately 10 min and consisted of a standard nine-point grid. Following the initial calibration the participant was given 10 practice trials to become familiar with the procedure before reading the experimental text fragments. The 24 experimental text fragments were embedded in a pseudo-random order in 120 filler text fragments. Each participant was presented one of the two possible variants of the text fragment according to a Latin square design.<sup>41</sup> Participants completed one session lasting about one hour, containing 144 pieces of text to read.

5.1.8. *Regions.* In the crucial third line of each text fragment reading times were recorded in four distinct regions. Let us resume the third line of the text in (15):

(16)

bezorgde de bond vanmiddag zijn eerste nederlaag. Maar de zaak is nog niet

In that line we distinguished the following four regions:

- (17) Region 1: bezorgde de bond vanmiddag  
gave the union (m.) this afternoon  
Region 2: zijn  
his  
Region 3: eerste nederlaag.  
first defeat  
Region 4: Maar de zaak is nog niet  
but the case is not yet

The other texts, all structurally identical (see 4. Composition of text fragments), were divided into the same four regions. Since the processing time of the (singular or plural) possessive pronoun is the focus of the study, region 2 and 3 will be of primary importance. Region 4 was included in order to be able to analyze possible late-time effects; region 1 will probably not be of interest, but contains the collective noun and was therefore analyzed separately, as a routine check.

5.1.9. *Eye movement measures.* In the four regions performance data were recorded by making use of four types of eye movement measures: *first fixation duration* (FFD), *gaze duration* (GD), *total fixation duration* (TFD), and *cumulative regression reading time* (CRRT) (see Figure 3). FFD measures the duration of the first fixation in the region, GD the sum of all fixations in the region during the first entry. TFD is the sum of all fixations in the region, no matter whether those fixations were produced during the first or a later entry. CRRT, finally, measures the duration of all fixations in a region, from the first fixation in the region until the first fixation in the following region. That implies that the duration of a regression starting in this region is included.

## 5.2. *Eye-tracking experiment: results and discussion*

In total, 40 participants with 24 texts each yields 960 trials. Of those 960 trials 53 (i.e., a normal proportion of 5.5%) were invalid due to calibration errors. The remaining data were subjected to an analysis of variance

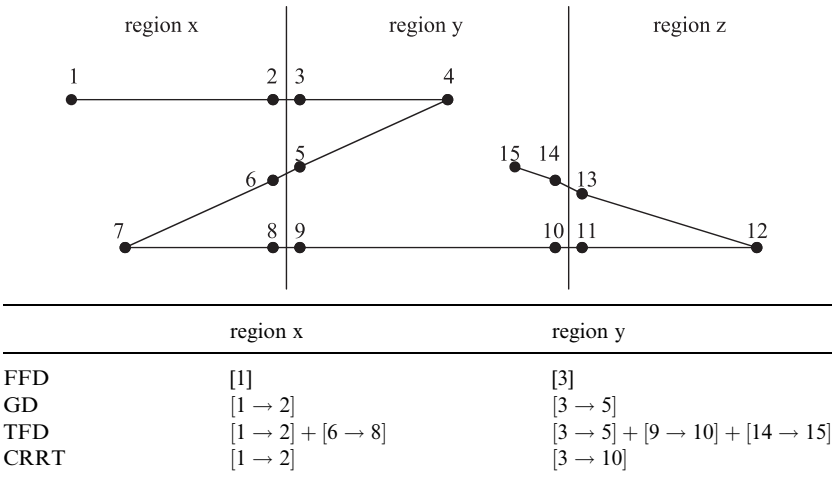


Figure 3. Four eye-movement measures

over subjects ( $F_1$ ) and stimuli ( $F_2$ ) and a planned comparison test. All results are presented in Appendix 3.

Analysis of the data shows the third region to be especially revealing. Table 7 presents the average TFD times in region 3 for the four distinct conditions:

Table 7. TFD in ms for region 3

	Possessive singular ( <i>zijn</i> 'his/its' / <i>haar</i> 'her')	Possessive plural ( <i>hun</i> 'their')
Low member level accessibility (type 1)	890	947
High member level accessibility (type 3)	738	759

Table 7 shows that sentences in which type 1 collective nouns (low member level accessibility) are combined with a plural possessive pronoun require more cognitive effort than sentences in which those collective nouns pattern with a singular pronoun. The difference in processing time between both conditions (57 ms) is significant ( $F_1(1, 39) = 4.12, p < 0.05$ ).<sup>42</sup>

For type 3 collective nouns (high member level accessibility), there is no such difference. Admittedly, average reading times for plural pronouns are slightly slower than for singular ones, but the difference between the two conditions (21 ms) is far from significant ( $F_1(1, 39) < 1; F_2(1, 11) < 1$ ).

Both observations confirm our hypothesis in 5.1: since the individual members are conceptually far from easily accessible for type 1 collective



nouns (*bond* ‘union’), but not for type 3 collective nouns, a plural possessive pronoun requires more cognitive processing time than a singular pronoun in the former case, but not in the latter.

The difference between low and high member level accessibility is even more pronounced if the second region (i.e., the region consisting solely of the possessive pronoun) is added to our analysis. Table 8 gives a survey of the average TFD times in the second region:

Table 8. TFD in ms for region 2

	Possessive singular ( <i>zijn</i> ‘his/its’ / <i>haar</i> ‘her’)	Possessive plural ( <i>hun</i> ‘their’)
Low member level accessibility (type 1)	257	248
High member level accessibility (type 3)	281	245

In 5.1 we (tentatively) hypothesised that, as far as highly accessible type 3 collective nouns are concerned, the plural possessive pronoun *hun* (‘their’) might yield even faster reading times than its singular counterpart. Table 8 appears to confirm this hypothesis: for collective nouns such as *echtbaar* ‘married couple’, singular TFD times are considerably slower than plural ones. The difference between the two conditions (36 ms) is significant ( $F_1(1, 29) = 5.22, p < 0.05$ ).<sup>43</sup> For type 1 collective nouns such as *bond* (‘union’) (low member level accessibility), plural pronouns are also faster than singular ones, but that difference (9 ms) is probably due to the difference in word length between, on the one hand, singular *zijn* (‘his/its’) or *haar* (‘her’) (4 characters), and, on the other hand, plural *hun* (‘their’) (3 characters).<sup>44</sup> It is clearly not significant ( $F_1(1, 33) = 1.35, p > 0.25$ ;  $F_2(1, 9) = 1.04, p > 0.30$ ).

Since only TFD times in zone 2 and 3 are significant, Tables 7 and 8 lead to the following general picture. Type 3 collective nouns (high member level accessibility) yield faster processing times for plural possessive pronouns than for singular ones; singular pronouns cause regressions from and to region 2. Further on, in region 3, reading time differences between the two conditions are no longer significant. Type 1 collective nouns (low member level accessibility) exhibit the opposite pattern: since the individual members are not easily accessible, singular possessive pronouns are processed faster than plural ones. Plural pronouns, which in region 2 are usually skipped, lead to regressions from and to region 3.

It appears that the basic claim of this article — that collective nouns can differ with regard to the conceptual accessibility of their member level — is confirmed by experimental research as well. However, the experimental

data mentioned above should be treated and interpreted with caution. Two remarks may help to put the results into perspective.

