Number Agreement in French Binomials

Aixiu An • Anne Abeillé

Abstract Number agreement in determiner plus bare-noun coordination (D N & N) varies across languages. This paper presents an empirical study of French in which the construction has been claimed to only be felicitous with plural nouns (e.g., Heycock & Zamparelli 2005, Le Bruyn & de Swart 2014). We use large-corpus data and ran two acceptability rating experiments, showing that singular split binomials allow for two agreement strategies for the shared determiner: Closest Conjunct Agreement (*votre nom et prénom* 'your.SG name and first name' or Resolution (*vos nom et prénom* 'your.PL name and first name'). We also show that speakers' preferences are sensitive to noun animacy and the syntactic function of the binomial. We assume that animate binomials give rise to a potential ambiguity between a joint and a split reading. We propose an HPSG analysis that assigns the same syntactic structure to split and joint binomials, while allowing for different agreement strategies.

Keywords number agreement · binominals · split reading · animacy · French · large corpora · experimental syntax

A. An, LLF/University of Paris 7-Diderot, aixiu0806@gmail.com
A. Abeillé, LLF/University of Paris 7-Diderot, abeille@linguist.univ-paris-diderot.fr

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 31–60. Paris: CSSP.
http://www.cssp.cnrs.fr/eiss12/

1 Introduction

In French, determiners and nouns agree in number.[1] Singular nouns require singular determiners and plural nouns require plural determiners (1a). And two coordinated singular NPs (with *et* 'and') trigger plural agreement (1b).

[1]French determiners also agree in gender. We ignore gender agreement here, but see An & Abeillé 2017 and Abeillé et al. 2018.

(1) a. un chien /des chiens
a.SG dog.SG /a.PL dog.PL
b. [un chien et un chat] intelligents
a.SG dog.SG and a.SG cat.SG clever.PL

Crosslinguistically, binomials display specific agreement properties. When two bare nouns are conjoined with a shared determiner (D N1 & N2), they can be coreferent (joint interpretation) (2a) or not (split interpretation) (2b) (e.g., King & Dalrymple 2004, Heycock & Zamparelli 2005, Le Bruyn & de Swart 2014).

(2) a. A friend and colleague is coming today.
b. My father and mother are separated.

In most languages, the determiner (D) agrees in number with both conjuncts in the joint interpretation, since they share the same index. However, languages have different strategies for the split interpretation: in English, as in Finnish and Hindi, the D is singular with singular nouns (3a), and nouns with different numbers cannot be coordinated with a shared D (3b), (3c) (King & Dalrymple 2004).

(3) a. This boy and girl are eating a pizza.
b. *these boys and girl
c. *this boy and girls

Other languages also use a singular D but allow for conjuncts with different numbers. In that case, Spanish (4a) and Portuguese (4b), for example, allow Closest Conjunct Agreement (CCA). The shared D agrees with the first noun.[2]

(4) a. Los ataques de la aviación y
the attacks of the.FSG aeroplanes.FSG and
helicópteros rusos fueron constantes.
helicopters.MSG Russian.MPL be.PST.3PL constant.PL

[2]In the glosses, FSG represents singular feminin gender, MPL represents plural maculin gender, and NPL represents plural neutral gender.

'The attacks of the Russian aeroplanes and helicopters were constant.' (Demonte & Perez-Jimenez 2012)

b. O presidente e amigo/amigos comeram
the.MSG president.MSG and friend.MSG/PL ate.PL
juntos.
together
'The president and friend/friends ate together.' (Villavicencio et al. 2005)

CCA is also attested in Serbo-Croatian (Wechsler & Zlatić 2003, Dalrymple & Hristov 2010). In (5), the prenominal Adj agrees with the first conjunct and the participle agrees with the second conjunct.

(5) Nova kaŭca i kola su kŏstal-a puno.
new.FSG house.FSG and car.NPL AUX.PL cost.NPL a.lot
'The new house and car cost a lot.' (Dalrymple & Hristov 2010)

Le Bruyn & de Swart (2014) assume that English-like and Spanish-like languages work similarly, since a plural D is not possible (4a), (6a), except for some rare cases with proper names ((6b) from Demonte & Perez-Jimenez 2012). They note that English may allow a mismatch of number in an appropriate context, as in (6c).

(6) a. *These boy and girl are eating a pizza.
b. ... los temibles Capirucho y Capirote ...
... the.MPL fearsome.PL Capirucho and Capirote ...
'... the fearsome Capirucho and Capirote '
c. These children and mother were living on charity of good people.

Assuming a hierarchical structure for coordination (Kayne 1994, Borsley 2005), examples (4a) and (5) can also be considered as cases of highest conjunct agreement (e.g., Marušič et al. 2007, Murphy & Puškar 2018): agreement with the first noun, which is hierarchically highest, or closer to D in terms of number of intervening nodes.

According to Le Bruyn & de Swart (2014), Russian is one of the "typologically rare languages" that allows for a plural D with a coordination of singular nouns (King & Dalrymple 2004) (7a), and French is supposed to forbid both a singular D and a plural D (7b), except in some "frozen" cases such as *vos père et mère* ('your.PL father and mother').

(7) a. èti mužčina i ženščina
this.PL man.MSG and woman.FSG
'these man and woman' (King & Dalrymple 2004:95)
b. *Ce/*Ces marin et soldat sont souvent ensemble.
this.SG/PL sailor and soldier are.PL often together
(Heycock & Zamparelli 2005:3)

Contrary to Heycock & Zamparelli (2005) and Le Bruyn & de Swart (2014), we have found that both a singular and a plural D are permitted in French with two non-coreferent singular nouns (An & Abeillé 2017), as in (8).

(8) Indiquez votre/vos nom et prénom.
indicate your.SG/.PL last name.MSG and first name.MSG
'Indicate your first and last name.'

However, this is not felicitous with all binomials. For instance, a plural D is better in (9a), and a singular D in (9b), respectively.

