ZEITSCHRIFT FÜR SPRACHWISSENSCHAFT
Journal of the Linguistic Society of Germany

Volume 45, Issue 3, Article 3

https://doi.org/10.18148/zs/2026-2012

Emma Kious (corresponding author), Laboratoire de Linguistique Formelle, Université Paris Cité, France, emma.kious@u-paris.fr

Anne Abeillé, Laboratoire de Linguistique Formelle, Université Paris Cité, France, anne.abeille@u-paris.fr

Gender mismatch and ellipsis: comparing French adjectives and nouns

Abstract: According to Merchant (2014), Greek adjectives and human nouns behave differently under ellipsis: while adjectives allow for gender mismatch, human nouns may fall into three classes: kinship and nobility nouns (*adherfos/aderfi* 'brother/sister') which allow no mismatch, role nouns (*dhaskalos/dhaskala* 'teacher.M/F') which allow asymmetric mismatch, and syncretic nouns (*jatros* 'doctor.M/F') which allow symmetric mismatch. We compare adjectives and nouns in French. We conducted two quantitative studies on several French corpora (FrTenTen, ORFEO: spoken-part, Frantext: after-1980, DUEL), we found more cases of adjective ellipsis (N=562) than of noun ellipsis (81 cases). For predicative adjectives (with human subjects), we also found a higher rate of gender mismatch: 40.14% (224/558 in the written corpora) than for predicative nouns (17.3%, N=14 in the written corpora). The same factors (with human pronominal subjects favoring mismatch) can be observed for adjectival and nominal predicates. We also show that French human nouns allow for more gender mismatches than previously thought, even for kinship nouns (*fils/fille* 'son/daughter') and unrelated pairs (*homme/femme* 'man/woman'). We conducted three experiments to test the acceptability of gender mismatches under ellipsis, compared with anaphora, for adjectives and role nouns. We found no gender mismatch penalty under ellipsis, no robust order effect, and a role of gender biases. This suggests that adjectives and (gendered) role nouns behave similarly under ellipsis, allowing for symmetric gender mismatch under ellipsis in French.

Keywords: ellipsis, mismatch, French, corpus data, experiment

Cite as: Kious, Emma & Anne Abeillé. 2026. Gender mismatch and ellipsis: comparing French adjectives and nouns. *Zeitschrift für Sprachwissenschaft* 45(3). https://doi.org/10.18148/zs/2026-2012.
Received: 29 September 2025; **accepted:** 27 August 2026; **published:** 16 November 2026
Published by: publikon / dgfs (Linguistic Society of Germany) under a Creative Commons 4.0 BY license

1. Introduction

Significatio ex nihilo (Merchant 1999), ellipsis illustrates a discrepancy between form and meaning, where the interpretation of the elliptical clause (also called "target" clause) depends on the content of a previous clause (the antecedent) or the situation (Poppels 2022). It omits forms while preserving the clausal meaning of the utterance and makes it possible to construct an autonomous and acceptable utterance.

Stripping ellipsis, or Bare Argument ellipsis (Culicover & Jackendoff 2005), is a head ellipsis construction with a unique nominal remnant (*Paul*) and a polarity adverb (*too*), which follows a complete antecedent clause with a correlate (*Mary* in 1a).

(1) a. *Mary is happy and Paul too.*
b. *Mary is happy and* [[*Paul too*]$_{\text{FocP}}$ - ~~is happy~~]$_{\text{CP}}$.
c. *Mary is happy and* [[*Paul*]$_{\text{NP}}$ *too*]$_{\text{S}}$.

While syntactic derivation theories proposed by Ross et al. (1969) or Merchant (2001) posit movement and identity under deletion of the second clause (1b), Abeillé & Kim (2022) adopt a referential, or direct interpretation approach (Ginzburg & Sag 2000) (1c). The direct interpretation approach posits no syntactic reconstruction: the elliptical predicate (*be happy* in 1a) is inherited from the antecedent clause (or the context) with the remnant's content as argument (Nykiel & Kim 2021).

1.1. Mismatch and ellipsis

Mismatches in grammatical form between antecedent and elliptical clauses are usually possible and found cross-linguistically. In English as in French gapping, inflectional feature mismatches are attested (e.g., number) (2a), but interpreted features, such as tense, are not (2b).

(2) a. *Paul va à Paris et ses enfants* [*vont*] *à Rome.*
Paul go.PRS.3SG to Paris and his.PL children [go.PRS.3PL] to Rome
'Paul is going to Paris and his children to Rome.'
(Abeillé et al. 2014: 240)

b. #*Paul est arrivé hier et Marie* [*arrivera*] *demain.*
Paul be.PRS.3SG arrive.PTCP.M.SG yesterday and Marie [arrive.FUT.3SG] tomorrow
'Paul arrived yesterday and Marie tomorrow.'
(Abeillé et al. 2014: 240)

A corpus study by Abeillé & Kim (2022) found similar mismatches in English and French stripping (3). They found that for French stripping, for instance, 70% of *aussi* ('too') and 63% of *non plus* ('neither') fragments feature mismatches in person, number, or gender.

(3) a. *I'm beginning to like the guy. You too* [*are beginning to like the guy*] ? [COCA, 1993]
b. *Elle était ravie et moi aussi* [*j'-étais ravi*].
she be.PST.3SG thrilled.F.SG and me too I-be.PST.1SG thrilled.M.SG]
'She was thrilled and me too.'
[Frantext, Dorin, 1084]

Mismatches between antecedent and fragment challenge syntactic reconstruction approaches (with identity-under-deletion): structuralist theories of ellipsis posit an identity relation between the ellipsis and its antecedent. This notion of identity is built on connectivity effects which, according to Culicover & Jackendoff (2005), give empirical support in favour of some invisible structure at the ellipsis site. Yet, the exact nature of this identity relation has been, and remains, the topic of numerous debates: from Fiengo & May's (1994) syntactic identity theory of 'vehicle change' (the ellipsis site and its antecedent allow for a mismatch if both elements are referentially equivalent), to Merchant's (2001) semantic identity condition with the notion of e-givenness (based on the givenness of elliptical clauses), to hybrid identity theories such as Chung et al.'s (2006) no-new-words constraint or Thoms' (2015) inferential antecedent accommodation proposal or Griffith's (2019) non-isomorphic reconstruction.

Unlike identity-based theories, referential theories rely on contextually accessible information. Building upon Ginzburg & Sag (2000), referential theories consider ellipsis to involve a discourse antecedent with a uniquely identifiable thematic referent, integrated into the Common Ground. Ellipsis is then a discursive reference, akin to anaphora (Hankamer & Sag 1976). Such an approach has received empirical support from exophoric VP ellipsis (Miller & Pullum 2014), exophoric stripping ellipsis (Abeillé & Kim 2022), and case differences between elliptical and non-elliptical clauses: *Who is coming? – Me/*I. – I/*Me am coming.*

1.2. Previous work on gender mismatch under stripping

There has been some experimental work on gender mismatch and stripping ellipsis, but no dedicated corpus work, to the best of our knowledge. In a written corpus study of English and French stripping (with pronominal remnants), Abeillé & Kim (2022) found different cases of mismatch, including one gender mismatch in French, for an adjectival predicate (4).

(4) *Cette fois je suis vraiment amoureuse, lui aussi* [*est vraiment amoureux*].
this.F time I be.PRS.1SG really in.love.F.SG him too [be.PRS.3S really in.love.M.SG]
'This time, I am really in love, him too.'
[Frantext, Prin 2005]

In Spanish, Aparicio et al. (2015) compared mismatch penalties for predicative adjectives under ellipsis ('surface anaphora', per Hankamer & Sag 1976) (5a) and under ('deep') anaphora, with the Spanish clitic *lo* (close to the French expletive pronoun *le*, neither varying in gender) (5b). Deep anaphora, which involves pragmatic control of a salient discursive antecedent, should be more permissive than ellipsis, or surface anaphora, which involves syntactic control.

(5) a. *El alumno es alto y la profesora también*
the.M.SG student.M.SG be.PRS.3SG tall.M.SG and the.F.SG teacher.F.SG too
[*es alta*].
[be.PRS.3SG tall.F.SG]
'The student is tall and the teacher is too.'
(Aparicio et al. 2015: 2)

b. *El alumno es alto y la profesora también lo*
the.M.SG student.M.SG be.PRS.3SG tall.M.SG and the.F.SG teacher.F.SG too it
es.
be.PRS.3SG
'The student is tall and the teacher is that too.'
(Aparicio et al. 2015: 5)

In two eye-tracking experiments on Spanish, Aparicio et al. (2015) found that mismatching features on the predicative adjectives carried a higher processing cost under ellipsis, suggesting that these features are not ignored for ellipsis resolution. Their results also pointed towards a stronger mismatch penalty for feminine antecedents (Feminine-Masculine ordered conjuncts). But they did not distinguish number and gender features.

Looking at modern Greek, Merchant (2014) suggested that gender mismatches under stripping ellipsis are allowed for adjectives (6a), but not necessarily for nouns (6b,c). He distinguishes three classes of human nouns: nobility and kinship nouns which prohibit gender mismatches (6b), non-alternating role nouns which allow for them (6c), and alternating role nouns which only allow asymmetric mismatches (6d) in Masculine-Feminine order.

(6) a. *I Maria ine ikani, ala O Alexandros dhen ine*
the.F.SG Maria be.PRS.3SG capable.F but the.M.SG Alexandros not be.PRS.3SG
[*ikanos*].
[capable.F]
'Maria is capable, but Alexandros is not.'
(Merchant 2014: 2)

b. **O Petros ine adherfos, ke I Maria episis* [*ine adherfi*].
the.M.SG Petros be.PRS.3SG brother and the.F.SG Maria too [be.PRS.3SG sister]
* 'Petros is a brother and Maria too.'