First, the effects discussed above appear to be quite subtle. That is evident from the fact that just one of the four eye movement measures, viz. TFD, turns out to be significant: only two of the 16 cells in Appendix 3 (4 measures  $\times$  4 regions), both TFD, yield significant results. TFD is a fairly general measure: it is the sole measure that possibly consists of discontinuous time recordings (see Figure 3). It seems, therefore, that reading time differences for plural pronouns and the relatively fast or slow conceptual “activation” of the member level cannot be pinned down to a certain region or be demonstrated by a systematically recurring type of eye movement. The different TFD times point to regressions from and to regions 2 and 3, but the nature of those regressions is not entirely systematic. Apparently, the phenomenon discussed is determined by individual differences: some participants produce longer fixations on the crucial regions, others predominantly make regressions, whereas still others scarcely have fixations or regressions, but reread the entire sentence at the end. Such individual differences make mean that FFD, GD, and CRRT yield insignificant results and that only the most general measure — the total fixation duration, no matter whether those fixations were recorded during the first or a later entry — is indicative of conceptual accessibility. Long total fixation duration hints at a general difficulty in conceptual activation, short duration does not. In fact, the subtlety of the reading effects is not really surprising: since the texts are only minimally different, viz. in one single word (*zijn* ‘his/its’ / *haar* ‘her’ vs *hun* ‘their’), it would seem unrealistic to expect highly differentiated results.

Our second note pertains to the differences between the individual participants. In order for any eye-tracking experiment to find significant effects in eye movements, it needs to be based on as many stimuli as possible. Only in that way is it possible to compensate for the variance in the data. It seems, therefore, that an enlargement of our set of stimuli could yield even more significant results. Though the  $F_1$  analyzes demonstrate a real and significant effect, the  $F_2$  analyzes do not reach a level of high significance ( $p < 0.05$ ). That hints at a power problem: the analysis over subjects ( $F_1$ ) confirms the hypotheses formulated in 5.1, but the analysis over stimuli ( $F_2$ ) indicates no more than a tendency. The hypotheses cannot yet be irrefutably confirmed due to the limited number of stimuli (six for each condition for each participant).

A follow-up experiment should therefore take the extension of the stimuli as its top priority. The extension could be achieved in two (complementary) ways: either by adding new examples to the list of collective nouns to be examined, or by extending the number of texts presented for

each collective noun. For the type 3 collective nouns (high member level accessibility) it might be interesting to make a division according to the feature ‘number of members’. The corpus data in Section 4 show that collective nouns with a fixed and small number of members (e.g., *duo* ‘duo, pair’, *echtpaar* ‘married couple’) have the highest scores for plural pronouns, higher than the scores for other type 3 collective nouns. Given the limited number of stimuli in the experiment described above, it does not seem feasible to split the data for type 3 collective nouns into two small groups (*echtpaar* ‘married couple’, *trio* ‘trio, threesome’, *tweeling* ‘twins’ vs *bemannig* ‘crew’, *gezin* ‘family’, *klied* ‘clique’), but if the set of stimuli were extended, such a split could be considered. Furthermore, it might be worthwhile to involve Dutch participants in the experiment, so that the results would be generalizable to the entire Dutch language area. We do not expect geographical differences, though, as such differences could not be inferred from the (partly Flemish, partly Dutch) corpus material in Section 4 either.

## 6. Questions for further research

The discussion on variable member level accessibility in (Dutch) collective nouns raises some interesting questions for further research. In what follows, we will outline what we see as the three most challenging questions.

### 6.1. *Conceptual motivations*

Describing and analyzing differences in member level accessibility is one thing, explaining them is another. Searching for conceptual motivations for the observations made in the previous sections, then, could be a first interesting direction for further research.

There are several possible lines of thought concerning motivations. First, as argued in Sections 4 and 5, it appears that the number of members could have an effect on member level accessibility: the smaller the number of individual members, the higher is the chance that they are conceptually profiled. Not surprisingly, collective nouns such as *duo* (‘duo, pair’) and *echtpaar* (‘married couple’) have the highest plural concord scores in our corpus data.<sup>45</sup> Conversely, it is plausible that a large number of members decreases the possibility of a highly accessible member level.<sup>46</sup> However, quite a number of differences in member level accessibility (e.g., *bemannig* ‘crew’ vs. *comité* ‘committee’; *publiek* ‘public’ vs. *vereniging* ‘association’) cannot be motivated in terms of the size of the collection.

A second possible line of thought might be to examine whether or not the individual members are involved in the collection's "origin process".<sup>47</sup> Some collections (e.g., *bemannig* 'crew') can only be composed (Dutch *samenstellen*), not founded (Dutch *oprichten*); others (e.g., *team* 'team') can both be composed and founded; and still others (e.g., *vereniging* 'association') can only be founded. If something is founded, there is no conceptual focus on the "building blocks" that one needs. If something is composed, however, there is such a focus. Not surprisingly, Geerts and Den Boon (1999: 2925) define *samenstellen* ("to compose") as "uit verschillende bestanddelen tot een geheel maken" [to make a whole out of parts; our underlining]. That is also why (18a) sounds perfectly normal, whereas (18b) does not:

- (18) a. We hebben een vereniging opgericht, maar wie de leden zijn staat nog niet vast.  
'We have founded an association, but the members are still unknown.'
- b. ?We hebben een bemanning samengesteld, maar wie de leden zijn staat nog niet vast.  
'?We have composed a crew, but the members are still unknown.'

Collections that do not require conceptual profiling of the members during their "origin process", are likely candidates for low member level accessibility in general.

The two tendencies mentioned above are only meant to be possible lines of further research; they are intuitively plausible but partial motivations and certainly require a more in-depth analysis. In that respect, one could ask to what extent linguists are able to motivate conceptual differences anyway. It is not unlikely that cooperation with other scientific disciplines such as psychology proves to be indispensable. Most interestingly, in the domain of social psychology quite a lot of research has centered on the question of how the human mind processes information about (types of) groups and individual persons.<sup>48</sup> Hamilton and Sherman (1996) offer an excellent survey. As appears from their discussion, a fundamental difference between individual persons and groups is that individual persons are in general considered to be more coherent; they are — more than groups — expected to show unity, coherence and consistency. Inconsistent behavior of a single person is, therefore, far more surprising than the same type of inconsistency within a group.

Interestingly enough, groups can be of different sorts. Campbell (1958) introduces the term "entitativity" for the degree to which a group is a coherent unity, so that:

a band of gypsies is empirically harder, more solid, more sharply bound than the ladies aid society, and the high-school basketball team [...] falls somewhere in between (Campbell 1958: 18).

Some groups are “entitative”, i.e., resemble individual persons, whereas others are not or to a lesser extent. That is also evident from the way in which information about groups is obtained and processed. Hamilton and Sherman (1996: 345) argue:

[...] that groups vary in the extent to which the perceiver assumes this kind of unity among their elements. Groups that are high in perceived entitativity are assumed to have unity and coherence, and their members are expected to show consistency among them. For such groups, like a fraternity, information about group members would be dealt with by processes similar to those engaged in forming impressions of individuals. The perceiver would assume consistency, would seek organization among the elements, would make online inferences about the group, inconsistencies would be surprising and likely to trigger attributional thinking, and so forth, just as in forming impressions of individuals. For other groups, the perceiver would presume less unity or entitativity and, hence, would be less likely to engage in these processes. In this case, one would see less evidence of an organized representation of the information about group members, judgments would more likely be memory based, and inconsistencies would be less likely to trigger any special processing.

The parallels with our description of member level accessibility are striking and certainly invite a more detailed comparison. On the other hand, the problem in social psychology seems to be similar to the one sketched above: the differences between groups and individual persons have been well *described*, but that does not mean that they have been satisfactorily *explained*. Hamilton and Sherman (1996: 348) admit that:

Still, many questions remain. [...] when we think about the range of social groups that we encounter and perceive in everyday life, what is it that gives some of these groups more essence, makes them more meaningful than others as perceived social units? Is it due to the physical proximity of the members to each other? Or to the interdependence among their members? Or to some common fate that they share? Or is it due to the similarity of the members to each other? If so, then similarity with respect to what? Their heritage? Their appearance? Their personalities? Their interests? It seems plausible that all of these factors probably contribute to the perception of entitativity in a group, at least under some conditions.

Their conclusion is that “there are multiple routes to perceiving entitativity in a group” (1996: 348). More research on those “routes” could provide useful insights and motivations, not only for the social psychological, but also for the linguistic differences. In the meantime it seems worthwhile to relate and compare the descriptions of these two disciplines.