(9) a. Les/?Le frère et sœur d'Emmanuel
the.PL/?.MSG brother.MSG and sister.FSG of Emmanuel
Macron n'ont pas souhaité être exposés
Macron NEG-AUX NEG want.PST-PT be exposed.MPL
durant cette campagne.
during this campaign
'The brother and sister of Emmanuel Macron didn't want to be exposed during this campaign.' (FrWAC, `www.gala.fr`)

b. La/?Les simplicité et beauté sont pour moi
the.FSG/.PL simplicity.FSG and beauty.FSG are for me
étroitement associées.
closely associated.PST-PT.FPL
'The simplicity and beauty are closely associated for me.'
(FrWAC, www.auto-evasion.com)

We present an empirical study based on large-scale corpus data and two controlled experiments, investigating D number agreement with French singular binomials, as well as its interactions with other factors, such as noun animacy and syntactic function. The article is organized as followed: in section 2, we discuss the backgrounds of our study, showing that the ambiguity between joint and split readings may come from frequency and context effects. With corpus data in section 3, we show that non-human nouns are more frequent with a split reading while human nouns favor a joint interpretation. With split singular binomials, a plural D is preferred for human nouns and a singular D for non-human nouns. In section 4, we report two experimental studies, testing the acceptability of singular/plural D with human and non-human nouns, in two different positions: subject and object. In section 5, we compare our corpus data and our experimental results and propose a formal analysis in HPSG.

2 Research Questions

2.1 Factors favoring Closest Conjunct Agreement

According to Corbett (1991), the factors favoring Closest Conjunct agreement (singular) or agreement resolution (plural) crosslinguistically, involve the agreement controller (the element which governs agreement) and the agreement target (the element which marks agreement). In a corpus study of Medieval Spanish, German, Russian and Serbo-Croatian, in which the predicate can be either singular or plural with two singular conjoined nouns, Corbett counted agreement examples in two possible orders, subject-predicate and predicate-subject, and showed that resolution (plural) was found

more frequently with animate nouns and in subject-predicate order.

Corbett also claims that the choice between resolution and non-resolution obeys an agreement hierarchy (10) (Corbett 1991). As we move rightwards along the hierarchy of targets, the likelihood of singular agreement decreases and that of resolution (plural) increases.

(10) attributive > predicative > relative pronoun > personal pronoun
⟵ non-resolution resolution ⟶

He does not mention determiner agreement. Since Determiner-Noun order is fixed, we predict that the only factor that interacts with agreement is animacy. Amongst the interactions between number agreement and noun animacy, we predict that a plural D will be favored by animate nouns over non animates.

2.2 Binomials and Semantic Ambiguity

Binominal constructions are ambiguous between a joint (coreferent) and a split (non coreferent) reading. We observe that human nouns are often role and function names, and thus more compatible with a joint reading than non-human nouns, referring to objects, properties or situations. The plural animate binominal is ambiguous between a joint reading and a split reading in (11b), while the inanimate binominal does not involve this ambiguity (11b).

(11) a. J' ai vu des collègues et amis
I see.PST a.PL colleagues and friends
'I saw some colleagues and friends.' (ambiguous)
b. J' ai vu des cahiers et crayons
I see.PST a.PL notebook.MPL and pencil.MPL
'I saw some notebooks ans pencils.' (split)

With singular nouns, both readings are possible with a singular D, while a plural D forces a split reading. To avoid this ambiguity, we suppose that singular animate Nouns favor a division of labor between plural D (split reading) (12a) and singular D (joint reading)

(12b).

(12) a. des collègue et ami
'a.PL colleague and friend.'
b. un collègue et ami
'a.SG colleague and friend.'

We will examine whether a biased context favoring the split reading will affect the acceptability of a plural D and section 4.

It should be noted that animacy is rather a continuum than a binary distinction between animate-inanimate (Comrie 1989, Croft 2002, Zaenen et al. 2004, Haspelmath 2013). In this study, we distinguish humans and non-humans for the reason that coreference is more likely to occur with human nouns. Animals are considered as non-humans for the same reason (*des chiens et chats* ('dogs and cats') cannot have a joint reading).

We begin with corpus data comparing the distribution of human and non-human binomials, as well as D agreement. Then, we report two experimental studies controlling syntactic and semantic factors, in order to analyze the different factors affecting the acceptability of French binomials.

3 Corpus study

3.1 Methodology

We use the web-based corpus FrWac (1.6 billion words, Baroni et al. 2009). We found 371 000 tokens with the request *D N et N*. We annotated the number of nouns and determiners with Flemm (Namer 2000). Table 1 reports the numbers of different types of binomials.

The plural binomials are the most frequent ones. If N1 and N2 are both singular, D can be either singular ((13), (14)) or plural (15). With Dsg, the binominal can have either a joint (13) or a split reading (14).

(13) Le chanteur et poète québécois Gilles Vigneault
the.MSG singer.MSG and poet.SG Quebec.M Gilles Vigneault

construction	types	tokens
Dsg Nsg et Nsg	31412	51711
Dpl Nsg et Nsg	1308	5137
Dsg Nsg et Npl	5742	9490
Dpl Nsg et Npl	724	1432
Dpl Npl et Nsg	7586	13460
Dpl Npl et Npl	55269	201503
total	102041	282733

Table 1 Numbers of different binominal types in FrWAC

publie en France un livre d' entretiens.
publish.3SG in France a book of interviews
'The Quebec singer and poet, Gilles Vigneault, publishes a book of interviews in France.'
(FrWAC, `republique-des-lettres.fr`)

(14) Présentez -vous à la date et lieu
introduce.imp yourself at the.FSG date.FSG and place.MSG
indiqué pour suivre votre formation.
indicated.MSG to follow.inf your training.
'Introduce yourself at the date and place indicated to follow your training.' (FrWAC, `secours57.fr`)

(15) Les lieu et programme seront précisés
the.PL place.MSG and program.MSG be.FUT.3PL specified.MPL
sur le bulletin.
on the bulletin
'The place and program will be specified on the bulletin.'
(FrWAC, `rao.free.fr`)

We also found cases with a number mismatch. When N1 is singular and N2 is plural, Dpl (16a) is less frequent than Dsg (16b), which can be considered an instance of Closest Conjunct Agreement, or highest conjunct agreement.

(16) a. L' atelier est fermé le dimanche et
the workshop AUX. close.PST the.MSG Sunday and
jours fériés.
holiday.PL
'The workshop is closed on Sundays and public holidays.' (FrWAC, `fram.fr`)

b. Cette publication comporte les nom et
This.FSG publication contains the.PL last name and
prénoms du débiteur.
first name of.MSG debtor
'This publication contains the first and last names of the debtor.' (FrWAC, `courdecassation.fr`)

When N1 is plural and N2 singular, D is plural(17), which can be considered as a case of Closest Conjunct Agreement or resolution.[3]

(17) Voici pour les frères et sœur des photos
here is for the.PL brother.FSG and sister.FSG some picture
faites tout exprès pour eux.
do.PST.FPL all especially for them
'Here are pictures made especially for the brothers and sisters.' (FrWAC, `catherine-de-mercueil.over-blog.fr`)

3.2 Animacy and Semantic Readings

We extracted the binomials with more than five occurrences (22600 tokens) and removed the errors (10640 tokens left).[4] We annotated noun animacy with an external dictionary (Olivier Bonami, pers. comm.) and the joint or split reading manually. Table 2 reports the

[3]In FrWac, 88 cases were labeled as Dsg Npl and Npl, but all involve obvious errors. Thus, they are not listed in table 1.