*I Maria ine adherfi, ke I Petros episis [ine adherfos].
the.F.SG Maria be.PRS.3SG sister and the.M.SG Petros too [be.PRS.3SG brother]
* 'Maria is a sister and Petros too.'
(Merchant 2014: 4)

c. O Petros ine kalos jatros, ala I Maria ine
the.M.SG Petros be.PRS.3SG good-M-SG doctor-SG but the.F.SG Maria be.PRS.3SG
mia kakia [jatros].
a.F.SG bad.F.SG [doctor.SG]
'Petros is a good doctor, but Maria is a bad one.'

I Maria ine kali jatros, ala O Petros ine
the.F.SG Maria be.PRS.3SG good.F.SG doctor.SG but the.M.SG Petros be.PRS.3SG
enas kakos [jatros].
a.M.SG bad.M.SG [doctor.SG]
'Maria is a good doctor, but Petros is a bad one.'
(Merchant 2014: 7)

d. O Petros ine kalos dhaskalos, ala I Maria ine
the.M.SG Petros be.PRS.3SG good.M.SG teacher.M.SG but the.F.SG Maria be.PRS.3SG
mia kakia [dhaskala].
a.F.SG bad.F.SG [teacher.F.SG]
'Petros is a good teacher, but Maria is a bad one.'

*I Maria ine kali dhaskala, ala O Petros ine
the.F.SG Maria be.PRS.3SG good.F.SG teacher.F.SG but the.M.SG Petros be.PRS.3SG
enas kakos [dhaskalos].
a.M.SG bad.M.SG [teacher.M.SG]
* 'Maria is a good teacher, but Petros is a bad one.'
(Merchant 2014: 8)

Sudo & Spathas (2020) refine Merchant's categories for Greek nouns, taking a gender competition approach where the masculine gender does not trigger any social gender inferences. The asymmetry in (6d) is then due to the feminine form implying the masculine form, while the reverse does not hold.

Looking at nominal ellipsis in Brazilian Portuguese, Nunes & Zocca (2009) also proposed a similar three-way distinction between human nouns for Brazilian Portuguese: those which license gender mismatches under ellipsis in both directions (7c, d), those (7e, f) whose masculine form sometimes allows a feminine subject (with the reverse being impossible), and those (7a, b) which allow no gender mismatches.

(7) a. **O Drácula* é *conde* *e* *a* *Mina* *também* é
the-m-sg Dracula be.PRS.3SG count.M.SG and the.F.SG Mina also be.PRS.3SG
[*condessa*].
[count.F.SG]
* 'Dracula is a count and Mina is too.'

b. **A Mina é condessa e o Drácula também* é
the.F.SG Mina be.PRS.3SG count.F.SG and the.M.SG Dracula also be.PRS.3SG
[*conde*].
[count.M.SG]
* 'Mina is a countess and Dracula is too.'

c. *O Paulo é médico e a Fernanda também é*
the.M.SG Paulo be.PRS.3SG doctor.M.SG and the.F.SG Fernanda also be.PRS.3SG
[*médica*].
[doctor.F.SG]
'Paulo is a doctor and Fernanda is too.'

d. *A Fernanda é médica e o Paulo também é*
the.F.SG Fernanda be.PRS.3SG doctor.F.SG and the.M.SG Paulo also be.PRS.3SG
[*médico*].
[doctor.M.SG]
'Fernanda is a doctor and Paulo is too.'

e. ?*O Paulo é ator e a Fernanda também é*
the.M.SG Paulo be.PRS.3SG actor.M.SG and the.F.SG Fernanda also be.PRS.3SG
[*atriz*].
[actor.F.SG]
? 'Paulo is an actor and Fernanda is too.

f. ??*A Fernanda é atriz e o Paulo também é*
the.F.SG Fernanda be.PRS.3SG actor.F.SG and the.M.SG Paulo also be.PRS.3SG
[*ator*].
[actor.M.SG]
?? 'Fernanda is an actress and Paulo is too.'
(Nunes & Zocca 2009: 216)

For Bobaljik & Zocca (2011), nobility (e.g. **conde*/**condesa*) and kinship nouns (e.g. *#tio*/*#tia*) always carry gender presupposition and allow for no mismatches. Alternating role nouns (e.g. ?*ator*/*#atriz*) show an asymmetric mismatch. For the latter, this asymmetry suggests a morpho-syntactic distinction between a marked [feminine] and an unmarked [masculine], where only [feminine] triggers a female referent presupposition. Other alternating nouns, in Brazilian Por-

tuguese, may also fall into the third class of nouns (which allows for symmetric mismatch), as it includes not only common gender nouns (as in Greek) but also role nouns (*médico/a*) that inflect for gender like adjectives, and do not carry gender presupposition.

Building upon Bobaljik & Zocca's (2011) and Merchant's (2014) proposals, Sprouse et al. (2022) empirically tested English nobility nouns (*prince/princess*) (8a, b), and alternating role nouns (*actor/actress*) (8c, d) under ellipsis.

(8) a. **John is a prince and Mary is* [*a prince*] *too.*
b. **Mary is a princess and John is* [*a princess*] *too.*
c. *John is an actor and Mary is* [*an actor*] *too.*
d. **Mary is an actress and John is* [*an actress*] *too.*
(Sprouse et al. 2022: 6)

A first acceptability study tested 16 pairs of nouns, with 8 belonging to a possible kinship and nobility class as in (8a, b), and 8 to a possible actor/actress class as in (8c, d), following a 2×2×2 design (with ±ORDER ±MISMATCH and ±ELLIPSIS as factors), and found an interaction between +MISMATCH and +ELLIPSIS for 7 of the 16 pairs. A second study tested 58 gendered pairs as in (8c, d) (with a 2×2 design: ±ORDER, ±MISMATCH), and found mismatch asymmetries for 35 of them, with a high inter-participant variation. They did not make a comparison with non-elliptical or anaphora counterparts and did not take gender biases (associated with profession names) into account. They did not confirm Haspelmath's relative frequency hypothesis, according to which the higher frequency of masculine forms makes Masculine-First ordering more acceptable and more efficient for communication.[1]

The present study seeks to further explore the question of gender mismatches under stripping ellipsis for adjectives and nouns in French, using both corpora (spoken and written) and experiments.

[1] Gandón-Chapela & Gallardo-del-Puerto (2025) tested gender mismatch under VP ellipsis: participants had to choose between strict and sloppy readings. While the sloppy reading was always preferred, it was more so with gender match (i) than with gender mismatch, and slightly more so in the Feminine-Masculine order (ii) than in (iii). Their results show participant sensitivity to gender mismatches under ellipsis.

(i) *Peter has voted for himself and Robert has too.*
a. *Robert has voted for Peter.* (strict reading)
b. *Robert has voted for Robert.* (sloppy reading)

(ii) *Elaine has voted for herself and Peter has too.*
c. *Peter has voted for Elaine.* (strict reading)
d. *Peter has voted for Peter.* (sloppy reading)

(iii) *Peter has voted for himself, and Elaine has too.*
e. *Elaine has voted for Peter.* (strict reading)
f. *Elaine has voted for Elaine.* (sloppy reading)

1.3. Testing French gender

Comparing French adjectives and nouns, the former have inflectional gender while the second are usually considered as having inherent gender (Haspelmath 2002; Zwanenburg 1988). We expect more gender mismatch with adjectives than with nouns.

But adjectival and nominal categories in French are somewhat permeable (Noailly 2021) and human nouns usually come in gendered pairs, with the same paradigmatic structure as adjectives (Bonami & Boyé 2019). Although adjectival gender is inflectional, the status of nominal gender is disputed for human nouns: derivational (Zwanenburg 1988) or inflectional (Spencer 2002). As inflectional (agreement) features should be ignored for ellipsis resolution (Merchant 2010), whether nouns pattern with adjectives or not for gender mismatches under ellipsis will shed light on this debate.

In the following, we distinguish grammatical gender (which triggers agreement) as a formal feature classifying nouns, from semantic gender as a conceptual classification according to a noun's lexical meaning, and social gender as a way one presents oneself in society, hinging upon world knowledge. For human nouns, grammatical gender encodes different meanings, implications, and connotations (Richy & Burnett 2019). For languages with grammatical gender, Corbett (1991)'s Semantic Core Generalisation states that for human nouns, social, and grammatical genders tend to correspond. Alternating noun pairs fall under this generalisation (9a): masculine grammatical gender corresponds to male social gender, and feminine grammatical gender to female social gender. Epicene nouns (9b) and mixed groups (9c), on the other hand, may not.

(9) a. *C'* *est* *une* *danseuse.*
that.SG be.PRS.3SG a.F.SG dancer.F.SG
'She is a dancer.'
C' *est* *un* *danseur.*
that.SG be.PRS.3SG a.M.SG dancer.M.SG
'He is a dancer.'

b. *Ces* *personnes* *sont* *excellentes.*
these.PL people.F.PL be.PRS.3PL excellent.M.PL
'Those people are excellent.'

c. *Ces* étudiants *sont* *impressionnants.*
these.PL student.M.PL be.PRS.3PL impressive.M.PL
'Those students are impressive.'

Role nouns also come with social biases, and stereotypical gender information may become part of the word's meaning (Mickus et al., 2019). A woman dancer (*danseuse*) is more stereotypical

than a man one (*danseur*), while a man banker (*banquier* 'banker.M) is more stereotypical than a woman one (*banquière* 'banker.F') (Misersky et al. 2014).

