

Agreement and interpretation of binominals in French

Abstract

This paper investigates the agreement of coordinated binominals in French in the form *Det N1 et N2*. We provide corpus data and experimental data to show that different agreement strategies exist in French and that they challenge previous syntactical analyses of binominals. We then propose an HPSG analysis to account for French binominals.

1 Introduction

Binomial coordinations such as *Det N1 et N2* have raised many discussions (Heycock & Zamparelli 2005; Le Bruyn & de Swart 2014; King and Dalrymple 2004, Arnold et al. 2006). Semantically, binominals can have two distinct readings: a joint reading (1a, colleague and friend are co-referent) and a split reading (1b, with two distinct individuals: boy and girl).

- (1) a. This friend and colleague is a nice person.
b. This boy and girl are coming tonight.

The split binominals are an instance of natural coordination (e.g., Haiman 1983, Wälchli 2005...), in which the coordinated parts express semantically closely associated concepts. Not all bare coordinations are equally felicitous (2a, b).

- (2) a. this boy and girl
b. *this boy and cat

Le Bruyn and de Swart (2014) develop special matchmaking semantics for split *Det N1 & N2* constructions: the discourse referent for the second conjunct is related to the (discourse) referent introduced by the DP in the first conjunct, and vice versa.

- (3) this brother and sister
(they are brother and sister of each other)

Syntactically, two structures have been proposed for split reading, either [*Det [N1 and N2]*] (King and Dalrymple 2004) or [*Det N1 and N2*] (Le Bruyn and de Swart 2014).

As for determiner agreement, languages may have different strategies: a singular determiner requires the conjuncts to be

singular in English, and conjuncts with different number cannot be coordinated (4a) (King and Dalrymple 2004), whereas Spanish exhibits closest conjunct agreement (4b) (Demonte et al. 2012)

- (4) a. This boy and girl.
*This boy and girls
* These boys and girl
b. [{EI/*Los} abdomen y
the.M.SG/M.PL abdomen.M.SG and
pecho] aparecen relativamente
chest.M.SG appear.PRS.3PL relatively
abultados.
swollen
'The abdomen and chest look relatively
swollen.'

This paper will make two points: on the one hand, we provide new data for French binominals, which challenge Le Bruyn and de Swart (2014)' syntactic analysis; on the other hand, we propose a preliminary HPSG analysis for French binominals.

2 Empirical evidence of binominal agreement in French

In French, for singular joint reading, like in many languages, only the singular determiner is allowed (5). For split reading, Heycock & Zamparelli (2005) and Le Bruyn & de Swart (2014) assume that French is an exception, as singular nouns are infelicitous (6a), while only plurals are allowed (6b)

- (5) Le/*les collègue et ami
the.SG/PL colleague.SG and friend.SG
de Jean est venu hier
of Jean PAST come.SG yesterday.
'The colleague and friend of Jean came
yesterday.'

- (6) a. *Ce/*Ces marin et
this.MSG/PL sailor.MSG and
soldat sont souvent ensemble.
soldier.MSG be.PL often together
b. Ces marins et soldats
this.PL sailor.MPL and soldier.MPL
sont souvent ensemble
be.PL often together.
'These sailors and soldiers are often together'

The examples in H&Z (2005) and B&S (2014) only consist of animate nouns. We will present a corpus study (corpus frWac) and an experiment of acceptability judgments challenging these data for singular nouns. For plural nouns, we assume that both joint and split readings are allowed: example (6b) is ambiguous between joint and split reading.

2.1 Corpus data

In frWac (1.6 billion words, Baroni et al. 2009), we found 371.000 tokens (96612 types) for the construction *Det N1 et N2*. We annotate the number of Det and Nouns with *Flemm* (Namer 2000). There are 51711 tokens (31412 types) for *Dsg N1sg et N2sg* with either joint reading or split reading (7, 8), 5137 tokens (1308 types) for *Dpl N1sg et N2sg* with only split reading (9).