[4]There are three major sources of errors: first, quite a lot of repeated tokens are found; second, some constructions which are in the form of *D N1 et N2* are not binomials, like *une heure et quart* (one.F hour.FSG and half.MSG); another major kind of errors is that the two bare nouns are not in the same constituent, like *mise sur le marché et prise en charge des vaccins* (placing.FSG on the.MSG market.MSG and taking.FSG care of the vaccines).

	joint		*split*		*total*	
	types	tokens	types	tokens	types	tokens
humans	196	2304	6	105	202	2409
non-humans	3	31	492	10535	495	10566
total	199	2335	498	10640	697	12975

Table 2 Readings of singular binomials in FrWAC

number of tokens with the joint/split reading for human and non-human nouns. For human nouns, 97.03% of the examples have a joint reading(13), while for non-human nouns, only 0.6% (18).

(18) Le restaurant et bar Starlight propose un
the.MSG restaurant.MSG and bar.MSG Starlight offer.3SG a
menu international.
menu international
'The restaurant and bar Starlight offers an international menu.'
(FrWAC, `expedia.fr`)

The results illustrate important frequency differences between human and non-human nouns regarding the interpretation. Human nouns favor the joint reading while non-human nouns favour the split reading. We can infer that human nouns may have more comprehension difficulties for the split reading because of the spurious ambiguity, which may lower its acceptability.

3.3 Animacy and Number Agreement

Then, we now examine the agreement strategies for the split reading with respect to human and non-human nouns. Our results (table 2) report more examples with a plural D for human nouns (19), even if both singular and plural D are quite rare. However, for non-human nouns, the singular D is more frequent (see (14) and (15)).

	Dsg		*Dpl*		*total*	
	types	tokens	types	tokens	types	tokens
human	1	6	5	99	6	105
non-human	439	7507	53	2997	492	10535
total	440	7513	58	3096	498	10640

Table 3 Dsg/Dpl with singular binomials in FrWAC

(19) a. Les mari et femme sont d' accord
the.PL husband.MSG and wife.FSG be.3PL of agreement
sur le partage des biens.
on the division of.PL property.PL
'The husband and wife agree on the division of these properties.' (FrWAC, `judiciaire.blog.20minutes.fr`)

b. Elle part loin à la recherche de nourriture pour
she leaves far to the research of food for
son mari et bébé.
her.MSG husband.MSG and baby.MSG
'She leaves far looking for food for her husband and baby.' (FrWAC, `harunyahya.fr`)

We suppose that this difference can be accounted for by a tendency to avoid ambiguity. For human nouns, the singular D would favor a joint reading, which explains why it is quite rare with a split reading in the corpus. However, for non-human nouns, both the singular and plural D are possible with a split reading, and the singular D is more frequent, since non-human binomials are not ambiguous (or rarely ambiguous) between a joint and a split reading.

3.4 Animacy and Syntactic Functions

The syntactic position of the binomial may play a role as well. Since split binomials yield plural verb agreement (15), this might also favor plural D agreement. In (14), the split binomial is a complement and has a singular D. Similarly for animates, in (19a), the split binomial

	subj	inv_ subj	comps	total
Dpl	42	17	503	562
Dsg	6	3	411	420
total	48	20	914	982

Table 4 Agreement and syntactic function of two singular binomials in FrWAC

is a subject with a plural D, and in (19b) it is a complement with a singular D.

Animates tend to be more often in subject position than inanimates (Clark & Begun 1971). Previous studies have examples of (animate) binomials in subject position only (e.g., see (7b) and Le Bruyn & de Swart 2014 and Heycock & Zamparelli 2005). So it may be the independent tendency for animates to be in subject position that favors the plural D.

Since human nouns have a strong preference for plural D, and the singular D is quite rare as illustrated in table 3, it is difficult to compare the interaction between number agreement and syntactic position for human nouns. Thus, we only compare D number agreement for inanimate nouns in different syntactic positions in FrWac. In order to balance the singular/plural D, we chose two pairs of inanimate nouns which appear with both singular and plural D: *D nom et prénom* ('D name and first name') (324 tokens of Dsg and 513 tokens of Dpl), *D date et heure* ('D date and hour') (96 tokens of Dsg and 49 tokens of Dpl) and manually annotated their syntactic position: (preverbal) subject, inverted-subject, complement.

Table 4 shows that there are significantly more examples in complement position, and the plural D is preferred in subject position. But in complement position, the difference between singular/plural D is not significant.

Our results are consistent with the claim that conjoined human nouns favor plural agreement crosslinguistically (Corbett 1991). How-

ever, there are not enough examples of human nouns for a statistical study, and other semantic or pragmatic factors, such as topicality, definiteness, distinctness, may play a role too. This is why we decided to run two acceptability rating experiments to test the interaction of animacy and agreement carefully controlling agreement, and syntactic functions.

4 Two Acceptability Rating Experiments

In this section, we report two quantitative gradient acceptability judgment studies for D agreement in French singular split binomials. The first study focuses on a set of sentences with the binomials in subject position, with a plural verb, with half of the examples human and the other half, non-human. It aims to test the role of animacy, excluding other potential linguistic factors in a context forcing the split reading.

The second study is designed to test the role of syntactic function, and to test acceptability of binomials without plural priming. As in the first study, the binomials in subject position involve plural verb agreement, whether the verb is singular or plural may have an effect on the NP-internal agreement. In Experiment II, the experimental items are objects of a singular verbs, without any other plural agreement. In comparing these two studies, we can test the effects of this syntactic function bias. Since the object context is neutral between a split and a joint reading, we chose human and non-human nouns which cannot have a joint interpretation.

4.1 Experiment I

4.1.1 Materials

We constructed two lists of 12 experimental items, one with 12 human singular binomials (20a) and the other with 12 non-human singular binomials (20b), inspired from the corpus FrWac. The experimental items are in subject position and the verb is collective and in the plural form in order to force a split reading. The D is either singular or plural.

(20) a. **D[-human]:** Il arrive souvent que
it happen.3SG often that
vos/votre identifiant et mot de passe ne
your.PL/your.SG username.MSG and password.MSG NEG
soient pas reconnus par le site.
be.SUBJV.3PL NEG recognized.PL by the site
'It often happens that your username and password are not recognized by the site.'

b. **D[+human]:** Les/Le directeur et
the.PL/the.MSG director.MSG and
sous-directeur du secteur se sont
assistant director.MSG of.MSG sector.MSG REFL AUX.3PL
mis d' accord sur le projet.
put.M of agreement on the project
'The director and assistant director of the sector agreed on the project.'