Based on corpus studies and experiments, we investigate which factors are at play for the production and comprehension of gender mismatches under stripping ellipsis. We will directly compare predicative adjectives and nouns in French, with the hypothesis that gender mismatch under ellipsis is possible for adjectival predicates but more constrained for nominal predicates. If French resembles (Brazilian) Portuguese more than (modern) Greek, another hypothesis is that symmetric mismatch may be possible beyond syncretic nouns. We will compare production data (from both written and spoken corpora) and comprehension data (from experiments).

2. Corpus study of French ellipsis

2.1. Methodology

We used different French corpora, of varying registers and modality: Frantext Contemporain (French literary texts after 1980, ATILF), FrTenTen (a French web-based corpus from the TenTen family, Kilgarriff et al. 2004), ORFEO (the spoken part, Benzitoun et al. 2016) and DUEL (a multimodal corpus of recorded interactions, Hough et al. 2016). As they are not annotated for ellipsis, extraction from the corpora was semi-automatic.[2]

We found 643 adjectival and 81 nominal ellipsis occurrences across the corpora, most of them in the written corpora (only five stripping utterances are found across spoken corpora).

We annotated occurrences for a variety of semantic, phonological and syntactic factors and, when such information was available, for the gender of the author or speaker (see Table 1). Beyond annotating the presence of mismatches (in gender and number), as well as the conjuncts order (Masc-Masc, Fem-Fem, Masc-Fem, Fem-Masc), the targets were annotated for morpho-phonology. For adjectival predicates, we distinguished homophones (e.g. *ému/e* 'moved.M/F') and syncretic pairs (e.g. *timide* 'shy.M/F') from alternating pairs. For nominal predicates, we contrasted common gender (such as *enfant* 'child.M/F') and epicene nouns from alternating pairs. We did this with the hypothesis that syncretism would favour mismatches. We also considered the relative frequency of the masculine form and the feminine form, (*Freq_Masc-Freq_Fem*), as well as the difference between the two forms and the lemma frequency. We used log-transformed frequencies from the corpus of French subtitles Lexique (New et al. 2001). We also annotated the subject of the full

[2] For Frantext, ORFEO and DUEL, we looked for all occurrences of *aussi/non plus* ('too/neither'), manually sorting through some 80k occurrences and keeping only stripping utterances (manual selection allowed us to account for disfluencies and large contexts). For FrTenTen, we used a CQL (Corpus Query Language) search for occurrences of *aussi/non plus* ('too/neither') in the vicinity of verbs such as être/sembler/para*itre* ('be/seem/appear') using 2–5-word jokers, and kept stripping constructions.

(antecedent) clause and the subject of the elided (target) clause for their category, nominal or pronominal, and pronominal subjects were annotated for their gender marking (in French, 1st and 2nd person do not inflect in gender), and humanness. For predicative nouns, we checked whether the target was a bare noun (10a) or had a determiner (which agrees in gender, 10b).

(10) a. *Mon grand-père était pasteur. Le vôtre aussi* [était
my.SG grandfather be.PST.3SG pastor.M.SG the.M.SG yours.SG too [be.PST.3SG
pasteur].
pastor.M.SG]
'My grandfather was a pastor. Yours too.'
[Frantext, de Monferrand 1990]

b. *Je suis le chef et le dernier de mon clan, toi aussi*
I be.PRS.1SG the.SG chief.M.SG and the.SG last.M.SG of my.SG clan you too
[*tu es le chef et le dernier de ton clan*].
[you be.PRS.2SG the.SG chief.M.SG and the.SG last.M.SG of your.SG clan]
'I am the chief and last of my clan, you too.'
[Frantext, Bauchau 1990]

Table 1: Summary of the main annotation factors

Annotation target	Annotation factor	Factor levels
Correlate, Remnant	Category	Nominal/Pronominal
	Gender markedness	Marked (3rd person/Unmarked (1st, 2nd person)
	Humanness	Human/Non-human
Constituent under ellipsis	Lexical frequency	Masc. /Fem./(Masc_freq-Fetn freq)/Lemma
	± Gender mismatch	Match/Mismatch
	± Number mismatch	Match/Mistmatch
	± Bare noun	Determiner/Bare
	Order	Masc-Masc/Fem-Fem/Masc-Fem/Fem-Masc
Occurrence	Author gender	Female/Male/Unknown

2.2. Hypotheses

We expect a lower mismatch rate for nominal predicates than for adjectives: gender is indisputably an inflectional feature for adjectives, but this is more controversial for French noun pairs. Overall, with human subjects, we expect more Masculine-First order, regardless of mismatch, due to a general ordering tendency for humans (Abeillé et al. 2018; Da Cunha 2024).

In mismatch cases, there are several hypotheses. For adjectives, we expect no interaction with ordering (as for other inflectional features). For human nouns, Merchant (2014) predicts more

Masculine-First, if the masculine noun is compatible with a female subject, but not vice versa. But referential theories would predict more Feminine-First if the (elided) masculine noun is more accessible than the feminine one at the ellipsis site. We also expect an interaction with lexical frequency, if masculine forms are more frequent than their feminine counterparts (Haspelmath 2020).

Given the specificities of French, we hypothesise that syncretic adjectives and common gender nouns favour gender mismatches under ellipsis, followed by homophone pairs: in such cases, mismatches would be less salient and therefore easier to produce. According to Noailly (2021), 18% of French adjectives are syncretic for gender (e.g. *timide* 'shy.M/F') and 66% are homophones (e.g. *joli/e* 'pretty.M/F') in standard French, leading us to posit that unmarked inflection—orthographically or phonologically—would favour mismatches, in line with identity theories of ellipsis. According to Bonami & Boyé (2019), half of human nouns are syncretic or common gender in (spoken) French.

We also expect that pronominal subjects unmarked for gender (1st and 2nd person) will favour mismatches, given that gender would only be marked by agreement. In contrast and in line with Corbett's (2023) Generalised Agreement Hierarchy, we expect gendered human subjects to disfavour mismatches, as they carry inferrable social gender.

2.3. Corpus results

2.3.1. Adjectival predicates under ellipsis

For predicative adjectives, we observe an overall high rate of gender mismatches[3]: 40.03% (N=225), and a higher count of Masculine-First occurrences (N=349, 62.1%) (13). Although the Feminine-First order is less frequent (N=213, 37.9%), the mismatch rate for Feminine-First is higher (55.4%, N=118/213) (14a), than with Masculine antecedents (30.7%, N=107/349). This order effect is statistically robust: Feminine-First is a robust predictor of gender mismatches. We report these results in Figure 1 and detail Bayesian inference models in Appendix A.[4] These results may be explained by the accessibility hypothesis if masculine referents are more frequent overall, they are more accessible, and thus more easily retrieved in the case of mismatches.

We found no robust effect of adjectival form: non-syncretic adjectives, regardless of the morphophonological distance (11, 13a–c) do not seem to disfavour gender mismatches. The same holds

[3] Number mismatch rates are much lower, with 19.2% (N=108) and, for occurrences with a number mismatch, the gender mismatch rate is very similar to that of the entire sample (41.6%, N=45), which would indicate no superadditive mismatch penalty for occurrences such as (11) below.

[4] We consider there to be robust and compelling evidence for an effect of a predictor if the probability of a posterior distribution estimate being different from 0 is superior to 95%.

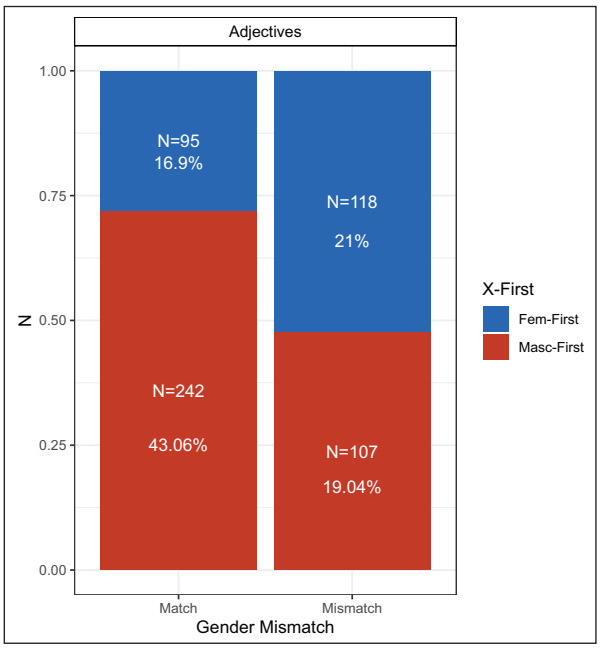


Figure 1: Order preferences for adjectival corpus data showing counts and proportions

for syncretic adjectives (12a) and homophones (12b), which do not affect gender mismatch rates: although theory predicts an advantage for such cases, the lack of syncretism effects may simply be due to the fact that adjectives in general, where gender is inflexional, do not penalise gender mismatches, regardless of more fine-grained morphophonological properties.