(7) Le chanteur et poète québécois
the.SG singer.SG and poet.SG Quebec
Gilles Vigneault publie en France
Gilles Vigneault publish.3SG in France
un livre d'entretiens
a.MSG book of interviews
'The singer and poet of Quebec, Gilles Vigneault, publishes a book of interviews in France'

(8) Présentez-vous à la date
introduce yourself at the.FSG date.FSG
et lieu indiqué pour
and place.MSG indicated.MSG to
suivre votre formation.
follow your.SG training.
'Introduce yourself at the date and place indicated to follow your training.'

(9) Les lieu et programme
the.PL place.MSG and program.MSG
seront précisés sur le bulletin.
be.FUT.PL specified.PL on the bulletin
'The places and programs will be specified on the bulletin'

We extracted the binominals with more than five occurrences and removed the errors. We annotated noun animacy with an external dictionary (Bonami pc.) and the joint or split reading manually.

The results (fig. 1) show that for the joint reading, only the Dsg is allowed, whereas both Dsg and Dpl are allowed for the split reading:

3084 tokens (60 types) for Dpl, 7563 tokens (456 types) for Dsg.

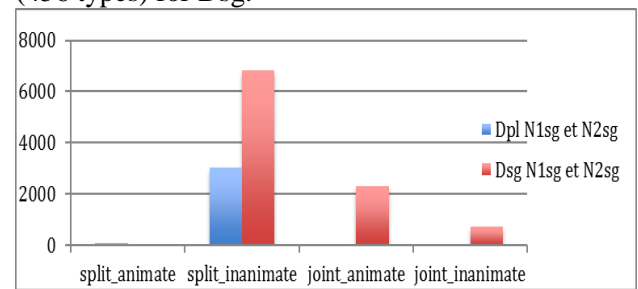


Figure 1. Animate/inanimate binominals and joint/split readings in frWac

Furthermore, there is an interaction with animacy: joint reading is more frequent with animate than inanimate nouns. For split reading, there is also an interaction between Det agreement and animacy: for split animate binominals, plural determiners are preferred in a two-tailed binomial test ($p < .001$), whereas singular determiners are preferred ($p < .001$) for split inanimate binominals.

Our hypothesis is that Dsg is dispreferred for split animates in order to avoid the joint reading.

2.2 Acceptability judgment experiment

We designed an experiment of acceptability judgment to test the acceptability of Dsg/Dpl for split singular binominals. We had 12 singular animate binominals (10) and 12 singular inanimate binominals (11), as well as 6 control items (12). These items were inspired by corpus data. Participants could only see one D number for each binomial, the number of which was counterbalanced across participants. The binominals are in subject position and the predicate is plural and collective, in order to force the split reading:

(10) Le/Les directeur et
the.SG/PL director.SG and
sous-directeur du secteur se
underdirector.SG of.MSG sector REFL
sont mis en accord sur le projet
PAST put in agreement on the project
'The director and underdirector of the sector agreed on the project.'

(11) Il arrive souvent que votre/vos
it happens often that your.SG/PL
identifiant et mot de passe ne
username.SG and password.SG NEG

soient pas reconnus par le site.
 be_SBJV.PL NEG recognized.PL by the site
 ‘It often happens that your username and password are not recognized by the site.’

(12) La tête dans le/les genoux,
 the.FSG head.FSG in the.MSG/PL knee.PL,
 je dormirais peut-être deux heures.
 I sleep.COND.1SG perhaps two hour.PL
 ‘With the head in the knee, I would sleep perhaps two hours.’

We had 42 participants, all French native speakers, recruited on the website RISC. The results (Fig.2) shows that the acceptability of experimental items slightly lower than good controls (without coordination) in green, but much higher than bad controls (in yellow)

We analyze the data using a mixed-effect linear regression model. Our dependent variable is participant’s acceptability judgment (range from 1-10), independent predictors are noun animacy and determiner number. We find significant effects for both animacy and determiner number and there is no interaction between these factors. Dpl is better judged than Dsg both for animate nouns and inanimate nouns. Meanwhile, we also find that animacy plays a role, since inanimate nouns are more acceptable than animate nouns.