## 6.2. Crosslinguistic perspectives

A second line of future research does not consist in adding to our description a number of conceptually plausible motivations, but in comparing Dutch to other languages. (British) English, for instance, seems to be an ideal starting point for a contrastive approach to member level accessibility. The data gathered in Levin (2001) and Depraetere (2003)<sup>49</sup> seem to suggest a gradient of member level accessibility in English that is quite similar to the one we have sketched for Dutch. Collective nouns such as *government* and *company* pattern significantly less with plural verbs and pronouns than *family* or *team*, and the same holds if we compare the latter pair to *crew* or *couple* ( $p < 0.001$ ):<sup>50</sup>

Table 9. Verbal and pronominal singular-plural oppositions for English collective nouns in The Independent 1995 (Levin 2001: 166)

	Verbs				Relative pronouns				Personal pronouns			
	Singular		Plural		Singular		Plural		Singular		Plural	
	n	%	n	%	n	%	n	%	n	%	n	%
<i>government</i>	345	94.5	20	5.5	16	94.1	1	5.9	77	85.6	13	14.4
<i>company</i>	316	97.8	7	2.2	86	92.5	7	7.5	79	81.4	18	18.6
<i>family</i>	109	63.0	64	37.0	11	42.3	15	57.7	7	10.6	59	89.4
<i>team</i>	91	62.8	54	37.2	21	42.0	29	58.0	15	19.5	62	80.5
<i>crew</i>	24	45.3	29	54.7	11	45.8	13	54.2	2	7.7	24	92.3
<i>couple</i>	9	15.8	48	84.2	0	0.0	25	100.0	0	0.0	41	100.0

Unlike for Dutch collective nouns, differences in member level accessibility in English are likely to show up more to the left of Corbett's (1979) Agreement Hierarchy. The differences in verbal singular-plural oppositions are in general the most significant ones (e.g., *government* / *company* vs. *crew* / *couple*:  $\chi^2 = 365.3$ ), though the data in Table 9 are evidently far too limited to make strong claims.<sup>51</sup>

In short, elaborate contrastive research on verbal and pronominal singular-plural oppositions in English and Dutch could provide even more evidence for our notion of variable member level accessibility, while at the same time it could demonstrate how two historically related languages are different with respect to the Agreement Hierarchy. Whereas for Dutch differences in member level accessibility are only visible in the right part of the hierarchy (possessive and personal pronouns), for English there are indications that they are particularly pronounced for verbs,

at the left side of the hierarchy. That would mean that concord in English is less syntactically (i.e., more conceptually) determined than in Dutch.<sup>52</sup>

### 6.3. *Formal-theoretical implications*

A third line of future research, finally, is an exploration in formal semantics. Collective nouns have been a topic of discussion in quite a few formal-semantic studies (e.g., Landman 1989; Krifka 1991; Dölling 1991, 1995; Barker 1992; Lønning 1997), one of the central questions being whether their “collectiveness” is comparable to the “collective readings” of plural NPs such as *the boys* (e.g., *The boys gather*). However, individual differences between collective nouns have been largely ignored<sup>53</sup> and therefore it seems worthwhile to relate our results to the models and principles in formal semantics, and to investigate whether they have theoretical implications.

*Received 31 March 2004*  
*Revised version received*  
*5 October 2004*

*University of Louvain*  
*University of Ghent*

## **Appendix 1. Corpus results**

For each of the 18 collective nouns selected, absolute and relative frequencies are listed for singular and plural verbs, singular and plural relative pronouns, singular and plural possessive pronouns, and singular and plural personal pronouns. The following principles were taken into account:

- Collective nouns were only considered in their singular form. Collective nouns in the diminutive form were left out of consideration.
- For each collective noun a maximum of 1500 sentences was analyzed.
- Examples in which the collective noun was used in a member identifying construction (e.g., *een team van deskundigen* ‘a team of experts’), were excluded.
- Also excluded from further analysis:
  - Plural non-third-person pronouns (e.g., *onze* ‘our’, *wij* ‘we’, *jullie* ‘you (plural)’), since they did not have a relevant singular counterpart;
  - The personal pronoun *men* ‘one’, since it is neither really singular nor really plural;
  - Pronouns ambiguous with regard to a singular or a plural interpretation (e.g., *de regering die hij verantwoordelijk houdt voor de dood van zijn dochter* (INL). Literally: ‘the government that/whom he holds responsible’);

- The possessive pronoun *hun* ('their') if it was dependent on the plural personal pronoun *ze* ('they'): *De bemanning wist dit gevaar te bezweren door het vuur te blussen en nadat ze hierop hun mijnen alsnog hadden gedropt bereikten ze veilig de thuisbasis* (INL). 'The crew ... after they had dropped their mines'.

Between brackets, absolute frequencies have been distinguished for two corpus components, respectively a formal component and an informal component. Usenet, Internet Relay Chat (Condiv), and youth news (INL) are considered to be informal. All other subcorpora belong to the formal component.

Mean A is the mean value of the four plural percentages. Mean B is the mean value of the last two plural percentages, i.e., the percentages for plural possessive and personal pronouns. Mean A and B give identical gradients. In what follows, the collective nouns have been ordered from high to low.

Condiv	<i>duo</i> ('duo, pair')			
Mean A = 43.7%	Singular		Plural	
Mean B = 86.1%	n	%	n	%
Verbs	326 (322+4)	97.3	9 (8+1)	2.7
Relative pronouns	16 (16+0)	100.0	0 (0+0)	0.0
Possessive pronouns	7 (7+0)	20.0	28 (28+0)	80.0
Personal pronouns	5 (5+0)	7.8	59 (57+2)	92.2

Condiv	<i>echtpaar</i> ('married couple')			
Mean A = 42.9%	Singular		Plural	
Mean B = 79.3%	n	%	n	%
Verbs	232 (228+4)	98.3	4 (3+1)	1.7
Relative pronouns	31 (31+0)	88.6	4 (2+2)	11.4
Possessive pronouns	14 (14+0)	23.3	46 (46+0)	76.7
Personal pronouns	6 (6+0)	18.2	27 (27+0)	81.8



Condiv + INL	<i>bemannig</i> ('crew')			
Mean A = 40.9%	Singular		Plural	
Mean B = 72.0%	n	%	n	%
Verbs	374 (317+57)	98.2	7 (6+1)	1.8
Relative pronouns	14 (13+1)	82.4	3 (2+1)	17.6
Possessive pronouns	4 (4+0)	26.7	11 (9+2)	73.3
Personal pronouns	17 (15+2)	29.3	41 (27+14)	70.7

Condiv	<i>gezin</i> ('family, household')			
Mean A = 31.9%	Singular		Plural	
Mean B = 63.6%	n	%	n	%
Verbs	282 (257+24)	99.6	1 (1+0)	0.4
Relative pronouns	38 (33+5)	100.0	0 (0+0)	0.0
Possessive pronouns	11 (11+0)	47.8	12 (12+0)	52.2
Personal pronouns	7 (7+0)	25.0	21 (16+5)	75.0

Condiv	<i>familie</i> ('family')			
Mean A = 31.2%	Singular		Plural	
Mean B = 58.2%	n	%	n	%
Verbs	289 (260+29)	99.0	3 (1+2)	1.0
Relative pronouns	62 (53+9)	92.5	5 (3+2)	7.5
Possessive pronouns	27 (27+0)	43.5	35 (34+1)	56.5
Personal pronouns	35 (34+1)	40.2	52 (43+9)	59.8

Condiv	<i>bende</i> ('gang')			
Mean A = 22.2%	Singular		Plural	
Mean B = 43.4%	n	%	n	%
Verbs	140 (129+11)	100.0	0 (0+0)	0.0
Relative pronouns	51 (47+4)	98.0	1 (0+1)	2.0
Possessive pronouns	11 (10+1)	84.6	2 (2+0)	15.4
Personal pronouns	8 (7+1)	28.6	20 (17+3)	71.4