(11) *Astruc est fou de bonheur, elles aussi* [*sont folles de bonheur*].
Astruc be.PRS.3SG crazy.M.SG of joy, they.F.PL too [be.PRS.3PL crazy.F.PL of joy]
'Astruc is overjoyed, them too.'
(Frantext, Gréco 1982)

(12) a. *Mon père était athée, moi aussi* [*j'étais athée*].
my.SG father be.PST.3SG atheist.M.SG me too [be.PST.1SG atheist.F.SG]
'My father was an atheist, me too.'
(Frantext, Montferrand 1990)

b. *Il était* ému, *moi aussi* [*j' étais* émue].
he be.PST.3SG moved.M.SG me too [I be.PST.1SG moved.F.SG]
'He was moved, me too.'
(Frantext, Chaix 1985)

(13) a. *Elle était vieille, et lui aussi [était vieux].*
she be.PST-3SG old.F.SG and him too [be.PST.3SG old.M.SG]
'She was old, and him too.'
[Frantext, Bianciotti, 1985]

b. *Je suis devenue folle et Nick aussi [est devenu fou].*
I be.PRS.1SG become.F.SG crazy.F.SG and Nick too [be.PRS.3SG become.M.SG crazy.M.SG]
'I became crazy and Nick too.'
[Frantext, Delperdange, 1992]

c. *La photo est belle, le garçon aussi [est beau].*
the.F.SG picture.F.SG be.PRS.3SG beautiful.F.SG the.M.SG boy too [be.PRS.3SG beautiful.M.SG]
'The picture is beautiful, the boy too.'
[Frantext, Barthes, 1980]

Looking at subject category and gender markings, we found that pronominal correlates robustly favour gender mismatches, and observe a mismatch rate of 51.25% (N=103). In our corpus, while only a minority of pronominal subjects are marked for gender (13a), we observe a slightly higher gender mismatch rate than the baseline for predicative adjectives. Namely, we found a 60.5% mismatch rate (N=49) for pronominal correlates, and a 57.1% mismatch rate (N=20) for pronominal remnants (11), contra our hypothesis. Although this factor is not robust, we discuss this result in section 2.4.

We looked at subject humanness, with the hypothesis that the uninterpretability of grammatical gender for inanimates would favour mismatches. We found no main effect of subject humanness (with rates similar to baseline for correlates (44.9%, N=169) and remnants (42.4%, N=173) alike), but inferential modelling shows a robust interaction between order and subject humanness: in the Feminine-First order, human remnants (13a, b) favour gender mismatches for predicative adjectives. Despite the strong correlation between grammatical and social gender for humans (Corbett 1991), humanness does not penalise mismatches.

To further investigate interactions with social gender, we looked at a possible role of author gender in their production of mismatches, when such information was available. Most of the occurrences were produced by male authors (N=231, or 57%), but we observe a tendentially higher mismatch rate for female authors (48.3%, N=84) than for male (36.8%, N=85), although this is not a robust effect. In cases of mismatch, we found no effect of author gender on ordering preferences: Feminine-First is preferred for male (54.11%, N=46) and female (53.57%, N=45) authors alike.

2.3.2. Nominal predicates under ellipsis

In our corpora, we found fewer nominal predicates under ellipsis (N=81) than predicative adjectives (N=562), with 15 occurrences of gender mismatches. With a rate of 18.5%, gender mismatches are rarer with predicate nouns than with adjectives, but are still produced (14a, b). Given the size of our dataset, we were not able to reach robustness thresholds for reliable predictors, and instead present tendencies and observations.

(14) a. *Je me sens un habitant du quartier […] Moi aussi* [*je me sens une habitante du quartier*].
I me feel.PRS.1SG a.M.SG inhabitant.M.SG from neighbourhood me too [I me feel.PRS.1SG a.F.SG inhabitant.F.SG from neighbourhood]
'I feel like an inhabitant of the neighbourhood. Me too.'
[CFPP]

b. *Ses parents* étaient *musiciens […] Vous aussi* [*étiez musicienne*]*?*
her.PL parents be.PST.3PL musician.M.PL you.SG too [be.PST.2PL musician.F.SG]
'Her parents were musicians. You too?'
[Frantext, Garat 2006]

c. *Elle avait l' air d' une enfant prise en faute. Moi aussi* [*j' avais l' air d' un enfant pris en faute*]
she have.PST.3SG the.SG look.M.SG of a.F.SG child.F.SG taken.F.SG in fault me too I have.PST.1SG the.SG look.M.SG of a.F.SG child.F.SG taken.F.SG in fault
'She had the look of a child caught red-handed. Me too.'
[Frantext, Férey 2012]

d. *Babet n' était pas cuisinière… Marie non plus* [*n' était pas cuisinière*].
Babet NEG be.PST.3SG NEG cook.F.SG Marie neither NEG be.PST.3SG NEG cook.F.SG
'Babet was not a cook… Marie neither.'
[Frantext, Chandernagor 1995]

Previous literature on gender mismatches under ellipsis highlighted the specificity of role nouns with regards to ordering constraints: in Romance languages more specifically, for alternating role nouns, mismatches should incur a stronger penalty in the Feminine-Masculine order.

Our results, in line with Sprouse et al. (2022), seem to indicate that the Masculine-Feminine ordering may indeed be preferred, found in 62.96% of cases (N=51/81), but does not account for all cases, indicating that ordering is a preference rather than a constraint. The overall preference for Masculine-First (75.31%, N=61/81) (14a), regardless of mismatch, in line with previous observations on the ordering of conjuncts, may also play a role. When Feminine is in first position

(N=20/81) (14c), we observe a mismatch rate of 25% (N=5/20), while the mismatch rate for Masculine-First is slightly lower (16.39%, N=10/61): this seems to indicate that Feminine-First may favour gender mismatch for nominal predicates, although, again, these contrasts are not robust.

Since some nominal lemma frequencies were not available in Lexique (New et al. 2001), we analysed the lexical frequency of 57 occurrences (45 lemmas, out of the 67 total lemmas in the corpus). Overall, only one role noun (*cuisinier/ère* 'cook.M/F') was more frequent in its feminine form than its masculine form and occurred in a matching occurrence (14d). When the masculine form is the most frequent of the pair (N=19), we find four cases of mismatch, and no Feminine-First ordering.

For analysis purposes, we distinguished cases where gender iconic pairs of nouns are possible (e.g. *habitant/e* 'inhabitant.M/F') which make up 40% of our sample (N=33), from those where such pairs are not available (N=48). In the latter case, gender may be arbitrary (as with inanimates), pairs may not be possible, as for *gentilhomme* ('gentleman') or common gender nouns (e.g. *sociologue* 'sociologist.M/F').

As with predicative adjectives, we do not observe a preference for homophones gender mismatch production (mismatch rate of 16.6%, N=8, close to the overall sample rate). We note here some atypical cases, such as *gentilhomme* ('gentleman') (15c), for which there exists no feminine form: the noun is lexically specified for [masculine] and may only refer to males. We also note (15a), the *homme/femme* 'man/woman' suppletive pair, as well as the kinship nouns *fils/fille* 'son/daughter' (15b). Both of these examples feature surprising gender mismatches, notably despite the impossibility of mismatch for kinship nouns predicted by Merchant (2014).

(15) a. *Vous êtes un homme très occupé*[5]. *En dépit* des *apparences: moi*
you be.PRS.2PL a.M.SG man very busy.M.SG in spite of appearances me
aussi [*je suis une femme très occupée*].
too [I be.PRS.1SG a.F.SG woman very busy.F.SG]
'You are a very busy man. Despite appearances: me too.'
[Frantext, Garat 2008]

b. *Maria Campallo serait une fille de disparus. Votre copain*
Maria Campallo be.COND.3SG a.F.SG daughter of missing.M.PL your friend.M.SG
Luz/Orlando aussi [*serait un fils de disparus*].
Luz/Orlando too [be.COND.3SG a.M.SG son of missing.M.PL]
'Maria Campallo would be a daughter of missing persons. Your friend Luz/Orlando too.'
[Frantext, Férey 2012]

[5] A reviewer notes that another reconstruction would be possible in (15a) *je suis très occupée* ('I am very busy.F'), but this would not work for (15b) where a noun is needed. A common gender noun might be possible (*serait un enfant de disparus* 'would be a child of missing persons'). Both these examples show that the gender of the antecedent noun can be ignored under ellipsis.

c. *Je vois que vous êtes, monsieur, maintenant, gentilhomme de*
I see.PRS.1SG that you.PL be.PRS.2PL Sir now gentelman by
naissance […] *et moi aussi* [*je suis maintenant gentilhomme de*
birth and me too I be.PRS.1SG now gentleman by
naissance].
birth
'I see that you are, Sir, now, a gentleman by birth, and me too.'
[Frantext, Koltès 1988]

Determiners, in French, agree with the noun they specify: the presence of a determiner marks gender when the noun does not alternate, and makes it possible to disambiguate nominal and adjectival uses (16). In that sense, we expected bare nouns, closer to adjectives, to favour gender mismatches, but we found identical gender mismatch rates of 18.5% with and without determiner.

Given the lack of semantic motivation for grammatical gender of non-human referents, we expected non-human subjects to favour gender mismatches. Our corpus had very few non-human subjects (6 correlates and 7 remnants) (16b), and no gender mismatches with non-human subjects.

Overall, our corpus data show a predominance of (human) pronominal subjects. We observe a non-robust but tendentially higher-than-baseline mismatch rate for pronominal correlates (26.7%, N=8) and remnants (30.4%, N=7) alike. For unmarked pronominal subjects (16c)[6], this rate appears to (24%, N=6 for correlates, 28.6%, N=6 for remnants), which, if verified through inferential modelling on larger samples, would indicate that gender-marked pronominal (16a) subjects favour mismatches.