We also included six control items without coordination. For each item, one version appears in a grammatical item (21a) and the other in an ungrammatical item (with number agreement error) (21b).

(21) a. **Control1:** La tête dans les genoux, je
the.FSG head.FSG in the.MPL knee.PL, I
dormirais peut-être deux heures.
sleep.COND.1SG perhaps two hour.PL

b. **Control2:** La tête dans le genoux, je
the.FSG head.FSG in the.MSG knee.PL, I
dormirais peut-être deux heures.
sleep.COND.1SG perhaps two hour.PL

We also included a set of 15 items from another unrelated experiment as fillers. Each item has three conditions and there are 45 sentences in total.

4.1.2 Procedure

43 subjects participated in the experiment, recruited from the website RISC (`http://www.risc.cnrs.fr/`). Participants have to fill a survey before starting the experiments, reporting mainly their mother language, age, sex, etc. One participant was removed as he was not a native speaker and 42 were retained. Participants were asked to rate the acceptability of each sentence, with a Likert scale from 1 (completely unacceptable) to 10 (perfectly acceptable), which is the usual scale in the French school system.

The items are generated using a Latin square. Participants could only see one possible *D* (singular/plural) for each item, the number of which was counterbalanced across participants. The order of experimental items is also randomized in each trial.

4.1.3 Results

The results (figure 1) show that plural D are more acceptable than singular D, for both human and non-human nouns. Non-human binomials with a plural D (mean: 7.87) are more acceptable than the grammatical control items (mean: 7.3). The singular D (mean: 6.96) is less acceptable, but better than the ungrammatical control sentences (mean: 6.0). For human binomials, a plural D (mean: 7.13) is also judged better than a singular D (mean: 6.39).

We ran a maximum mixed-effects linear model with items and participants included as random factors (Barr et al. 2013), using *lmer4* package in R. The dependent variable is the mean of participants' ratings, and independent predictors are Noun animacy and Dsg/Dpl. There are significant main effects of D number ($p = 0.03$) and N animacy ($p = 0.01$): non-human nouns are more acceptable than the human nouns. And plural D are significantly more acceptable than singular ones. We didn't detect any effect of participants' age or social gender.

4.1.4 Discussion

The results of Experiment I are consistent with our corpus data in the sense that inanimate (split) binomials are more acceptable than

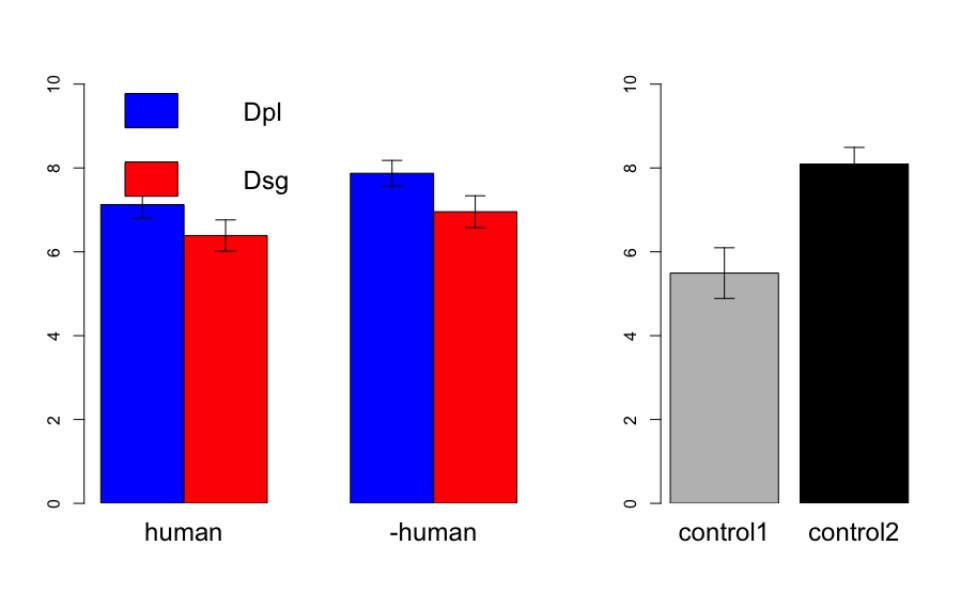

Figure 1 Acceptability ratings of subject binomials with Dsg/Dpl

animate (split) binomials.

Furthermore, the plural D is preferred for human nouns, which is in line with our hypothesis because the singular D can be ambiguous between a split and joint reading. However, contrary to the corpus data, the plural D is also preferred for inanimate nouns in this experiment. We suppose that this may come from a semantic bias: we use plural verbs to force a split reading, so the plural verb agreement may have an effect on the preference for a plural D. We thus ran a second experiment with the binomials in object position.

4.2 Experiment II

We used the same experimental items as Experiment I but in object position. It is not possible to use the same context as Experiment I, but we tried to keep the contexts similar. We took the same pairs of nouns, 12 animate and 12 inanimate.[5]

[5]In the glosses, REF.1SG represents the reflexive pronoun of first person singular.

(22) a. **D[-human]:** Chaque fois que je me connecte, je
each time that I REF.1SG log-in, I
dois taper mes/mon identifiant et
must type my.PL/my.MSG username.MSG and
mot de passe.
password.MSG
'Every time I log in, I need to type in my username and password'

b. **D[+human]:** Il faudrait pouvoir prévenir
it should be-able inform
les/le directeur et sous-directeur
the.PL/the.MSG director.MSG and assistant director.MSG
de l'établissement.
of the establishment
'It is necessary to be able to inform the director and assistant director of the establishment.'

The control items are the same as in Experiment I. We also included 24 items, for a total of 48 sentences as fillers.

4.2.1 Results

The procedure is the same as in Experiment I. We had 51 participants recruited from the RISC, all of them native speakers of French.

Figure 2 reports the mean for each condition of Experiment II. For non human nouns, both plural D (mean: 6.99) and singular D (mean: 6.83) are acceptable, but a little lower compared to the grammatical sentences (mean: 8.09). Human nouns are less acceptable than non-human nouns, both for plural D (mean: 5.97) and singular D (mean: 5.47), and singular D is even worse than the ungrammatical items (mean: 5.49).