Condiv + INL	<i>delegatie</i> ('delegation')			
Mean A = 21.0%	Singular		Plural	
Mean B = 37.7%	n	%	n	%
Verbs	498 (484+14)	99.4	3 (3+0)	0.6
Relative pronouns	35 (33+2)	92.1	3 (3+0)	7.9
Possessive pronouns	17 (16+1)	77.3	5 (5+0)	22.7
Personal pronouns	18 (18+0)	47.4	20 (20+0)	52.6

Condiv	<i>team</i> ('team')			
Mean A = 17.4%	Singular		Plural	
Mean B = 34.6%	n	%	n	%
Verbs	273 (223+50)	99.6	1 (1+0)	0.4
Relative pronouns	160 (128+32)	100.0	0 (0+0)	0.0
Possessive pronouns	20 (16+4)	90.9	2 (2+0)	9.1
Personal pronouns	14 (13+1)	40.0	21 (10+11)	60.0

Condiv	<i>publiek</i> ('public')			
Mean A = 17.4%	Singular		Plural	
Mean B = 33.1%	n	%	n	%
Verbs	443 (374+69)	99.3	3 (2+1)	0.7
Relative pronouns	74 (62+12)	97.4	2 (0+2)	2.6
Possessive pronouns	20 (19+1)	95.2	1 (1+0)	4.8
Personal pronouns	27 (25+2)	38.6	43 (28+15)	61.4

Condiv	<i>firma</i> ('firm')			
Mean A = 16.5%	Singular		Plural	
Mean B = 32.4%	n	%	n	%
Verbs	456 (392+64)	100.0	0 (0+0)	0.0
Relative pronouns	101 (77+24)	99.0	1 (0+1)	1.0
Possessive pronouns	34 (31+3)	89.5	4 (1+3)	10.5
Personal pronouns	16 (14+2)	45.7	19 (13+6)	54.3

Condiv + INL	<i>koor</i> ('choir')			
Mean A = 16.2%	Singular		Plural	
Mean B = 32.3%	n	%	n	%
Verbs	506 (499+6)	100.0	0 (0+0)	0.0
Relative pronouns	52 (52+0)	100.0	0 (0+0)	0.0
Possessive pronouns	26 (24+2)	89.7	3 (3+0)	10.3
Personal pronouns	16 (16+0)	45.7	19 (18+1)	54.3

Condiv	<i>leger</i> ('army')			
Mean A = 14.2%	Singular		Plural	
Mean B = 26.7%	n	%	n	%
Verbs	343 (279+64)	99.7	1 (0+1)	0.3
Relative pronouns	29 (25+4)	96.7	1 (0+1)	3.3
Possessive pronouns	16 (13+3)	80.0	4 (3+1)	20.0
Personal pronouns	16 (15+1)	66.7	8 (2+6)	33.3

Condiv	<i>club</i> ('club')			
Mean A = 13.8%	Singular		Plural	
Mean B = 25.9%	n	%	n	%
Verbs	340 (244+96)	99.4	2 (0+2)	0.6
Relative pronouns	179 (110+69)	97.3	5 (1+4)	2.7
Possessive pronouns	76 (63+13)	97.4	2 (2+0)	2.6
Personal pronouns	31 (20+11)	50.8	30 (11+19)	49.2

Condiv	<i>bond</i> ('union')			
Mean A = 10.1%	Singular		Plural	
Mean B = 20.2%	n	%	n	%
Verbs	435 (382+53)	100.0	0 (0+0)	0.0
Relative pronouns	31 (22+9)	100.0	0 (0+0)	0.0
Possessive pronouns	35 (28+7)	97.2	1 (0+1)	2.8
Personal pronouns	20 (17+3)	62.5	12 (7+5)	37.5

Condiv	<i>vereniging</i> ('association')			
Mean A = 7.1%	Singular		Plural	
Mean B = 14.3%	n	%	n	%
Verbs	453 (408+45)	100.0	0 (0+0)	0.0
Relative pronouns	111 (93+18)	100.0	0 (0+0)	0.0
Possessive pronouns	82 (75+7)	96.5	3 (3+0)	3.5
Personal pronouns	39 (33+6)	75.0	13 (6+7)	25.0

Condiv	<i>maatschappij</i> ('company')			
Mean A = 6.2%	Singular		Plural	
Mean B = 12.1%	n	%	n	%
Verbs	324 (298+26)	99.4	2 (2+0)	0.6
Relative pronouns	82 (70+12)	100.0	0 (0+0)	0.0
Possessive pronouns	66 (63+3)	98.5	1 (1+0)	1.5
Personal pronouns	34 (33+1)	77.3	10 (2+8)	22.7

Condiv	<i>comité</i> ('committee')			
Mean A = 4.7%	Singular		Plural	
Mean B = 9.2%	n	%	n	%
Verbs	400 (394+6)	99.8	1 (1+0)	0.2
Relative pronouns	93 (90+3)	100.0	0 (0+0)	0.0
Possessive pronouns	46 (46+0)	92.0	4 (4+0)	8.0
Personal pronouns	43 (42+1)	89.6	5 (3+2)	10.4

Condiv	<i>regering</i> ('government')			
	Singular		Plural	
Mean A = 4.0%				
Mean B = 7.7%	n	%	n	%
Verbs	648 (576+72)	99.2	5 (0+5)	0.8
Relative pronouns	58 (48+10)	100.0	0 (0+0)	0.0
Possessive pronouns	108 (95+13)	97.3	3 (2+1)	2.7
Personal pronouns	125 (113+12)	87.4	18 (5+13)	12.6

## Appendix 2. Eye-tracking experiment: composition of text fragments

The text fragments printed below were constructed following a strict set of principles. The most important “guidelines” were the following:

1. All text fragments had five lines and each line consisted of not more than 75 characters (spaces included).
2. The contexts used were neutral with regard to collection or member level interpretations (see 5.1).
3. The crucial sentence, i.e., the one that contained collective noun and pronoun, was in all 48 text fragments structurally similar: subject + finite verb + definite collective noun used as a direct object + temporal adverb + indirect object starting with a possessive pronoun.
4. The crucial sentences started at the end of the second line, so that the introduction was always similar in length and — more importantly — so that the collective noun and the rest of the sentence (in which the most relevant reading times were to be expected) were situated in the middle of the third line. That excluded possible distortions caused by the so-called “return sweep”, i.e., the eye movement that returns the participants’ eyes to the beginning of the next line. As the exact “launching” and “landing” sites of return sweeps may vary considerably, it is of vital importance that the relevant parts of the sentence are situated in the middle of a line.
5. In the crucial sentence, the (default) gender of the subject noun was different from the collective noun’s gender (e.g., *koningin* ‘queen’ (feminine) vs. *bond* ‘union’ (masculine)), in order to exclude coreference problems.
6. As variation between singular and plural verbs in English is typical for definite collective nouns (and not for indefinite ones, see Levin 2001: 121–125, Depraetere 2003: 96–97), all collective nouns were used in a definite noun phrase. There was one problem, though: definiteness implied that the collective noun in question had already been introduced. To re-

solve this problem (and to avoid artificial contexts) a context was created in which the reader knew from the beginning that not every definite noun phrase could be traced back in the text.

7. The temporal adverbs separating direct and indirect object in the crucial sentence, were similar as far as frequency (Baayen et al. 1993) and length were concerned: *binnenkort* ‘soon’ (799, 10 characters), *indertijd* ‘at the time’ (727, 9 characters), *vanmiddag* ‘this afternoon’ (726, 9 characters), *gisteravond* ‘yesterday evening’ (704, 11 characters).
8. The indirect object was structurally built up as follows: possessive pronoun + adjective + noun.
9. The head nouns in the indirect object were abstract nouns. It is plausible that in general, concrete nouns would more easily yield member level interpretations.
10. The verbs used in the crucial sentence were typically three-place verbs: *iemand iets ontzeggen* (‘to deny someone something’), *iemand iets vergeven* (‘to forgive someone something’), *iemand iets bezorgen* (‘to give someone something’), and so on.
11. In order to be able to analyze possible late-time effects, all sentences following the crucial sentence started with *Maar de* (‘But the’).