(16) a. *Il ne faut pas me prendre pour un con, ni elle non plus*
IMPRS not must NEG me take.INF for a.M.SG idiot.M.SG nor her either
[*il ne faut pas la prendre pour une conne*].
[IMPRS not must NEG her.REFL take.INF for a.F.SG idiot.F.SG]
'Don't take me for a fool, nor her either.'
[Frantext, Bayon 1987]

b. *Tous les livres de Queneau ne sont pas oulipiens, tous les*
all the.PL books of Queneau not be.PRS.3PL NEG oulipian.M.PL all the.PL
miens non plus [*ne sont pas oulipiens*].
mine either [not be.PRS.3PL NEG oulipian.M.PL]
'All of Queneau's books are not oulipian, all of mine neither.'
[Frantext, Perec, 2004]

[6] Note that in cases where the remnant is unmarked for gender, the referent's gender was recovered from the context in the annotation phase.

c. *T' es le fils d' un voleur et moi aussi [je suis*
you be.PRS.3PL the.M.SG son of a.M.SG thief.M.SG and me too I be.PRS.3PL
le fils d' un voleur].
the.M.SG son of a.M.SG thief.M.SG
'You are the son of a thief. Me too.'
[Frantext, Portante 1993]

Where such information was available, we looked at author gender as a possible factor influencing mismatches and ordering preferences. While most authors are men (77%, N=30), mismatch rates appear to be similar for male (23.3%, N=7) and female authors (22.3%, N=2), although given the size of our sample, this was not verified by our model. From Figure 2, we can observe that the proportion of masculine and feminine subjects may differ depending on author gender. Masculine-Everywhere is the predominant case for male authors, who never produce the Feminine-Everywhere. Female authors, on the other hand, appear to equally use Masculine-Feminine and Masculine-Masculine, with no Feminine-First mismatches.

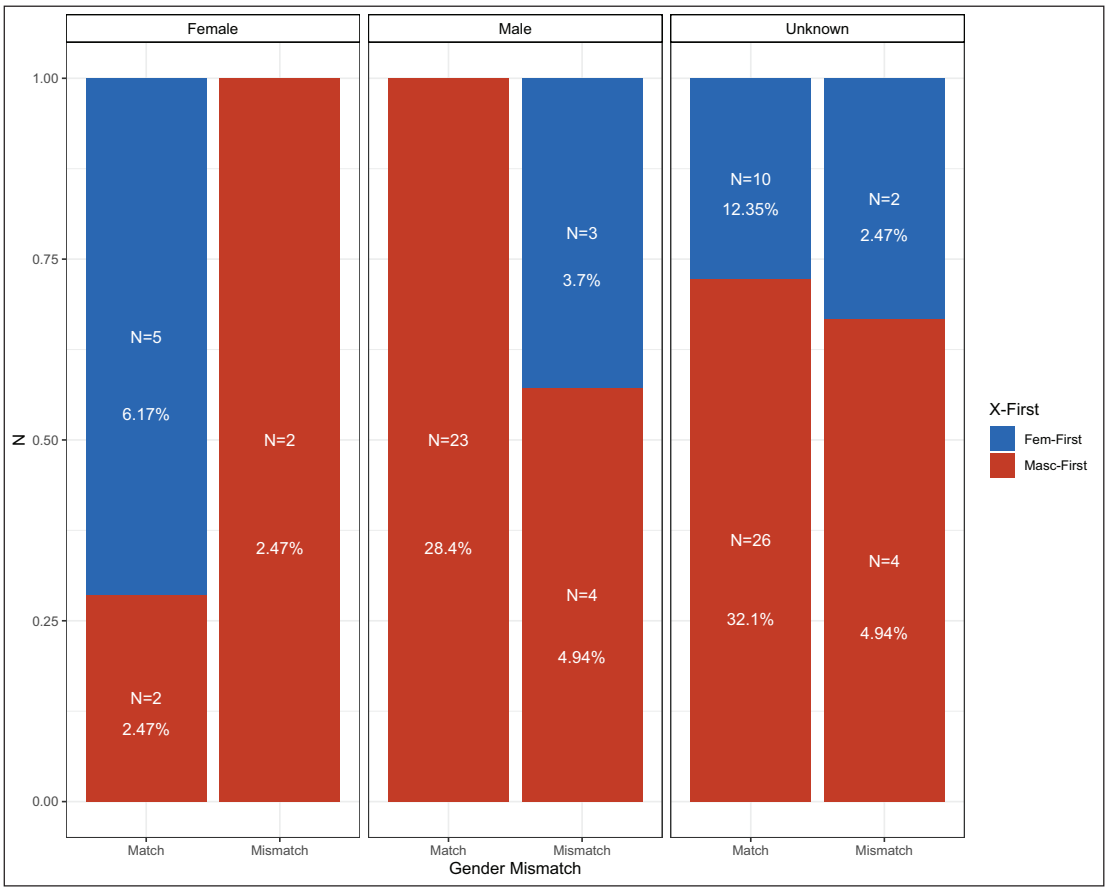


Figure 2: Order preferences per author gender for nominal corpus data

2.4. Interim discussion

Our corpus findings reveal that gender mismatch is well attested in French adjectival and nominal predicate ellipsis, in both orderings. We also found several tendencies for the production of gender mismatches under ellipsis of predicative adjectives (mismatch rate of 40.03%, N=225) and nouns (mismatch rate of 18.5%, N=15), which we present below.

Overall, animate and pronominal subjects favour mismatches, and author gender has an effect on the ordering of the conjuncts. Formal factors of the target predicate such as syncretism, homophony, as well as lexical frequency, had no robust effect, but some trends emerge, for which we evoke several hypotheses, ahead of experimental data.

First, contrary to initial hypotheses, pure indexicals (1st and 2nd person pronouns) which are unmarked for gender do not seem to favour mismatches, while marked pronouns do. While these pure indexical assume some associated demonstration for referent identification, 3rd person pronouns are directly referential, with no extra demonstration needed (Kaplan 1989). For these pronouns, which are not bound variables, gender features may be syntactically optional, but they are semantically interpreted, with the contrast between [feminine] and [masculine] being highly accessible.

The lexical frequency hypothesis posits frequency-based coding efficiency; and the ellipted material must be accessible in addressees' active memory. Haspelmath's (2020) theory would predict differential marking between the Masculine-First ordering and its marked counterpart, Feminine-First. Corpus data reveal a preference for the Feminine-First pattern in cases of mismatch, despite a strong mention bias in favour of masculine referents. The frequent ellipsis of masculine predicates (Feminine-First) is unsurprising given that they are more easily predicted and more accessible, per lexical frequency hypotheses.

Another possible explanation for ordering preferences may be syntactic: biases have been observed in French coordinate structures by Abeillé et al. (2018) and Da Cunha (2024), where male speakers clearly preferred Masculine-First orders, hinting at a syntactic gender bias correlated to the masculine social gender. Within stripping constructions, with two conjunct clauses, it is not surprising to observe the same preference for Masculine-First (in line with Aparicio et al. 2015, Sprouse et al. 2022, and Gandón-Chapela & Gallardo-del-Puerto 2025): this may be an effect of coordinate structures rather than a product of ellipsis or gender mismatches.

Our data show that gender mismatches are possible for French predicate adjectives and nouns, in both orders, even for alternating and suppletive pairs, suggesting grammatical gender may be an inflectional feature on both adjectives and human nouns. Given the small sample of predicative nouns, definitive conclusions are difficult based on corpus data alone. This is why we also ran acceptability studies.

3. Experiments on French ellipsis

We ran experiments with predicative adjectives and role nouns, to evaluate relative penalties (or meliorations) of gender mismatches per clause type (stripping ellipsis, *le* anaphora, and full repetitions), ordering, and participant gender, among other factors.

3.1. Experiment on predicative adjectives

Having collected corpus-based data on elliptical utterances, we collected acceptability data comparing 'deep' and 'surface' anaphora (Hankamer & Sag 1976), to see whether gender mismatched predicative adjectives are more acceptable in anaphoric or elliptical conditions.

3.1.1. Methodology

We ran an online acceptability judgment task (1–10 scale), coded and hosted on IbexFarm (Drummond 2007). After ethics approval by University Paris Cité Ethics Committee, we recruited 94 self-declared native French speakers on the RISC website (https://expesciences.risc.cnrs.fr/). After filtering out participants who did not meet the necessary criteria[7], we analysed the data from 89 participants, 57 of whom were women (64%), 30 were men and 2 were non-binary.

The experiment followed a 2×2 Latin square design with ±ELLIPSIS and ±MISMATCH as factors, for a total of 16 target items. The experiment also included 12 control items from An & Abeillé (2021) which tested gender agreement of adjectives, and 16 filler items on extraction out of *wh-* questions.[8]

For the target items, the +ELLIPSIS condition featured a simple stripping construction (Table 2), and the anaphoric condition featured the clitic pronoun *le*. Order was used as an intra-item variable, with half of the items following a Masculine-First order. We used human, non-pronominal subjects (using names with unambiguous social and grammatical gender). Only adjectives with an audible gender alternation were used (non-syncretic, non-homophones).

3.1.2. Hypotheses

We expect an overall preference for ellipsis: in an acceptability task, Bîlbîie (2023) found an interaction between gender mismatches and sentence types in Romanian, where non-elliptical

[7] Each item was followed by a simple yes/no comprehension question, participants who scored below 85% accuracy were not kept for analysis. At the end of each experiment, participants were presented with a short meta-data questionnaire (age, gender, native language, education level and country of residence): we thus also excluded non-native speakers.

[8] Experimental items (including filler and control items), data and scripts can be found on the project's OSF repository: https://osf.io/5azyd/.