Table 6 reports the results of the maximum mixed-effects linear model with the participants' rating as the dependent variable and D number and N animacy as independent variables. Only animacy has a significant effect ($p < 0.001$): non-human nouns are more acceptable than human nouns. However, we don't detect a significant effect

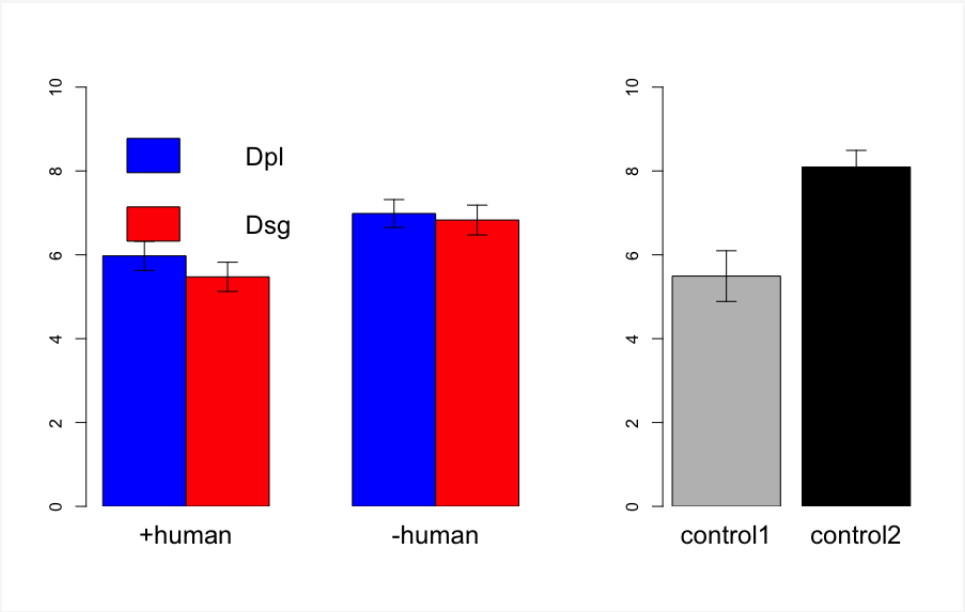

Figure 2 Acceptability ratings of object binomials with Dsg/Dpl

of D number ($p = 0.19$): there are no significant differences between singular and plural D in object position.

In object position, the preference for plural D doesn't exist anymore. Consistent with the first experiment, non-human nouns are still more acceptable than human nouns. However, contrary to Corbett (1991), we didn't detect any interaction between D agreement and N animacy, which means that the plural and singular D are judged in a similar way for both human and non-human nouns, something for which we are not able give an explanation in this paper.

5 General Discussion

5.1 Corpus Data and Experimental Data

We did a statistical comparison between Experiments I and II using a maximum mixed-effects linear model. The fixed effects consist of three factors: animacy, syntactic function and D number, as well as the interaction among them. For random effects, we included D

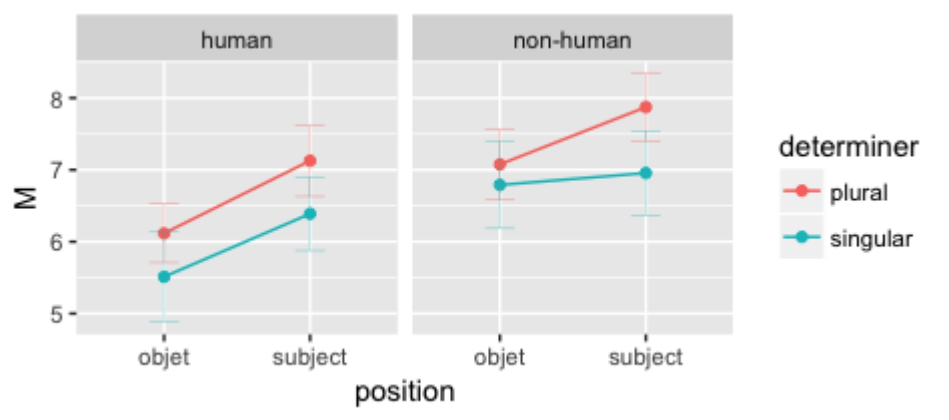

Figure 3 ggplot of the maximum linear model comparing Experiments I and II

number and noun animacy and their interaction for subjects, and the syntactic function of N and D number as well as their interaction for items.

As showed in figure 3, there are significant effects for D number ($p = 0.03$), animacy ($p < 0.001$) and syntactic function ($p < 0.001$). The interaction between animacy and syntactic function is also significant ($p = 0.04$).

We conclude that the acceptability of singular binomials is not a categorical phenomenon, but sensitive to several factors. In what follows, we would like to offer two considerations. First, the effects of animacy and syntactic position can be explained in terms of language processing difficulties caused by a tendency to avoid potential ambiguity. As a result, the difference of number agreement preferences between the corpus data and the experimental results may come from the bias that animate binomials are more frequent in subject position.

5.1.1 Semantic Ambiguity

The corpus data (section 3.2) have shown the different preferences for joint/split reading regarding animacy. Split is a higher, frequent dominant reading with non-human nouns, while for human nouns,

it is a lower-frequency subordinate reading. When the second conjunct of the coordination appears, which was always congruent with the split reading, this will give rise to processing difficulties especially for the human nouns. Thus, this incremental processing difficulty may lead to lower judgments of acceptability.

Furthermore, recent models of ambiguity resolution (MacDonald et al. 1994, Tabor et al. 1997, McClelland et al. 1989) agree that readers are able to make use of available contextual information to help them activate and integrate the appropriate meaning of an ambiguous word or an ambiguous syntactic structure. In Experiment I, we use plural verbs to force the binomials in the subject position to have a split reading. The result illustrates that it is more acceptable with a context biased toward split reading (subject position with a plural verb) than with a neutral context (object position after a singular verb). We then explain the effects of syntactic function by its bias for the split reading: the binomials are more acceptable when the contextual information provides information that facilitates resolving the joint and split reading ambiguity.

We also detect a significant interaction between animacy and syntactic function. This means that the differences between subject and object position are greater for human nouns than for non-human nouns. As shown in figure 3, there are important differences between subject and object position for human nouns, while the differences are quite tiny for non-human nouns, especially for the singular determiner. We can interpret this effect by saying that the acceptability of human binomials depends more on the disambiguated context.

5.1.2 Agreement and Animacy

The empirical data revealed that the determiner can agree in number with the whole coordination (plural), or with the closest conjunct (singular). The preference for singular or plural D varies according to the syntactic position. Experiment I shows that a plural D is more acceptable with either human or non-human nouns in the subject position, while there is no preference for singular or plural D in

object position in Experiment II.