Below, the singular and plural variant of the same text have only been printed once. Possessive pronouns have been separated by ‘/’ and collective nouns have been underlined. Of course, in the experiment itself that was not the case.

Al heeft hij officieel nog niets gehoord, toch maakt de bondsvoorzitter zich ernstig zorgen over een nakende beslissing van het hof. De koning verbiedt de bond binnenkort zijn / hun zondagse activiteiten. Maar de rest van het koningshuis is het daar absoluut niet mee eens. Zij menen dat de loopwedstrijden geen hinder veroorzaken en dat een verbod dus te zwaar is.

Jarenlang leek het een kat-en-muisspel waarbij geld, macht en corruptie de overhand zouden krijgen. Tot de uitspraak van vandaag. Een topadvocate bezorgde de bond vanmiddag zijn / hun eerste nederlaag. Maar de zaak is nog niet beslecht, aangezien de bond van plan is in beroep te gaan. Volgens de meeste gerechtsdeskundigen blijven de kansen in dat geval fifty-fifty.

Na weken gebakkelei en gegalaver is er nog altijd geen oplossing. Komt er nu een uitgebreid onderzoek in de Kamer of niet? De kamervoorzitter vroeg de partij vanmiddag haar / hun algemene opinie. Maar de eensgezindheid binnen de partij bleek zeer ver te zoeken. De conservatieve groep rond de ex-premier is gewonnen voor een alternatieve straf, de jongere garde niet.

Door zijn tactische sluwheid en zijn charismatische persoonlijkheid werd zijn macht over leger, politiek en bevolking groter en groter. De dictator ontzegde de partij gisteravond haar / hun verworven vetorecht. Maar de bevolking blijkt daar nu tegen in opstand te komen, want vandaag krijgen we berichten binnen van stevige rellen in de hoofdstad en in de grote kuststeden.

Canada zit erg verveeld met de smeufge pedofiliezaak. Intussen zitten er dertig bezwarende getuigenissen in het gerechtelijke dossier. De premier ontnam de vereniging gisteravond haar / hun laatste beroepsmogelijkheid. Maar de maatregel is op stevig protest onthaald, aangezien op die manier een aloud rechtsprincipe overboord wordt gegooid.

Fijngevoelig en vrouwvriendelijk kon je het in ieder geval moeilijk noemen: blote borsten en platvloerse grappen swingden de pan uit. Een journalist verweet de vereniging gisteravond haar / hun wansmakelijke reclamestunt. Maar de voorzitter verdedigde zich door te verwijzen naar de reclameactie van de concurrentie een jaar geleden, waarop niemand schijnbaar kritiek had.

Eergisteren was er de spreekwoordelijke druppel. Toen sneuvelde er verschillende grote ruiten uit een regeringsgebouw. Een toppoliticus ontzegde de organisatie vanmiddag haar / hun wettelijke stakingsrecht. Maar de timing van die beslissing, net op de vooravond van de verkiezingen, is op zijn zachtst gezegd ongelukkig te noemen.

De stad heeft een clean, braaf en groen imago en de burgemeester wil dat zo houden. Al jarenlang woedt er dan ook een heftige vete. De burgemeester verbood de organisatie indertijd haar / hun jaarlijkse optochten. Maar de rechter gaf hem later ongelijk: de optochten werden weer toegelaten, met de huidige verhitte acties en provocaties tot gevolg.

Het ging al jaren minder goed en de inkomsten daalden zienderogen, tot plots de redding kwam en wel uit zeer onverwachte hoek. Een boekhouder bezorgde de firma indertijd haar / hun kolossale beurswinst. Maar de identiteit van de “weldoener” zou nog jarenlang geheim blijven. Achteraf bleek het om McDough te gaan. Het salaris van de man was intussen vertienvoudigd.

Windschade en de bijbehorende vertragingen maken het concert nog steeds twijfelachtig, zodat extra mankracht een absolute must is. De organisator vroeg de firma vanmiddag haar / hun bereidwillige medewerking. Maar de vakantie zorgt ervoor dat de kleine minderheid die nog aan het werk is, heel hard zal moeten doorwerken om alles op tijd klaar te krijgen.

Een zeer lieve man, drie schatten van kinderen en een aantrekkelijke job: de ex-topzwemster heeft het erg naar haar zin in Zwolle. De Marokkaanse bezorgde de club indertijd zijn / hun eerste kampioenstitel. Maar de clubleiding is dat, in tegenstelling tot de supporters, blijkbaar al lang vergeten. Op de eeuwfeestviering werd de vroegere vedette niet uitgenodigd.

De soap duurde een aantal weken, totdat drie dagen geleden in een Brusselse gravin dan toch een potentiële geldschietter werd gevonden. De gravin vergoedt de club binnenkort zijn / hun hoge gerechtskosten. Maar de vraag is of dat zal volstaan om in eerste klasse te blijven. Pas na de winterstop beslissen de hoge bondsbonzen over de licentie van de club.

Ook in de huiskamers begint de spanning stilaan te stijgen. De reis is nog niet binnen, want er volgt nog één ultieme hindernis. De presentatrice geeft het echtpaar binnenkort zijn / hun laatste opdracht. Maar de kans dat die opdracht goed afloopt, blijkt de afgelopen weken sterk gedaald te zijn. Toch nog even alles geven en hopelijk is de wereldreis dan een feit.



Waarzeggers en sterrenwichelaars zijn allemaal bedriegers die alleen maar uit zijn op geld en veel leed veroorzaken, luidt het nu. Een waarzegster ontnam het echtpaar indertijd zijn / hun innige kinderwens. Maar de medische vooruitgang staat voor niets: door middel van in-vitrofertilisatie werd Els toch zwanger en gisteren werd Jonas geboren.

Terwijl heel wat kinderen het schip bezochten, merkten twee begeleidende leerkrachten een verdrinken vrouw op in het water. De politiecommissaris verweet de bemannig gisteravond haar / hun grote laksheid. Maar de zaak moet eerst verder worden onderzocht, voordat er tot eventuele arrestaties kan worden overgegaan. Het is al het derde slachtoffer in één maand tijd.

De afreis naar het verre, warme Zuid-Amerika is voorlopig gepland voor volgende woensdag, tenminste als het weer het dan toelaat. De kapitein gunde de bemannig gisteravond haar / hun laatste uitspatting. Maar de drank aan boord bleek gelimiteerd te zijn, zodat de nachtwinkels erg goede zaken deden. Het feestje duurde tot vroeg deze morgen.

Zijn leven werd totaal verwoest en die gedachte is moeilijk te verdragen. Veel wil Tom niet meer kwijt over de mensonterende pesterijen. De jongen vergaf de klied indertijd haar / hun publieke lastercampagne. Maar de psychische gevolgen die hij eraan overhoudt, zijn niet te onderschatten: nachtmerries, slaapstoornissen en paniekaanvallen maken zijn leven ondraaglijk.

De overheid is tot nog toe te laks geweest. Criminaliteit viert hoogtij in de wijk en het gaat zeker niet om een alleenstaand geval. Een wijkopzichter gaf de klied indertijd haar / hun eerste waarschuwing. Maar de baldadigheden bleven voortduren, zodat gisteren, na een nieuwe rel, de maat vol was voor enkele buurtbewoners. Zij trokken nog maar eens naar de politie.

Zeer blijde gezichten vandaag, want de kogel is eindelijk door de kerk: er komen nu toch meer financiële middelen voor de dienst. De secretaresse vergoedt het trio binnenkort zijn / hun dagelijkse reiskosten. Maar de andere eisen, waaronder een loonsverhoging van minstens twee procent, een nieuwe medewerker en beter betaalde overuren, worden voorlopig niet ingewilligd.