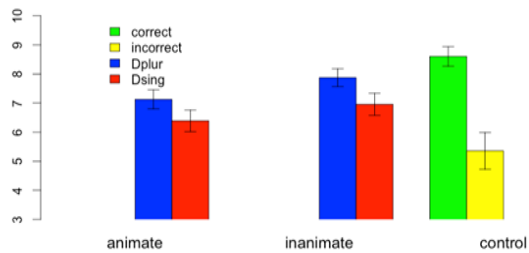


Figure 2 Acceptability judgement for D Nsg et Nsg for split reading

3 Syntactic structures of binominals in French

Le Bruyn and de Swart (2014) propose two different syntactic structures depending on the meaning. For joint reading, Det lives in a position above the joint Coord phrase (13a). For split reading, Det combines with the first conjunct only (13b), predicting thus ungrammaticality of Dpl when followed by two coordinated Nsg, as in English (14a), Spanish (14b) and supposedly for French (14c).

(13) a. joint reading: [DP D [CoordP NP and NP]]

b. split reading: [CoordP [DP D NP] and NP]

- (14) a. *These boy and girl are eating a pizza
 b. *Los abdomen y pecho
 the.M.PL abdomen.MSG and chest.MSG
 c. *Les homme et femme sont venus
 the.PLman andwoman PAST come.PL
 ‘The man and woman have come.’
 (B&S’s judgement)

According to B&S (2014), (13b) is also compatible with bare binomials N1 et N2 (Roodenburg 2005), which only have the split reading (*Nom et prénom doivent être écrits en noir*. ‘Name and first name must-pl be written in black’)

We agree with B&S’s syntactic structure for joint reading. However, for split reading, we argue that there is a Det above coordinated nouns for split reading as well. On the one hand, the data presented above show that Dpl is acceptable in French, as long as the two N form a natural pair (15a). Moreover, plural numerals may be used: example (15b) refers to a pair, one brother and one sister.

- (15) a. Les mari et femme sont
 the.PL husband.SG and wife.SG be.PL
 d’accord sur le partage des biens.
 agree on the division of.PL property.PL
 ‘The husband and wife agreed on the division of these property.’
 b. [Mes deux [frère et sœur]]
 my.PL two brother.SG and sister.SG

Our data show that Dsg is also possible for split reading in French, at least with inanimates (524 tokens for *vos nom et prénom* (‘your.PL name.SG et surname.SG’), 383 for *votre nom et prénom* (‘your.SG name.SG et surname.SG’)). When there is a mismatch of number, the determiner agrees with the closest conjunct.

- (16) a. La plupart de nos établissements sont
 the most of our facilities be.PL
 ouverts tous les jours y compris
 open.PL all the.PL days including
 le dimanche et jours fériés
 the.SG Sunday.SG and day.PL holiday.PL
 ‘Most of our facilities are open every day including Sunday and public holidays.’ (frWac)
 b. Chacun essaye de trouver sa place
 everyone tries to find his place.SG
 en fonction de ses dons et
 in accordance of his.PL gift.PL and

charisme.
 charisma.SG
 ‘Everyone tries to find his place according
 to his gifts and charisma.’ (frWac)

We thus assume that Det agreement for split binominals does not involve an abstract structure but a fairly superficial structure (Borsley 2009), where the Det number is influenced by the adjacent conjunct.

We assume that Dsing is disfavored for split singular animates since Dsing favors joint reading, and most animate nouns are ambiguous between a referential use (*Un collègue est venu*. ‘A colleague has come’) and a predicative use (*Jean est un collègue*. ‘Jean is a colleague’)

4 HPSG analysis

4.1 Previous work

As is generally accepted in HPSG, we use two distinct agreement features, CONCORD for morphosyntactic agreement and INDEX for semantic agreement (Pollard & Sag 1994, Wechsler & Zlatić 2000). Nouns, determiners, and attributive adjectives carry a CONCORD feature, closely related to inflection. INDEX agreement is more semantic, whose value is related to the referential/semantic possibilities of the associated nominal.

King & Dalrymple (2004) propose an LFG analysis where CONCORD features are distributive. The conjuncts require the Det to have as the same CONCORD value as the conjuncts. INDEX features are non-distributive, representing the set formed by the coordinating structure and triggering verb agreement.