The experimental data (section 4) are consistent with the corpus data (section 3) in that for non-human nouns, the plural D is preferred in subject position, while in object position, there is no difference between singular and plural D.

5.2 French Data and Previous Analyses

5.2.1 Optimality Theory

Le Bruyn & de Swart (2014) develop an analysis that supposes a different syntactic structure depending on the meaning. Using Optimality Theory (OT), they propose a ranking of constraints (23a), based on the typology of de Swart & Zwarts (2008) and de Swart & Zwarts (2009). The ranking of a specific constraint for coordination (*FunctSCoordP) with respect to a more general markedness constraint (*FunctN) and three faithfulness constraints ((23b)–(23d)) determine the article use with split and joint coordination readings.

(23) a. {FPl, *FunctSCoordP} >> {FDr, FDef} >> *FunctN
b. FPl: reference to a plurality of individuals must be reflected in the form
c. FDr: the presence of a discourse referent in the semantics corresponds to an expression that carries discourse referential force
d. FDef: reference to discourse unique individuals requires the use of an expression of definiteness

*FunctN is a general constraint to avoid functional structure (D) in the nominal domain, of which *FunctNSCoordP is a specific subconstraint to avoid functional structure (D) on top of a split coordination, thus yielding *Man and wife were happy together*, in languages like English which do not allow bare nominal arguments (with singular count nouns). Since *FunctNSCoordP is ranked above *FunctN, coordination is an exception to the general rule enforced by the faithfulness constraints (requiring an explicit marking of definiteness). Since it is ranked above the faithfulness constraints FDr and FDef, split co-

ordination relaxes the usual requirements on definite and indefinite articles.

The ranking of constraints may vary crosslinguistically. In languages like English or French, *FunctN is ranked low with respect to the faithfulness constraints governing article use and *FunctNSCoordP is ranked high, thus Split-Coordination Phrases behave differently from Joint-Coordination Phrases. Le Bruyn & de Swart (2014) proposed that for the joint reading, there is a DP projection above the coordination, and for the split reading (24a), the first noun is combined with the D to form a DP, which is combined with the second noun (24b).

(24) a. $[_{\text{CoordP}}[_{\text{DP}}$ D [NP and$_{\text{joint NP}}$ NP]]
b. $[_{\text{CoordP}}[_{\text{DP}}$ D NP] and$_{\text{split}}$ $[_{\text{NP}}$ NP]]

The analysis of Le Bruyn & de Swart 2014 relies on linguistic intuitions and is based on the assumption that only one option is posible in a given language, which is unable to account for our results. Our empirical results suggest that number agreement for the joint reading is categorical, since only a singular D is allowed. However, both singular and plural D are allowed for the split reading, and which one is used is rather a matter of preference (cf. Gries 2003, Bresnan & Nikitina 2009, Bresnan & Hay 2008). Multiple factors, like semantic interpretation, animacy and the syntactic function of the binominal, play a role for computing agreement.

These constraints would probably be best represented with a competition algorithm (cf. MacWhinney et al. 1984, MacWhinney & Bates 1989, MacDonald et al. 1994). We rely on previous work within LFG (see Bresnan 1982, Dalrymple & Hristov 2010 and Kathol 1997 for an overview) and HPSG (Pollard & Sag 1994) to sketch an HPSG analysis.

5.2.2 An HPSG Analysis

An & Abeillé (2017) have proposed a simplified HPSG analysis for joint and split binomials. We use a general schema for coordina-

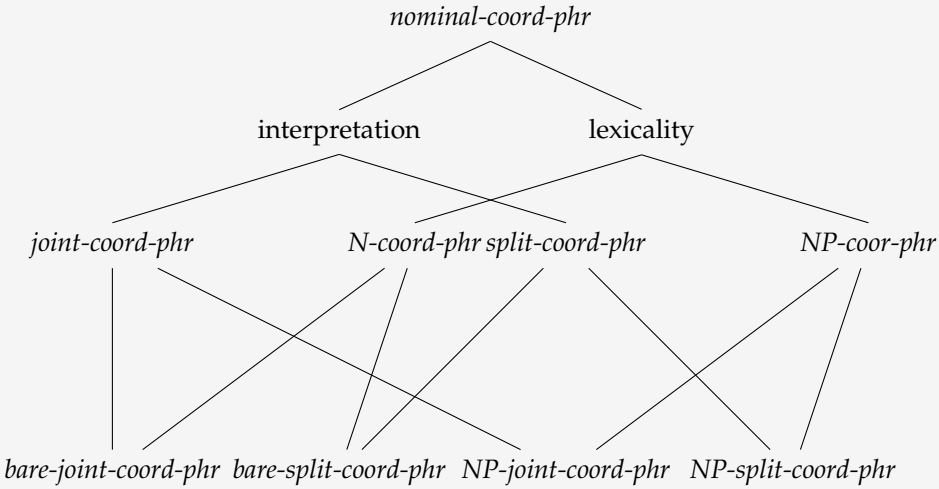

Figure 4 Hierarchy of nominal coordinate phrases

tion, which is n-ary phrase (Borsley 2005, Abeillé 2005). The conjunction forms a subconstituent with the following conjunct, but it is not the head of the coordination, which is a subtype of *unheaded phrase*. We use a CONJ feature to distinguish conjuncts with a conjunction, and conjuncts without. *nelist* stands for non-empty-list, and DTRS (Daughters) for immediate constituents.

(25) $$\textit{coord-phrase} \Rightarrow \left[\text{DTRS} \left\langle \textit{nelist} \left(\left[\text{CONJ nil}\right]\right) \oplus \textit{nelist} \left(\left[\text{CONJ} \neq \text{nil}\right]\right)\right\rangle\right]$$

The syntactic structure is the same for split and joint coordination. We propose a cross-classification that distinguishes interpretation (split or joint) and lexicality (bare or NP coordination). This leads to the hierarchy in figure 4.

As proposed in Kathol 1997 and Wechsler & Zlatić 2003, we distinguish CONCORD and INDEX features. As illustrated in example (26) (repeated from (3a)), the verb shows INDEX agreement (plural) while the determiner shows CONCORD agreement (singular).

(26) This [boy and girl] are eating a pizza.

For *joint-coord-phr*, the schema imposes sharing of CONCORD and INDEX features:

(27) *joint-coord-phr* $\Rightarrow$
$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CONCORD} & \boxed{1} \\ \text{INDEX} & \text{i} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{CONCORD} & \boxed{1} \\ \text{INDEX} & \text{i} \end{bmatrix} \cdots \begin{bmatrix} \text{CONCORD} & \boxed{1} \\ \text{INDEX} & \text{i} \end{bmatrix} \right\rangle \end{bmatrix}$$

This is compatible with bare nouns (28a) and NP coordination (28b).