Deze trieste zelfmoord is nog maar eens het zoveelste bewijs dat pesterijen in het onderwijs niet streng genoeg kunnen worden aangepakt. De leraars vergaf het trio indertijd zijn / hun zware uitlatingen. Maar de valse pesterijen, de gemene insinuaties en intimidaties staken minder dan een maand later weer de kop op. Dat bleek voor haar de genadeslag.

Vanavond doken zelfs geluiden op van een mogelijke rechtszaak tegen de organisator van de wedstrijd, de gemeentelijke basisschool. De directrice overhandigde de tweeling vanmiddag zijn / hun eerste hoofdprijs. Maar de uitslag werd meteen aangevochten door enkele ouders, die de jury beschuldigden van favoritisme. Velen namen zelfs de term “omkoperij” in de mond.

De moeder, in een niet zo ver verleden zelf nog ooit Vlaams turnkampioene, vindt lichaamsbeweging erg belangrijk, voor jong en oud. Een sportlerares geeft de tweeling binnenkort zijn / hun eerste zwemles. Maar de vader is daar niet zo meteen voor te vinden: hij is van mening dat kinderen zelf het initiatief moeten nemen en dat dwang of druk op lange termijn niet helpt.

Financiële schulden, een onder water gelopen huis en de dood van de oma: na alle ellende dit jaar zal die adempauze zeker niet ongelegen komen. De buurvrouw gunt het gezin binnenkort zijn / hun korte herfstvakantie. Maar de honden te eten geven, durft ze niet. De drie rottweilers van het gezin boezemen haar al jaren angst in en gaan voor een weekje naar een neef.

Rond de middag bleef alles nog binnen de grenzen van het welvoeglijke, maar elf uur later stroomden veel klachten binnen en was het prijs. Een agente verbood het gezin gisteravond zijn / hun lawaaierige tuifneestje. Maar de gasten waren nog niet zo snel te overtuigen. Uiteindelijk moest er een extra politiewagen aanrukken voordat iedereen teleurgesteld afdroop.

### Appendix 3. Eye-tracking experiment: reading times in ms

		Region 1		Region 2		Region 3		Region 4	
		Sing	Plur	Sing	Plur	Sing	Plur	Sing	Plur
FFD	Low	197	196	226	226	237	240	224	228
	High	184	181	246	226	232	242	230	227
GD	Low	1051	1069	234	228	664	665	527	565
	High	1029	1031	251	229	536	579	580	610
TFD	Low	1257	1263	257	248	890	947	601	634
	High	1175	1166	281	245	738	759	633	644
CRRT	Low	1125	1143	261	239	968	1017	599	620
	High	1058	1077	294	257	819	829	640	648

Significant cells ( $p < 0.05$ ) have been italicized.

### Notes

- \* This article is a revision of Joosten (2003: Ch. 7). The authors wish to thank Joosten's dissertation director Willy Smedts, Ilse Depraetere, Géry d'Ydewalle, Jan Hulstijn, Gert Storms, Piek Vossen and three anonymous reviewers for their useful suggestions and comments, and John Osborne for his stylistic corrections. The research was supported by a research grant from the Fund for Scientific Research — Flanders. Correspondence address: Frank Joosten, Boekstraat 61, 3500 Hasselt, Belgium. E-mail: fjoosten@yahoo.com.

1. See also Geerts (1977: 169) and Forsmark (n.d.: 2).

2. See also Wierzbicka (1991: 374).
3. Vossen (1995: 35) defines *conceptual individuation* as “the entities we have in our minds as a result of interpreting terms in normal circumstances”. See also Section 3.
4. English differs from languages such as French, German, and Dutch in that English collective nouns can pattern with both a singular and a plural verb (*the committee is/are*). See Section 2.
5. The short lines connecting the members represent the external (spatio-temporal, social, cooperative, functional) relations between them.
6. See Joosten (2003: Ch. 2).
7. Rare exceptions are Poutsma (1914), Jespersen (1924), and Persson (1989): they advocate a semantic definition that includes inanimate collective nouns.
8. Three random definitions in French, German and Dutch studies: “un nom singulier qui représente un regroupement d’éléments d’une même catégorie” (Borillo 1997: 106), “Lexeme [...] [die] referieren auf einzelne Individuen, die also “versammelt”, d.h. als diskreter Gegenstand konzeptualisiert werden” (Mihatsch 2000: 39), “benamingen van een aantal gelijksoortige wezens of dingen die tezamen een eenheid vormen” (Haeseryn et al. 1997: 140) [names for a multiplicity of similar animate beings or things that together form a unity].
9. See Section 4 for a more nuanced view.
10. Flaux (1999) is to our knowledge the only study based on form-meaning relationships that seriously questions a broad semantic definition of collective nouns as well.
11. The term *property* is used in a fairly broad sense here, for anything that is or may be predicated of a noun, no matter whether that predication has an accidental (e.g., *drunk(en), to say*) or a more permanent character (e.g., *big, to live in England*) and whether it is expressed by an adjective or a verb.
12. A syntactic consequence is that these collective nouns, in a member identifying construction, are regularly followed by plural verbs, a phenomenon that is often named *attraction* or *proximal concord* (see e.g., Levin 2001). One Dutch corpus example: “[...] *dat de groep sympathisanten de openbare zitting bijwoonden.*” (Condiv) ‘that the group of sympathizers attended (past third-person plural) the public session’.
13. In her terminology both are called *Kollektiva*, following the “continental” tradition.
14. This method of combining collective nouns with a fixed set of properties is comparable to using a set of different drills in order to measure the thickness of a wall. For the drill metaphor see Joosten (2001).
15. The *result* column indicates whether the combination of the property in question with a collective noun is always semantically acceptable (+) or not (+ or –).
16. Two persons can buy a house together, or move a table together, but it is impossible for two persons to perform one “act” of (or share one state of) being drunk. Thus, *dronken* and *blond* are properties typically associated with one person at a time. Similar properties are *slapen* (‘to sleep’) or *lachen* (‘to laugh’).
17. The abbreviations *c* and *m* indicate collection and member level interpretations, while a minus sign and a question mark mean that the combination of property and collective noun is semantically unacceptable (–) or questionable (?). If both collection and member level interpretations are plausible, *c/m* is used. That also holds true for *c(/m)*, in which case collection level interpretations are more likely.
18. See e.g., Barker (1992: 86–87).
19. A possible exception is *familie* (‘family’). *Een eeuwenoude familie* (‘an age-old family’) seems semantically acceptable, but as property distribution with *jong* (‘young’), *blond* (‘blond’), and *dronken* (‘drunk(en)’) follows the pattern of type 2 rather than type 1, classification as a type 2 collective noun seems justified.