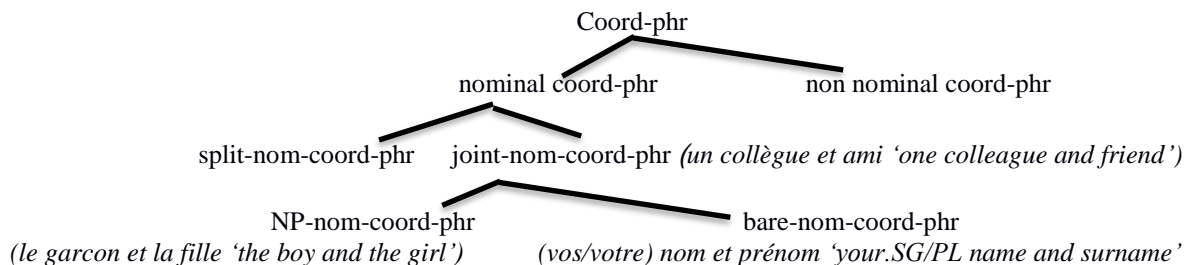


Figure 3. Hierarchy of French coordinated phrases in HPSG

We follow Borsley (2005) who argues that coordinated phrases are analysed as unheaded. Disregarding conjunction features, SLASH features are shared between the conjuncts and the coordinate phrase (Mouret 2007) and VALENCE features are shared by default (\setminus) (17).

Arnold et al. (2006) show that in Portuguese, the determiner always agrees in gender with the first conjunct, and in number either with the first conjunct (16a) or with the coordinated structure (16b).

- (16)a. No povo
 on.the.M.SG population.M.SG
 e gente hebreia
 and people.F.SG Hebrew.F.SG
 (Arnold et al. 2006)
 ‘on the Hebrew population and people’
- b. Os provaveis diretor
 the.MPL probable.PL director.MSG
 e ator principal
 and actor.MSG principal.MSG
 ‘the likely director and main actor’

In addition to CONCORD and INDEX, Arnold et al. (2006) propose two new features: L-AGR for the leftmost conjunct, R-AGR for the rightmost conjunct. In closest conjunct agreement, Det agrees with the first N via L-ARG. L-ARG and R-ARG are head features; the value of L-ARG of the coordination structure comes from the L-AGR of the leftmost daughter, while the CONCORD value reflects the resolved agreement features of the coordinate structure, with identical values of INDEX.

4.2 Our analysis

To deal with the agreement of binominals, we propose a hierarchy of nominal-coordinate-phrases (Fig. 3)

$$(17) \text{ Coord-phr} \Rightarrow \left[\begin{array}{l} \text{VAL} \setminus [2] \\ \text{SLASH} [3] \\ \text{DTRS} \left\langle \left[\begin{array}{l} \text{VAL} \setminus [2] \\ \text{SLASH} [3] \end{array} \right], \dots, \left[\begin{array}{l} \text{VAL} \setminus [2] \\ \text{SLASH} [3] \end{array} \right] \right\rangle \end{array} \right]$$

In this paper, we only deal with binominals whose head is a noun.

(18) nom-coord-phr => [HEAD noun]

For joint nominal coordinated phrases, CONCORD and INDEX features are also shared (19):

(19) joint-nom-coord-phr =>

$$\left[\begin{array}{l} \text{HEAD } [\text{CONCORD } [1]] \\ \text{INDEX } i \\ \text{DTRS list } \left(\left[\begin{array}{l} \text{CONCORD } [1] \\ \text{INDEX } i \end{array} \right] \right) \end{array} \right]$$

For split nominal coordination phrases, the coordinated phrase has a different INDEX value than the conjuncts and a plural CONCORD value (20).

(20) split-nom-coord-phr =>

$$\left[\begin{array}{l} \text{HEAD } [\text{CONCORD } [\text{NUM } pl]] \\ \text{INDEX } k = i + \dots n \\ \text{DTRS } < [\text{INDEX } i], \dots [\text{INDEX } n] > \end{array} \right]$$

For NP coordination (*le frère et la soeur* ‘the brother and the sister’), the VALENCE of the daughters is saturated and the number is plural. We then consider bare binominals, which can only have a split reading (Roodenburg 2005).