(28) a. Un [collègue et ami] est venu hier.
a.MSG colleague and friend AUX. come.PST yesterday
'One colleague and friend come yesterday.'

b. [Un grand poète et un grand homme]
a.MSG great.MSG poet.SG and a.MSG great.MSG man
est mort hier.
is dead yesterday
'A great poet and a great man is dead yesterday.'

For *split-coord-phr*, the schema is (29) and the INDEX features are not shared:

(29) *split-coord-phr* $\Rightarrow$
$$\begin{bmatrix} \text{CONT} & \left[\text{INDEX} \quad \text{i} \oplus \ldots \oplus \text{n}\right] \\ \text{DTRS} & \left\langle \left[\text{INDEX} \quad \text{i}\right] \ldots \left[\text{INDEX} \quad \text{n}\right] \right\rangle \end{bmatrix}$$

For bare split nominals, the schema in figure 5 allows both plural and Closest Conjunct Agreement. Following Villavicencio et al. (2005), we use RAGR for right-agreement (rightmost noun, N2) and LAGR for left-agreement (leftmost noun, N1). We consider the nouns to be the syntactic head, and use the SPR feature (Specifier) for D selection. We ignore gender features here, for which Abeillé et al. (2018) show that only Closest Conjunct Agreement is allowed (recall (14)).

The schema *bare-split-coord-phr* is illustrated in figure 6 for *votre/vos*

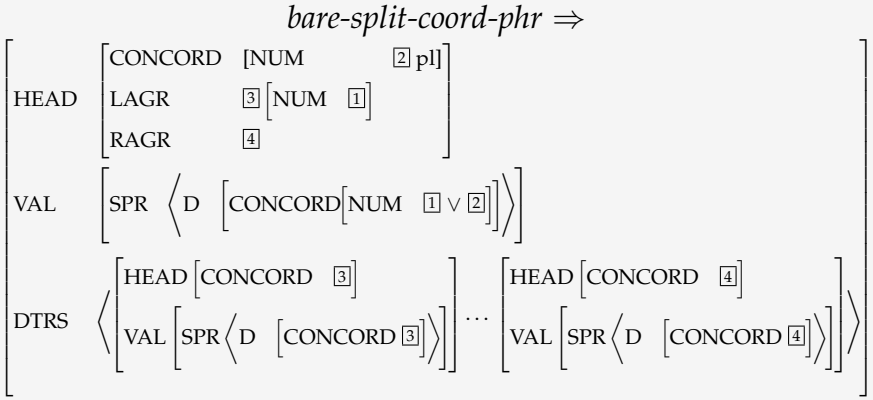

Figure 5 Schema for *bare-split-coord-phr*

nom et prénom. Notice that the closest conjunct (the leftmost noun) is also the highest.

However, this analysis also has its limits. It accounts for the categorical difference between joint and split reading, and for the different agreement possibilities, but it does not take into account the preferences for the singular/plural D in the split case.

6 Conclusion

Previous studies have addressed number agreement of binominal coordinations in different languages (e.g., Le Bruyn & de Swart 2014, King & Dalrymple 2004, Demonte & Perez-Jimenez 2012, Villavicencio et al. 2005). We use large-corpus data and experimental results to investigate this question. Furthermore, we explored semantic and syntactic factors that may affect number agreement of the shared determiner.

Our empirical findings are twofold. First, we show that Closest Conjunct Agreement is attested in French and is not superficial since it is sensitive to syntactic (word order and syntactic function) and semantic factors (noun animacy). Since in D-Noun agreement the target is before the controller, the possibility of Closest Conjunct Agreement is expected by Corbett's agreement hierarchy, especially for non-human nouns. As the closest noun is also the highest one,

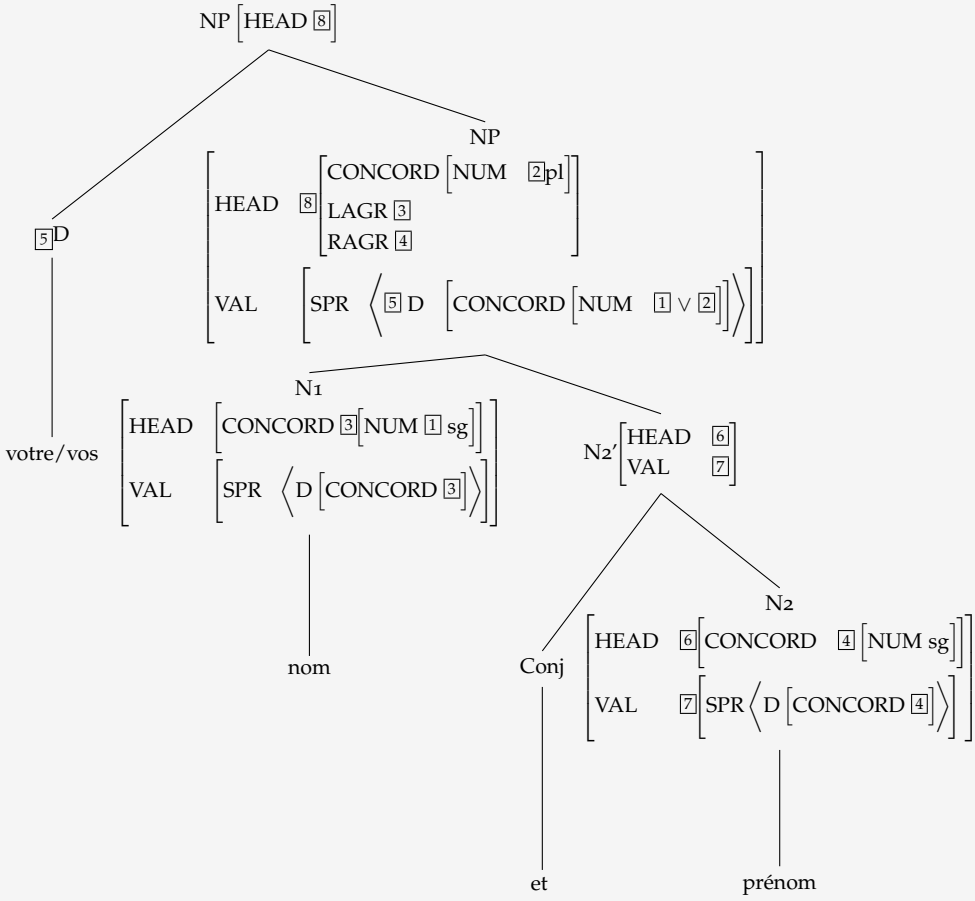

Figure 6 An illustration of *bare-split-coord-phr*

it is also expected under a structural account of agreement (Marušič et al. 2007). However, a plural D is preferred in subject position for both human and non-human nouns, contrary to other Romance languages, but no differences are observed in the object position.