Table 2: Example items for each condition (Experiment 1)

	UTTERANCE	±MISMATCH	ORDER	Item
1a	Ellipsis	Mismatch	Masc-Fem	Paul est grand, et Marie aussi.
1b	Ellipsis	Match	Masc-Masc	Paul est grand, et Michel aussi. *Paul is tall.M, and Marie/Michel too.*
1c	Anaphora	Mismatch	Masc-Fem	Paul est grand, et Marie l'est aussi.
1d	Anaphora	Match	Masc-Masc	Paul est grand, et Michel l'est aussi. *Paul is tall.M, and Marie/Michel (it) is too.*
2a	Ellipsis	Mismatch	Fem-Masc	Alice est belle, et Victor aussi.
2b	Ellipsis	Match	Fem-Fem	Alice est belle, et Léa aussi. *Alice is beautiful.F, and Victor/Léa too.*
2c	Anaphora	Mismatch	Fem-Masc	Alice est belle, et Victor l'est aussi.
2d	Anaphora	Match	Fem-Fem	Alice est belle, et Léa l'est aussi. *Alice is beautiful.F, and Victor/Léa (it) is too.*

utterances carried a significant repetition penalty. These results were likely due to a redundancy penalty, which Amsili et al. (2016) argued to be heightened by the presence of the additive adverb *aussi* 'too'. Similarly, results from Aparicio et al. (2015) showed ellipsis sensitivity to an antecedent's syntactic representation, with results highlighting different processes for anaphora and ellipsis.

In line with syntactic, semantic, and discursive parallelism constraints in coordinate structures (Frazier et al. 1984; Bîlbîie 2021), matching cases should be more acceptable overall, given that they satisfy all relevant constraints. Building on Hankamer & Sag's (1976) deep and surface anaphora distinction, we expect a stronger parallelism constraint for surface anaphora (stripping), where mismatches should incur a higher cost. If ellipsis patterns with anaphora, we should not observe any interaction between the utterance type and the mismatch acceptability, with an overall preference for +MATCH.

Finally, the preference for Masculine-First in nominal coordination (Abeillé et al. 2018) as well as syntactic preferences for masculine subjects (da Cunha & Abeillé 2022, da Cunha 2024) lead us to hypothesise that Masculine-First will be rated higher in all conditions. In line with corpus results, this may also interact with participant gender.

3.1.3. Results

All conditions received high acceptability ratings, far above ungrammatical controls (shown in Figure 3). We found a robust main effect of mismatch (a mismatch penalty), as well as a robust main effect of the utterance type (a preference for ellipsis), compatible with a redundancy penalty caused by the pronominal *le*, which retrieves the material in the antecedent clause, as well as the repeated copula. Interaction coefficients show no robust interaction between ellipsis and mismatch, no robust effect of ordering, and no robust interaction between order and ellipsis. We report these results in Figure 3 and detail all model results in the Appendix (Appendix B). Regarding order, we found a non-robust tendency for Masculine-First to be preferred, across ±MISMATCH and ±ELLIPSIS.

Mismatches under ellipsis were rated higher than ungrammatical controls (Figure 3), and tendentially higher than under anaphora (non-robust interaction), in both orders, unlike in Spanish, where Aparicio et al. (2015) found that the Feminine-Masculine order was more penalised than other mismatch conditions (Masculine-First, and number mismatches). Our results are compatible with both the direct interpretation approach (without syntactic reconstruction), and the deletion-under-identity approach (assuming identity ignores adjective agreement inflectional features).

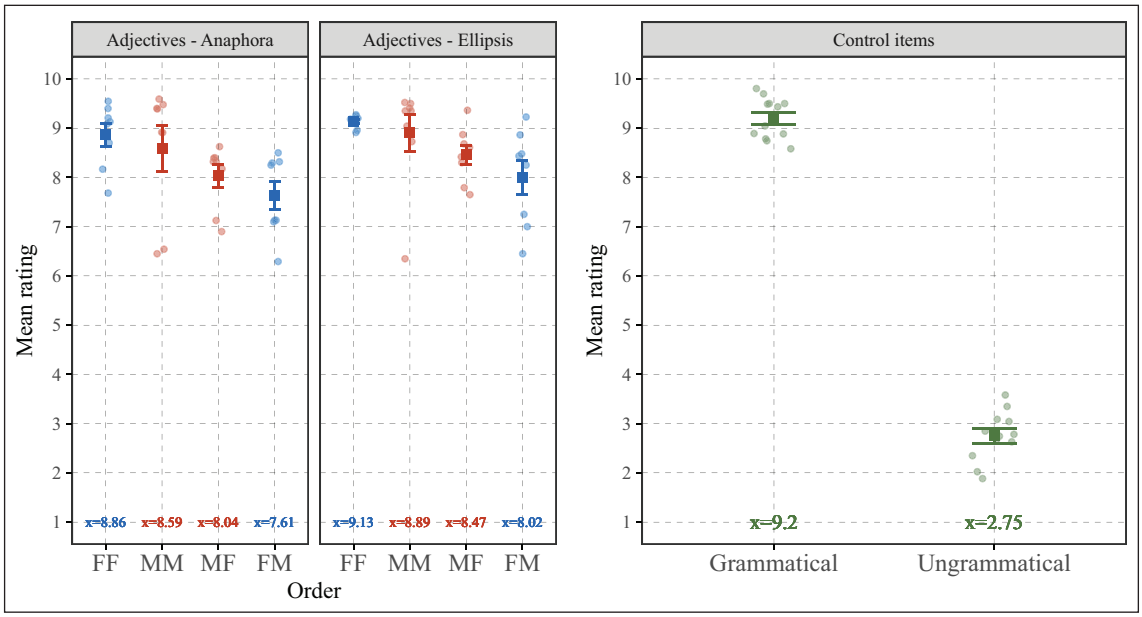


Figure 3: Mean judgments per sentence, order for adjectives (Experiment 1) and control items (right). Dots represent means per item, and boxes represent means per condition

We found no robust effect of participant gender and only observe a tendency for female participants to give slightly lower ratings to mismatches, especially under anaphora. Male participants showed a slight preference for +MATCH under anaphora, and -MATCH under ellipsis. We found no robust effect of gender on order: all participants seem to prefer Masculine-First across all conditions.

3.2. Experiments on predicative nouns

We sought to test the acceptability of gender mismatches for predicative role nouns, while accounting for gender stereotypes.

3.2.1. Methodology

To control the effect of gender stereotypes on role noun interpretation, we used normed scores from Misersky et al. (2014) to distinguish non-biased role nouns (with stereotypicality scores indicating that they do not carry stereotype information, e.g. *étudiant/e* 'student.M/F') from male-biased (e.g. *chirurgien/ne* 'surgeon.M/F') and female-biased (e.g. *infirmier/ère* 'nurse.M/F') role nouns. Experiment 2 tested 24 non-biased role nouns, and Experiment 3 tested 24 biased ones (12 female-biased and 12 male-biased role nouns).

Participants were randomly assigned to one of the two experiments, with post-hoc counterbalancing. Both experiments were online acceptability judgment tasks (1–10 scales), coded and hosted on IbexFarm (Drummond 2007). Following ethics approval, we recruited 85 self-declared native French speakers, and kept 83 for analysis, following the exclusion of participants who had obtained low comprehension scores, using the same methodology as for Experiment 1. Among these participants, 40 were women (48.2%), 41 were men (49.4%), and 2 were non-binary (2.4%), for a median age of 30.

Each experiment followed a 2×3 Latin Square design with UTTERANCE (Ellipsis/Anaphora/Full repetition) and ±MISMATCH as factors, totalling 6 conditions for 24 target items (Table 3). For both experiments, the 24 target items were supplemented with 12 control items (used for Experiment 1) and 16 filler items on the French active/passive alternation.[9]

In the Full repetition condition, both the determiner and the noun are repeated: for -MATCH, we expect this condition to serve as an ungrammatical control. For each item, we used nominal human subjects, with gender-unambiguous names. For Experiment 2, order was encoded as an intra-item variable. For Experiment 3, order was congruent with stereotypicality: items featuring a male-biased role noun followed a Masculine-First order, female-biased ones were Feminine-First.

[9] Experiment items (including filler and control items, Miserksy scores), data and scripts can be found on the project's OSF repository: https://osf.io/5azyd/.

Role nouns were selected for each experiment according to their Misersky score (Misersky et al. 2014): non-biased nouns were those with a score between 0.45 and 0.55; biased nouns had scores above 0.6 (female bias) or below 0.4 (male bias). We also controlled for morpho-phonological alternations, excluding syncretic and homophone forms (epicenes and common gender nouns): half of the noun pairs had a final consonant alternation (e.g. *assistant/assistante* 'assistant.M/F'). For the remaining nouns, all of which ended in *-eur* in the masculine, half of these took *-euse* (e.g. *chanteur/chanteuse* 'singer.M/F'), and half took *-rice* (e.g. *instituteur/institutrice* 'school-teacher.M/F') endings in the feminine. These correspond to the most frequent alternations for gender iconic pairs in the Morphalou database (Bonami & Boyé 2019). Twelve grammatical controls were also included, identical to those described in Section 3.1.1.

Table 3: Example items for each experimental condition

	UTTERANCE	±MISMATCH	Item
1a	Ellipsis	Match	Jeanne est une lectrice rapide, et Clara aussi.
1b	Ellipsis	Mismatch	Jeanne est une lectrice rapide, et Corentin aussi. *Jeanne is a fast reader.F, and Clara/Corentin too.*
1c	Anaphora	Match	Jeanne est une lectrice rapide, et Clara l'est aussi.
1d	Anaphora	Mismatch	Jeanne est une lectrice rapide, et Corentin l'est aussi. *Jeanne is a fast reader.F, and Clara/Corentin (it) is too.*
1e	Full repetition	Match	Jeanne est une lectrice rapide, et Clara est une lectrice rapide aussi.
1f	Full repetition	Mismatch	Jeanne est une lectrice rapide, et Corentin est une lectrice rapide aussi. *Jeanne is a fast reader.F, and Clara/Corentin is a fast reader.F too.*

3.2.2. Hypotheses

We only took (non-syncretic) alternating role nouns and tested several hypotheses. If French follows Greek (Merchant 2014) and English (Sprouse et al. 2022), these nouns should license mismatches only in the Masculine-Feminine order: ellipsis should be penalised, compared to anaphora, in the Feminine-Masculine order. If one assumes syntactic reconstruction and gender as a derivational feature on nouns, ellipsis with mismatch should be rated like full repetition. If on the other hand, gender is an agreement inflectional feature for these nouns like for adjectives and/or ellipsis is interpreted directly (without syntactic reconstruction), we expect no ellipsis mismatch penalty, in either order.