20. As far as *groot* ('big') is concerned, a fixed number of members automatically implies member level interpretations, but not vice versa. That means that for type 3 collective nouns without a fixed number (e.g., *publiek* 'public') both member and collection level interpretations are possible.
21. Levin (2001) mentions other corpus studies on English collective nouns, but as is evident from his comments, those studies are not concerned with individual differences either. Most of them consider regional or stylistic factors (e.g., British vs. American English, written vs. spoken language).
22. All data were electronically annotated in *Abundantia Verborum*, a computer tool designed by Speelman for carrying out corpus-based linguistic case studies (see Speelman 1997).
23. It will become clear at the end of this section why the division is based on mean B rather than on mean A.
24. In terms of percentage, the borders between type 1 and 2 (e.g., *firma* (32.4%) vs. *team* (34.6%)) and between type 2 and 3 (e.g., *familie* (58.2%) vs. *gezin* (63.6%)) are far from clear-cut, but that confirms rather than contradicts our point that member level accessibility is a gradual notion.
25. For brevity's sake, *attributive* is left out of consideration here. Dutch resembles English in that determiners and quantifiers preceding Dutch collective nouns are always singular (e.g., *dat* / \**die* / *een* / \**tien* *comité* ('that' / \*those' / a' / \*ten committee')). In all our corpus data not a single counterexample was found.
26. Counterexamples are very rare. For six collective nouns (*duo* 'duo, pair', *gezin* 'family, household', *team* 'team', *maatschappij* 'company', *comité* 'committee', *regering* 'government') relative pronouns are proportionately more plural than verbs, but for five of them that difference is less than 1% and it seems that a major factor is the considerably smaller number of relative pronouns (which means that the chance of a plural form is also smaller.) The only irrefutable counterexample is *bemannings* ('crew'), for which possessive pronouns are proportionately more plural than personal pronouns (respectively 73.3% and 70.7% plural). We lack a plausible explanation for this irregularity.
27. Consequently, the examples in which a shift takes place from singular to plural are legion (e.g., *De bemanning is tevreden. In het gebied waar zij opereerden is het embargo in elk geval gewaarborgd* (INL). [The crew is ... they ...]. The reverse pattern, a shift from plural to singular, is very rare (see e.g., Levin 2001: 110–121).
28. It appears that in the Middle Dutch and Early New Dutch period collective nouns occurred with plural verbs far more often (Geerts 1977: 166).
29. If collective nouns occur in a member identifying construction (e.g., *een team van deskundigen* 'a team of experts', see Section 2) — a condition that was not taken into consideration in this analysis (see Appendix 1) — then plural verbs are less exceptional, although they are generally considered to be less "correct" than singular ones by most Dutch language users (see Haeseryn et al. 1997: 1149).
30. Compare the following figures:
  - high vs. medium plural concord; possessive:  $\chi^2 = 105.81$ ; personal:  $\chi^2 = 13.21$ .
  - medium vs. low plural concord; possessive:  $\chi^2 = 11.66$ ; personal:  $\chi^2 = 64.76$ .
31. That explains why we prefer mean B to distinguish "high", "medium," and "low" plural concord.
32. Furthermore, our data suggest that the claim put forward by Bock et al. (1999) should be slightly modified: not all pronouns are governed by the speaker's meaning. Dutch relative pronouns resemble verbs in that their number features may be retrieved under control from the utterance's form. Possessive and personal pronouns, on the other

- hand, exhibit the same behavior as the English reflexive and tag pronouns in Bock et al. (1999): they are more prone to conceptual factors.
33. As we do not discriminate between *semantic* and *conceptual*, Corbett's (1979) "semantic agreement" is, for simplicity's sake, renamed *conceptual*. An asterisk (\*) indicates ungrammatical constructions (e.g., *\*het comité hebben* 'the committee have' (plural verb)).
  34. In contrast to English, Dutch has a gender system that discriminates between masculine, feminine and neuter. In this example *ze* ('she') refers to feminine *bemannig* ('crew').
  35. See also Humphreys and Bock (2005) on the complicating effects of contextual variation and the need for minimally contrastive contexts.
  36. The opposite, however, does not hold true: profiling of the member level does not exclude singular verbs or pronouns. There are quite a few contexts that block plural forms, but singular ones are almost always possible. See Levin (2001: 151).
  37. This notion of *distance* is different from Corbett's (1979). Corbett discusses "syntactic distance", i.e., distance in terms of syntactic relations between words. Here, *distance* should be interpreted fairly literally as 'real, actual, linear distance' (e.g., the number of words between collective noun and verb or pronoun). As both (12a) and (12b) have a personal pronoun, syntactic distance is similar. See also Nixon (1972), Forsmark (n.d.), and Levin (2001: 92–99).
  38. Except for *publiek* ('public').
  39. In our hypothesis, singular pronouns are not problematic for type 3 collective nouns in the same way that plural pronouns are problematic for type 1 collective nouns. The reason for this is that singular pronouns can be syntactically determined, whereas plural ones cannot (as far as singular collective nouns are concerned). In other words, if a type 1 collective noun (e.g., *bond* 'union') is combined with a plural pronoun, then this combination is not only conceptually problematic, but also syntactically deviant. However, if a type 3 collective noun (e.g., *echtpaar* 'married couple') is combined with a singular pronoun, then the combination may be hard to conceptualize, but from a syntactic point of view a type 3 collective noun can perfectly combine with a singular pronoun (see e.g., [11]). Our hypothesis is that singular pronouns will therefore be relatively less problematic.
  40. Translation: "It had been a kind of cat and mouse game for years, a game in which money, power and corruption seemed to prevail. Until today's verdict. A top lawyer was responsible for the union's first defeat this afternoon. But the case isn't closed yet, since the union intends to appeal against the decision. Legal experts say that in that case the odds are fifty-fifty."
  41. A Latin square design, in which texts have two variants, means that the first participant receives text 1 in the first variant, text 2 in the second, text 3 in the first, and so on. The second participant will read text 1 in the second variant, text 2 in the first, text 3 in the second, and so on. In such a design every participant reads a text with a singular and a text with a plural possessive pronoun for all 12 collective nouns selected. All conditions (low accessibility / singular; low accessibility / plural; high accessibility / singular; high accessibility / plural) are equally represented (6 text fragments).
  42. As for  $F_2$ , the difference is not significant ( $F_2(1,11) = 2.23$ ,  $p > 0.10$ ). An explanation will be given at the end of the section.
  43. Here as well,  $F_2$  is not significant ( $F_2(1,11) = 2.76$ ,  $p > 0.10$ ). See the explanation at the end of the section.
  44. Differences in word length also affect the initial skipping of region 2 (i.e., GD = 0 ms). *Hun* (3 characters) is (in the first entry) skipped in 75.1% of the trials (341 times out of 454), *zijn/haar* (4 characters) in only 62.3% of the trials (282 times out of 453).

45. In Levin (2001) it is *couple* that has the highest scores for plural concord. See also Poutsma (1914: 284): “[The plural construction] is the usual construction when the collective noun denotes a small body of persons.”
46. See Levin (2001: 144) on the predominance of singular verbs for *army*, *audience*, *faculty*, and *population*: “Collectives comprising very large numbers of individuals approach mass nouns in that the constituent members are less likely to be highlighted.” Compare also Poutsma (1914: 283).
47. Comments on the “origin process” of collections can be found in e.g., Jespersen (1924: 195–196), Cruse (1986: 175–177), and Biber et al. (1999: 247–250).
48. *Group* is used in its (social) psychological sense here, though — as appears from Wilder (1981: 215–216) — the term is far from unequivocal (much like *collective noun* in linguistics). Wilder (1981: 216) defines a *group* as “a collection of persons who share (or are thought to share) some set of characteristics and who may (but not necessarily will) interact with one another.”
49. As already mentioned, Depraetere (2003) does not deal with pronominal concord. Levin (2001) analyzes verbs, relative pronouns and personal pronouns, but neglects possessive pronouns.
50. In Forsmark’s (n.d.) small-scale questionnaire survey on Swedish collective nouns *besättning* (‘crew’) exhibits higher plural concord scores than *familj* (‘family’), which, in its turn, is more plural than *församling* (‘assembly’).
51. Related to this discussion is the interesting issue of the (morpho)syntactic differences between Dutch and English type 3 collective nouns. As *crew* can directly combine with cardinal numerals (e.g., *twenty crew*) is there a conceptual difference with Dutch *bemanning* (e.g., *\*tuintig bemanning*)? Why isn’t there an English collective noun counterpart for Dutch *tweeling* (‘twins’)? And if, at first sight, (morpho)syntactic differences between English and Dutch are typical for type 3 collective nouns (and less so for type 1 and 2), is that not because (morpho)syntactic variation becomes more likely in cases of obvious multiplicity (Vossen 1995: 202–204)?
52. However, despite this difference, English concord is more syntactically determined than generally presumed. The data discussed in Depraetere (2003) suggest that English is developing in the direction of Dutch. A few exceptions such as *couple* notwithstanding, English collective nouns are usually followed by singular verbs.
53. Landman (1989: 742) briefly discusses differences in “intentionality” between *committee*, *group* and *deck*.