(21) a. Nom et prénom doivent être
 name.SG and surname.SG must be
 écrits en noir.
 written.PL in black
 ‘Name and surname must be written in
 black.’

(22) bare-nom coord-phr=>

$$\left[\begin{array}{l} \text{HEAD } \left[\begin{array}{l} \text{CONCORD } [\text{NUM } [0]] \\ \text{LAGR } [\text{NUM } [1]] \\ \text{RAGR } [\text{NUM } [2]] \end{array} \right] \\ \text{VAL } [\text{SPR } < (D[\text{NUM } [0] \vee [1]]) >] \\ \text{DTRS } \left\langle \left[\begin{array}{l} \text{HEAD } \left[\begin{array}{l} \text{CONCORD } \text{NUM } [1] \\ \text{LAGR } \text{NUM } [1] \end{array} \right] \right], \dots, \left[\begin{array}{l} \text{HEAD } \left[\begin{array}{l} \text{CONCORD } \text{NUM } [2] \\ \text{LAGR } \text{NUM } [2] \end{array} \right] \right] \right\rangle \\ \text{VAL } [\text{SPR } < D[\text{NUM } [1]] >] \end{array} \right]$$

b. Votre/Vos nom et prénom doivent
 your.SG/PL name.SG and surname.SG must
 être écrits en noir.
 be written.PL in black
 ‘Your name and surname must be written in
 black.’

For them, we assume the SPR value to be optional. We propose that a split bare coordinated phrase does not necessarily share its CONCORD feature with the conjuncts. It can either inherit the CONCORD value of the first conjunct, expecting a singular Det if it is singular, a plural Det if it is plural, or have a resolved CONCORD value, expecting a plural Det (22).

5 Conclusion

On the basis of large corpus data, we argue that singular split binominals do exist in French, and both singular and plural determiners are possible. Our experimental data further show that animacy play a role in the acceptability judgments: inanimate binominals are better accepted than animate binominals. The fact that the determiner can agree with the coordinated phrase suggest that the determiner is placed above the coordinated nouns.

We also propose an HPSG analysis with different subtypes for split reading and joint reading. For split reading, following Arnold et al (2006), we use L-ARG feature to capture the closest conjunct agreement fact.

This paper presents on-going work, which leaves a number of questions open. We have identified the role of animacy for the determiner agreement, whether there are other factors should also be explored.

References:

- Arnold, D, L Sadler, A Villavicencio. 2005. *An HPSG account of closest conjunct agreement in NP coordination in Portuguese*. In 12th International Conference on Head-driven Phrase Structure Grammar, ed. Stefan Müller, 427–447. Stanford, CA: CSLI Publications.
- Baroni, M., Bernardini, S., Ferraresi, A. et Zanchetta, E. (2009). *The WaCky wide web: a collection of very large linguistically processed web-crawled corpora*. *Language Resources and Evaluation*, 43(3):209–226
- Borsley, R. 2005. Against ConjP, *Lingua* 115, 461-482
- Borsley, R. 2009. *On the superficiality of Welsh agreement*. *NLLT*, 27: 225–265.
- de Swart, H, Le Bruyn, B. 2014. *Bare coordination: the semantic shift*. *NLLT* 32 (4): 205-1246
- Demonte, V, I Pérez-Jiménez. 2012. *Closest conjunct agreement in Spanish DPs. Syntax and beyond*. *Folia Linguistica* 46(1): 21–74.
- Haiman, John. 1983. *Iconic and economic motivation*. *Language* 59(4): 781–819.
- Heycock, C, R Zamparelli. 2005. *Friends and colleagues. Plurality, coordination and the structure of DP*. *Natural Language Semantics* 13: 201–270.
- King, T, M Dalrymple. 2004. *Determiner agreement and noun conjunction*. *J of Linguistics* 40 (1): 69–104.
- 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. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press
- Roodenburg, J, 2005, Une coordination particulière : les syntagmes N Conj N en français. *Langages* 160 : 93-109.
- Wälchli, Bernard. 2005. *Co-compounds and natural coordination*. Oxford: Oxford University Press
- Wechsler, S, Zlatić L. 2000. *A theory of agreement and its application to Serbo-Croatian*. *Language* 76:799–832.