Second, the binomial is ambiguous between a joint or a split reading, and human nouns favor the joint reading whereas non-human nouns favor the split reading. In the split context, human nouns are less acceptable than non-human nouns. Furthermore, they are more acceptable in a disambiguated context, for instance, in subject position.

Third, we argue that a traditional approach using uncontrolled grammatical judgments can lead to empirical inadequacy as in the OT approach of Le Bruyn & de Swart (2014).

Finally, new questions arise: whether Closest Conjunct Agreement is also possible in other types of agreement in French, such as noun-adjective agreement and subject-verb agreement, and whether it is sensitive to the same factors.

Acknowledgments We thank the editor, Christopher Pinon, as well as the anonymous reviewer for their feedback on previous versions of this paper. This work is supported by strand 2 of LabEx EFL(ANR-10-LABX-0083) and China Scholarship Council. Special thanks to Olivier Bonami, Berthold Crysmann, Jean-Pierre Koenig, Jeremy Kuhn and Bert Lebruyn for helpful comments and discussions. We are also grateful to the audience of the CSSP 2017 conference.

References

Abeillé, A. (2005). Les syntagmes conjoints et leurs fonctions syntaxiques. *Langages*, *160*(4), 42–66.

Abeillé, A., An, A., & Shiraïshi, A. (2018). L'accord de proximité du déterminant en français. *Discours*, *22*.
URL `http://journals.openedition.org/discours/9542`

An, A., & Abeillé, A. (2017). Agreement and interpretation of binominals in French. In S. Müller (Ed.) *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky,*

Lexington, (pp. 26–43). Stanford, CA: CSLI Publications.
URL http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-an-abeille.pdf

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: A collection of very large linguistically processed web-crawled corpora. *LREC Conference*, *43*(3), 209–226.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Borsley, R. D. (2005). Against ConjP. *Lingua*, *115*(4), 461–482.

Bresnan, J. (Ed.) (1982). *The Mental Representation of Grammatical Relations*. The MIT Press.

Bresnan, J., & Hay, J. (2008). Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua*, *118*(2), 245–259.

Bresnan, J., & Nikitina, T. (2009). The gradience of the dative alternation. In L. U. . L. H. Wee (Ed.) *Reality Exploration and Discovery: Pattern Interaction in Language and Life*, (pp. 161–184). Stanford, CA: CSLI.

Clark, H. H., & Begun, J. S. (1971). The semantics of sentence subjects. *Language and Speech*, *14*(1), 34–46.

Comrie, B. (1989). *Language Universals and Linguistic Typology: Syntax and Morphology*. The University of Chicago Press.

Corbett (1991). *Gender*. Cambridge University Press.

Croft, W. (2002). *Typology and Universals*. Cambridge University Press.

Dalrymple, M., & Hristov, B. (2010). Agreement patterns and coordination in Lexical Functional Grammar. In M. Butt, & T. H. King (Eds.) *Proceedings of LFG10*, (pp. 187–207).

de Swart, H., & Zwarts, J. (2008). Article use across languages: An OT typology. In A. Grønn (Ed.) *Proceedings of Sinn und Bedeutung 12*, (pp. 628–644). University of Oslo.

de Swart, H., & Zwarts, J. (2009). Optimization principles in the typology of number and articles. In B. Heine, & H. Narrog (Eds.) *The Oxford Handbook of Linguistic Analysis*, (pp. 555–581). Oxford University Press.

Demonte, V., & Perez-Jimenez, I. (2012). Closest conjunct agreement in Spanish DPs: Syntax and beyond. *Folia Linguistica*, *46*(1), 21–73.

Gries, S. T. (2003). Towards a corpus-based identification of prototypical

instances of constructions. *Annual Review of Cognitive Linguistics*, (pp. 1–27).

Haspelmath, M. (2013). Occurrence of nominal plurality. In M. S. Dryer, & M. Haspelmath (Eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL `https://wals.info/chapter/34`

Heycock, C., & Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural Language Semantics*, *13*(3), 201–270.

Kathol, A. (1997). Agreement and the syntax-morphology interface in HPSG. In R. Levine, & G. Green (Eds.) *Studies in Contemporary Phrase Structure Grammar*, (pp. 223–274). Cambridge University Press.

Kayne, R. S. (1994). *The Antisymmetry of Syntax*. The MIT Press.

King, T. H., & Dalrymple, M. (2004). Determiner agreement and noun conjunction. *Journal of Linguistics*, *40*(1), 69–104.

Le Bruyn, B., & de Swart, H. (2014). Bare coordination: The semantic shift. *Natural Language and Linguistic Theory*, *32*(4), 1205–1246.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

MacWhinney, B., & Bates, E. (1989). *The Crosslinguistic Study of Sentence Processing*. Cambridge University Press.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 127–150.

Marušič, F., Nevins, A., & Saksida, A. (2007). Last-conjunct agreement in Slovenian. In R. Compton, M. Goledzinowska, & U. Savchenko (Eds.) *Formal Approaches to Slavic Linguistics*, vol. 15, (pp. 210–227). Ann Arbor: Michigan Slavic Publications.

McClelland, J. L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, *4*(3-4), 287–335.

Murphy, A., & Puškar, Z. (2018). Closest conjunct agreement is an illusion. *Natural Language and Linguistic Theory*, *36*(4), 1207–1261.

Namer, F. (2000). FLEMM: un analyseur flexionnel du français à base de règles. *Traitement automatique des langues*, *41*(2), 523–547.

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, *12*(2-3), 211–271.

Villavicencio, A., Sadler, L., & Arnold, D. (2005). An HPSG account of closest conjunct agreement in NP coordination in Portuguese. In S. Müller (Ed.) *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, (pp. 427–447). Stanford, CA: CSLI Publications.

Wechsler, S., & Zlatić, L. (2003). *The Many Faces of Agreement*. Stanford, CA: CSLI Publications.

Zaenen, A., Carletta, J., Garretson, G., Bresnan, J., Koontz-Garboden, A., Nikitina, T., O'Connor, M. C., & Wasow, T. (2004). Animacy encoding in English: Why and how. In *Proceedings of the 2004 ACL Workshop on Discourse Annotation*, (pp. 118–125). Association for Computational Linguistics.