## References

- Baayen, Harald R.; Piepenbrock, Richard; and van Rijn, Hedderik (1993). *The CELEX Lexical Database*. Philadelphia: Linguistic Data Consortium.
- Bache, Carl and Davidsen-Nielsen, Niels (1997). *Mastering English: An Advanced Grammar for Non-Native and Native Speakers*. Berlin and New York: Mouton de Gruyter.
- Barker, Chris (1992). Group terms in English: representing groups as atoms. *Journal of Semantics* 9, 69–93.
- Benninger, Céline (2001). Une meute de loups / une brassée de questions: collection, quantification et métaphore. *Langue française* 129, 21–34.
- Biber, Douglas; Johansson, Stig; Leech, Geoffrey; Conrad, Susan; and Finegan, Edward (1999). *Longman Grammar of Spoken and Written English*. Harlow: Longman.

- Bock, Kathryn; Nicol, Janet; and Cutting, J. Cooper (1999). The ties that bind: creating number agreement in speech. *Journal of Memory and Language* 40, 330–346.
- Borillo, Andrée (1997). Statut et mode d'interprétation des noms collectifs. In *Co-texte et calcul du sens. Actes de la table ronde tenue à Caen les 2 et 3 février 1996*, Claude Guimier (ed.), 105–121. Caen: Presses Universitaires de Caen.
- Campbell, Donald T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science* 3, 14–25.
- Chelaru-Ioniță, Oana and Bantaș, Andrei (1981). English versus Romanian collectives. *Papers and Studies in Contrastive Linguistics* 13, 223–248.
- Corbett, Greville G. (1979). The agreement hierarchy. *Journal of Linguistics* 15, 203–224.
- Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Depraetere, Ilse (2003). On verbal concord with collective nouns in British English. *English Language and Linguistics* 7, 85–127.
- Dölling, Johannes (1991). Group expressions and the semantics of plurals and collective nouns. In *Syntax und Semantik der Substantivgruppe*, Ilse Zimmermann (ed.), 147–183. Berlin: Akademie Verlag.
- (1995). Ontological domains, semantic sorts and systematic ambiguity. *International Journal of Human-Computer Studies* 43, 785–807.
- Flaux, Nelly (1999). À propos des noms collectifs. *Revue de Linguistique Romane* 63, 471–502.
- Forsmark, Beata (n.d.). Kollektiver, animatet och numeruskongruens. Unpublished manuscript, University of Stockholm.
- Geerts, Guido (1977). Het collectivum als haar-syndroom. *Tijdschrift voor Nederlandse Taal- en Letterkunde* 93, 153–200.
- and den Boon, Ton (1999). *Van Dale Groot Woordenboek der Nederlandse Taal*, 13th revised ed. Utrecht and Antwerp: Van Dale Lexicografie.
- Gil, David (1996). Maltese 'collective nouns': a typological perspective. *Rivista di Linguistica* 8, 53–87.
- Grondelaers, Stefan; Deygers, Katrien; Van Aken, Hilde; Van den Heede, Vicky; and Speelman, Dirk (2000). Het CONDIV-corpus geschreven Nederlands. *Nederlandse Taalkunde* 5, 356–363.
- Haeseryn, W.; Romijn, K.; Geerts, G.; de Rooij, J.; and van den Toorn, M. C. (1997). *Algemene Nederlandse Spraakkunst*, 2nd revised ed. Groningen and Deurne: Martinus Nijhoff and Wolters Plantyn.
- Hamilton, David L. and Sherman, Steven J. (1996). Perceiving persons and groups. *Psychological Review* 103, 336–355.
- Humphreys, Karin R. and Bock, Kathryn (2005). Notional number agreement. *Psychonomic Bulletin & Review* 12(4), 689–695.
- Jespersen, Otto (1924). *The Philosophy of Grammar*. London: George Allen & Unwin.
- Joosten, Frank (2001). De 'boorbaarheid' van Nederlandse collectiva. *Over Taal* 40, 77–79.
- (2003). Collectiva en aggregaatsnamen in het Nederlands: begripsbepaling en typologie. Unpublished doctoral dissertation, University of Louvain.
- Juul, Arne (1975). *On Concord of Number in Modern English*. Copenhagen: Nova.
- Krifka, Manfred (1991). How to get rid of groups, using DRT: a case for discourse-oriented semantics. *Texas Linguistic Forum* 32, 71–110.
- Kruyt, J. G. and Dutilh, M. W. F. (1997). A 38 million words Dutch text corpus and its users. *Lexikos* 7, 229–244.
- Kuhn, Wilfried (1982). Kollektiva und die Technik KOLLEKTION am Beispiel des Deutschen. In *Apprehension. Das sprachliche Erfassen von Gegenständen. Teil I: Bereich und Ordnung der Phänomene*, Hansjakob Seiler and Christian Lehmann (eds.), 84–97. Tübingen: Gunter Narr.

- Landman, Fred (1989). Groups I, II. *Linguistics and Philosophy* 12, 525–605, 723–744.
- Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar*, vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lecolle, Michelle (1997). Étude des noms collectifs en français. Unpublished master's thesis, University of Toulouse.
- Leisi, Ernst (1975). *Der Wortinhalt: Seine Struktur im Deutschen und Englischen*. Heidelberg: Quelle & Meyer.
- Levin, Magnus (1998). Manchester United are my team: on concord with collective nouns. *Moderna Språk* 92, 14–18.
- (2001). Agreement with collective nouns in English. Lund Studies in English 103. Unpublished doctoral dissertation, University of Lund.
- Lønning, Jan Tore (1997). Plurals and collectivity. In *Handbook of Logic and Language*, Johan F. A. K. van Benthem and Alice G. B. ter Meulen (eds.), 1009–1053. Amsterdam: Elsevier.
- Michaux, Christine (1992). The collectives in French: a linguistic investigation. *Linguisticae Investigationes* 16, 99–124.
- Mihatsch, Wiltrud (2000). Wieso ist ein Kollektivum ein Kollektivum? Zentrum und Peripherie einer Kategorie am Beispiel des Spanischen. *Philologie im Netz* 13, 39–72.
- Nixon, Graham (1972). Corporate-concord phenomena in English. *Studia Neophilologica* 44, 120–126.
- Perlmutter, David M. (1972). A note on syntactic and semantic number in English. *Linguistic Inquiry* 3, 243–246.
- Persson, Gunnar (1989). On the semantics of collective nouns in English. In *Instead of Flowers: Papers in Honour of Mats Rydén on the Occasion of his Sixtieth Birthday, August 27, 1989*, Bengt Odenstedt and Gunnar Persson (eds.), 179–188. Stockholm: Almqvist and Wiksell.
- Pollard, Carl and Sag, Ivan A. (1988). An information-based theory of agreement. *Papers from the 24th Annual Regional Meeting of the Chicago Linguistic Society (Parasession on Agreement in Grammatical Theory)*, 236–257.
- Poutsma, H. (1914). *A Grammar of Late Modern English*, part 2: *The Parts of Speech*. Groningen: P. Noordhoff.
- Quirk, Randolph; Greenbaum, Sidney; Leech, Geoffrey; and Svartvik, Jan (1985). *A Comprehensive Grammar of the English Language*. London and New York: Longman.
- Rayner, Keith (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124, 372–422.
- Speelman, Dirk (1997). *Abundantia verborum: a computer tool for carrying out corpus-based linguistic case studies*. Unpublished doctoral dissertation, University of Louvain.
- Vossen, Piek (1995). *Grammatical and Conceptual Individuation in the Lexicon*. Amsterdam: IFOTT.
- Wierzbicka, Anna (1991). Semantic rules know no exceptions. *Studies in Language* 15, 371–398.
- Wilder, David A. (1981). Perceiving persons as a group: categorization and intergroup relations. In *Cognitive Processes in Stereotyping and Intergroup Behavior*, David L. Hamilton (ed.), 213–257. Hillsdale, NJ: Erlbaum.