We do not test common gender nouns, but we expect the mismatch penalty to grow with the formal distance between the two forms. Moreover, given that gender stereotypes are inherent to the interpretation of grammatical gender for human nouns (Richy & Burnett 2019; Mickus et al. 2019), we expect an effect of gender stereotypes on the acceptability of all utterances, with the hypothesis that the stereotype violation adds a superadditive penalty in cases of mismatching gender.

Finally, as was the case for adjectival predicates in French (Experiment 1) and Spanish (Aparicio et al. 2015), we expect an overall preference for the Masculine-First condition, interacting with participant gender.

3.2.3. Results

For both experiments, we found a robust main effect of utterance types, with a repetition penalty. We observe an acceptability cline, such that ellipsis is preferred to anaphora (regardless of ±MISMATCH), and both are preferred to full repetition. For both experiments, we also find a robust main effect of +MISMATCH corresponding to a mismatch penalty across all conditions. For both experiments, we found no robust effect of noun morphology, unlike what was proposed for English (Sprouse et al. (2022) and Greek (Merchant 2014), namely that symmetrical mismatch is only allowed for common gender nouns. It seems more in line with Bobaljik & Zocca's (2011) proposal for Brazilian Portuguese, that some alternating role nouns have inflectional gender like adjectives. But while half of the nouns we tested have a consonant gender alternation (*assistant/e* 'assistant.M/F') like most French adjectives (*grand/e* 'tall.M/F'), the other half has alternating types (*chanteur/euse* 'singer.M/F', *lecteur/rice* 'reader.M/F') with no equivalent for adjectives. No difference was found between these two morphological subtypes in both experiments[10].

Observations on mean judgments show MISMATCH interacting with gender biases: biased role nouns (Experiment 3) received lower ratings under mismatch than non-biased role nouns (Experiment 2). For both experiments, mismatched full repetitions are close to ungrammatical controls (Figure 4).[11]

For non-biased nouns (Experiment 2, Figure 4), as with predicative adjectives (Experiment 1), mismatches under ellipsis received higher ratings than mismatches under anaphora, in both orders. This casts doubt on the syntactic reconstruction hypothesis for ellipsis and/or the analysis of nominal gender as derivational. We also found a robust interaction between MISMATCH and anaphora/ellipsis, both corresponding to an acceptability increase compared to the MISMATCH:full repetition condition, which acted as an acceptability penalty ceiling.

[10] Item-per-item results for all three experiments are available on the project's OSF repository.
[11] Inferential modelling results are detailed in Appendix B.

While order was randomly encoded intra-item for Experiment 2, we found a robust main interaction between order and utterance type: Feminine-Masculine elliptical and anaphoric items were preferred to Feminine-Masculine full repetition sentences. In the full repetition condition, ratings were at ceiling for a mismatch penalty: masculine nouns are not suitable with a feminine subject, and feminine nouns are not suitable with a masculine subject (Figure 4). The suspension of this gender competition effect under ellipsis is in line with Sudo & Spathas's (2020) account.

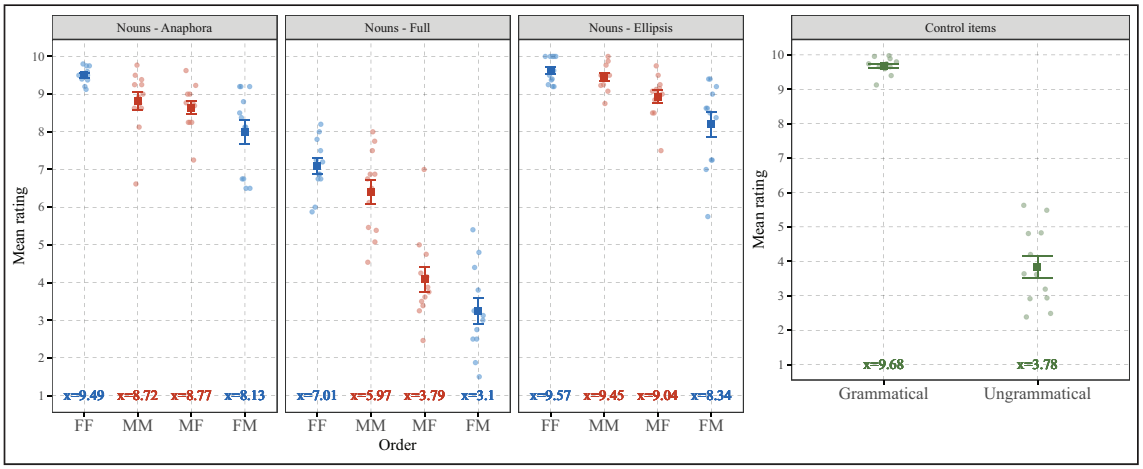


Figure 4: Mean judgments per utterance, order for non-biased nouns (Experiment 2) and control items. Dots represent means per item, and boxes represent means per condition

For biased nouns (Experiment 3, Figure 5), we found a robust MISMATCH penalty (stronger than for non-biased nouns) across all conditions, with a tendential melioration for anaphora compared to ellipsis, although both are robustly preferred to full repetition sentences. For this experiment, order was congruent with stereotype: male-biased nouns only appeared in the Masculine-First order, and female-biased nouns in the Feminine-First order. We found a robust order effect across all conditions (both directions showed equal MISMATCH penalties, with no preference between Masculine or Feminine antecedents). We found no robust interaction of order and utterance type, but observe tendencies in Figure 5.

We found no robust effect of participant gender for either experiment, but plot acceptability tendencies in Figure 6, where we observe an overall stronger penalty of MISMATCH from male participants.

Overall, experimental results are in line with corpus results and confirm the acceptability of gender mismatches under nominal ellipsis, with a tendential preference for Masculine-First for predicate nouns, for all corpus authors and experiment participants. Corpus data follow judgments for MISMATCH, with a clear preference for MATCH in experimental conditions. As with corpus data, we observe no robust effect of predicate noun morphology.

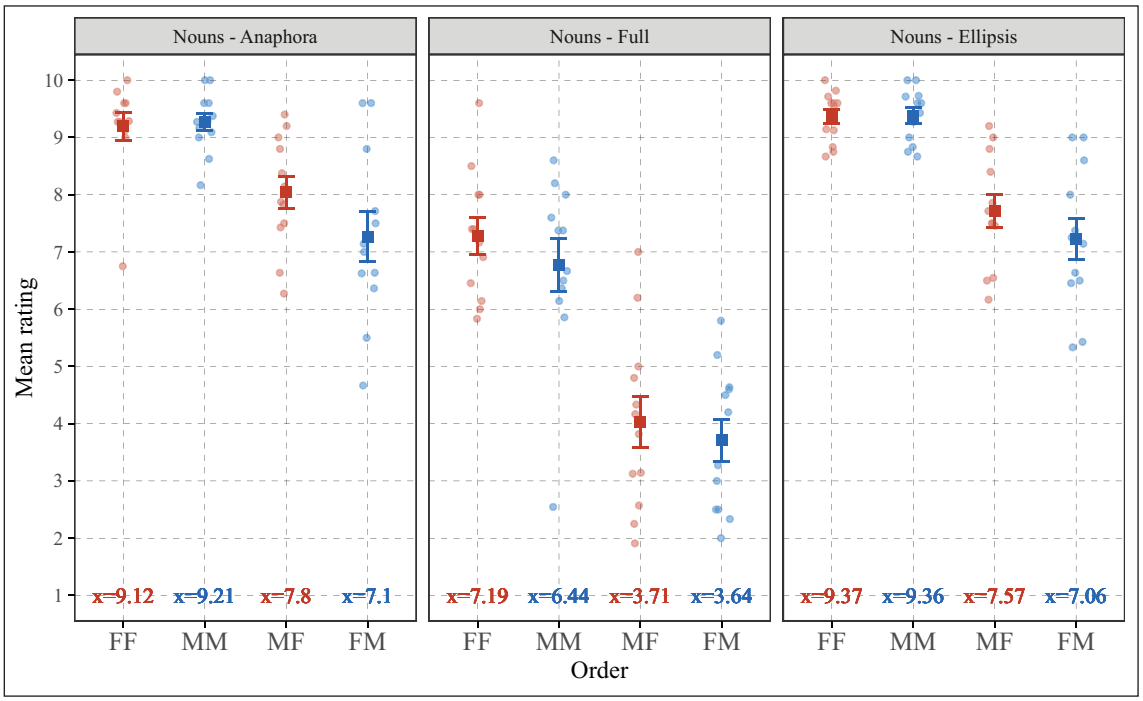


Figure 5: Mean judgments per utterance and order for biased nouns (Experiment 3). Dots represent means per item, and boxes represent means per condition

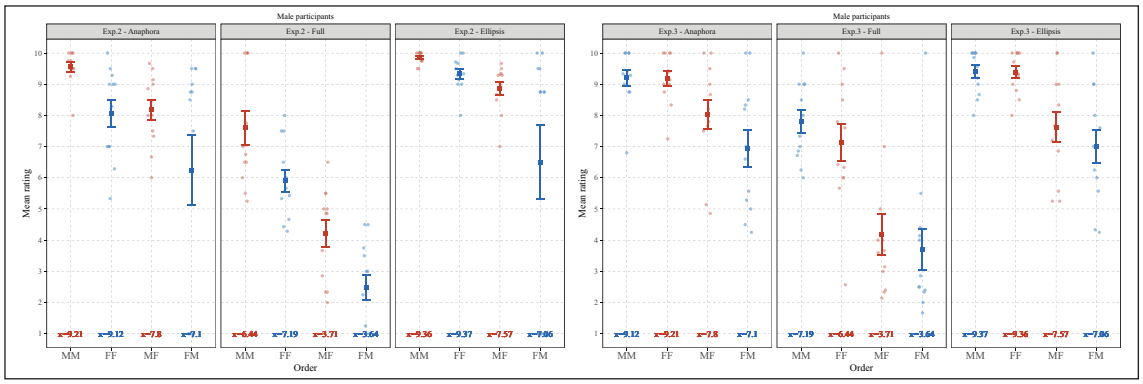


Figure 6: Mean judgments per utterance, order, and for male participants (Experiments 2 and 3)

4. Discussion

The interaction between utterance type and mismatch acceptability, which emerges for predicative uses of both adjectives and nouns, seems to confirm the hypothesis that adjectives

and human nouns pattern together in French. This is at odds with Merchant's (2014) proposal for Greek and Sprouse et al.'s (2022) for English. Our results are compatible with two possible theories. Either gender is inflectional (an agreement feature) for French predicative nouns as it is for adjectives, which would allow an analysis in terms of syntactic reconstruction under identity; or, one might assume that nominal ellipsis falls under a direct interpretation approach, with antecedent accommodation similar to pronominal resolution.

Moreover, it appears that ellipsis and anaphora pattern alike in French, which challenges previous predictions based on syntactic reconstruction at the ellipsis site. For French adjectives and nouns, the mismatch penalty under ellipsis is even smaller than under anaphora. Parker (2019) offers a possible explanation to this penalty asymmetry. Anaphora resolution based on a cue-based retrieval mechanism rests upon a cognitive principle which encodes elements in active memory. Cues allow us to extract and activate an antecedent when its correspondence with the anaphoric form is highest: according to Parker (2019), mismatches and partial matches disturb sentence processing of anaphora, which may incur acceptability penalties.

Regarding order effects, several explanations are possible at this stage, compatible with markedness and accessibility theories. It is possible to consider them a consequence of syntactic markedness, where the formally marked construction (Feminine-First) corresponds to the informationally marked construction (Comrie 1989).

For French corpus and experimental data, the Masculine-Everywhere preference is maintained, and while Feminine-everywhere is rare in corpora, it is rated above mismatched conditions by participants in experimental conditions.

In cases of mismatch, ordering preferences are reversed between corpus data (Feminine-First) and experiments (Masculine-First). This effect may be attributed to speaker gender: male authors account for an overwhelming majority of our sample, while female participants are more numerous, but this likely interacts with other factors, as Feminine-Masculine order is more frequent than Masculine-Feminine in corpora (despite a majority of male authors), while the reverse is true in experiments (despite a majority of female participants). It is possible that the grammatical masculine form is unmarked for men (if only in production data) which, consequently, renders the masculine form more accessible for ellipsis; for women, both forms would be more symmetrical in their markedness. The preference asymmetries we found are also compatible with a difference between production and comprehension: if our active memory stores some ungendered lemma, and not the grammatical form, the lemma (which is identical to the masculine form in French) is then more accessible during production.

5. Conclusion

We found that gender mismatches are attested in French corpora for predicative adjectives and nouns, and are judged acceptable in experimental studies, in both orders (Masc-Fem, Fem-Masc), across morphological types and gender stereotypes. Experimental results also reveal an interaction between mismatch acceptability and utterance type, with a general ellipsis preference for nouns and adjectives alike:

+MISMATCH:

ADJECTIVES: Ellipsis > Anaphora ≫ Ungrammatical controls

NOUNS: Ellipsis ≥ Anaphora ≫ Ungrammatical controls ≥ Full repetition

Gender mismatch under predicate ellipsis (stripping) was rated higher than with predicate anaphora, unlike what was found for Spanish adjectives (Aparicio et al. 2015). This is compatible both with deletion-based theories and direct interpretation of ellipsis for adjectives, since adjective gender is an agreement feature (contextual, inflectional) to be ignored under ellipsis. But for nouns, this challenges previous theories (Merchant 2014; Sprouse et al. 2022), which derive ellipsis by deletion-under-identity and assume nominal gender as inherent or derivational. This seems more in line with Boballjik & Zocca's (2011) proposal for Brazilian Portuguese that gender is inflectional for some alternating role nouns. The *médico/médica* class would have to be extended for French to nouns which do not morphologically patterns like adjectives, and we do not find evidence for an asymmetric *ator/atriz* class.

Overall, for corpus and experiments alike, we observe no robust effect of the morphophonology of nouns, nor do we observe a robust effect of lexical frequency. For adjectives, corpus data show that syncretic and homophone adjectives are not a factor favouring mismatches under ellipsis, and experimental data confirm that mismatched alternating pairs are fully acceptable under ellipsis, unlike with full repetition.

Contrary to Greek and English, Masculine-Feminine ordering was not the only possible order for gender mismatch for (alternating) nouns under ellipsis in French. In our corpus data, Fem-Masc ordering was among the factors favouring gender mismatch under ellipsis for both adjectives and nouns. In our experimental data, both orderings were acceptable, with a preference for Masc-Fem for ellipsis and anaphora alike, for both adjectives and nouns. The main factor disfavouring gender mismatch in our experiments was nominal gender stereotypes, which was not taken into account in previous studies.

The present study accounts for several factors which are at play for the production and comprehension of gender mismatches under stripping ellipsis, and brings empirical support for these conclusions, in a quantitative syntax (corpus studies) as well as an experimental syntax

(acceptability studies) approach. Two conclusions are suggested, which are not incompatible: the gender of French alternating role nouns is contextual and inflectional (Spencer 2002; Bonami & Boyé 2019) and thus ignored under ellipsis, and/or predicate ellipsis (stripping) is directly interpreted with no syntactic reconstruction (Ginzburg & Sag 2000; Culicover & Jackendoff 2005; Nykiel & Kim 2021). The second conclusion would apply to other languages as well, which remain to be tested empirically.

Acknowledgements: We thank the anonymous reviewers, as well as the audience at Experimental and Corpus-Based Approaches to Ellipsis, for their helpful comments and discussions.

Funding: This work has received support from the investment programme "France 2030" as part of the IdEx programme (ANR-18-IDEX-0001) implemented by Université Paris Cité, under which the inIdEx project EFL is conducted.

Appendix

Appendix A

We adopt a Bayesian framework for our inferential statistical modelling, to translate data towards a precise estimate of the proportional distribution of the data. We used the 'brms' package (Bürkner 2017) in R, and the Markov chain Monte-Carlo sampling method (MCMC) to get a representative sample of the posterior probability distribution.

To model corpus data, which are count data, we use the Bernoulli likelihood, corresponding to a 0 (-MISMATCH) or 1 response with a probability of θ (in the log-odds space). We use weakly informative priors (normal distribution centred on 0) and default priors for thresholds and standard deviations of random effects (occurrence and author ID).

For each model, we ran four sampling chains of 2000 iterations, with 1000 warm-up iterations, for a total of 4000 chains per parameter tuple.

Posterior probabilities are obtained using the hypothesis testing function of brms, based on evidence ratios. Posterior predictive checks can be found in the project's OSF.

Model 1: Adjectival Gender mismatch

> GEN_MISMATCH~XFIRST*(NUM_MISMATCH+HOMOPHONE+SYNCRETIC+SUBJ1_CAT+SUBJ2_CAT+SUBJ1_HUM+SUBJ2_HUM+SUBJ1_GENDERED+SUBJ2_GENDERED+AUTHOR_GENDER-)+(XFIRST*(NUM_MISMATCH+HOMOPHONE+SYNCRETIC+SUBJ1_CAT+SUBJ2_CAT+SUBJ1_HUM+SUBJ2_HUM+SUBJ1_GENDERED+SUBJ2_GENDERED+AUTHOR_GENDER)|AUTHOR_

ID)+(XFIRST*(NUM_MISMATCH+HOMOPHONE+SYNCRETIC+SUBJ1_CAT+SUBJ2_CAT+SUBJ1_HUM+SUBJ2_HUM+SUBJ1_GENDERED+SUBJ2_GENDERED+AUTHOR_GENDER)|LEMMA)

Reference levels of predictors are set to 'Match' (GENDER_MISMATCH & NUMBER_MISMATCH), 'alternating' adjectives (non HOMOPHONE, non SYNCRETIC), 'nominal' 'non-human' and 'gendered' subjects (SUBJ1, SUBJ2), and 'unknown' AUTHOR_GENDER.

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
$\beta_{\text{Fem-First}}$	2.08	0.78	[0.55;3.62]	1*
$\beta_{\text{Nb-Mismatch}}$	0.32	0.77	[−1.2;1.82]	0.67
$\beta_{\text{Homophone}}$	0.15	0.72	[−1.29;1.49	0.59
$\beta_{\text{Syncretic}}$	−0.94	0.78	[−2.43;0.6]	0.88
$\beta_{\text{Pro-subj1}}$	1.58	0.78	[0;3.12]	0.98*
$\beta_{\text{Pro-subj2}}$	0.60	0.80	[−0.99;2.16]	0.78
$\beta_{\text{Human-subj1}}$	1.03	0.75	[−0.45:2.5]	0.91
$\beta_{\text{Human-subj2}}$	0.13	0.79	[−1.4;1.72]	0.56
$\beta_{\text{Ungen-subj1}}$	0.14	0.82	[−1.47;1.76]	0.57
$\beta_{\text{Ungen-subj2}}$	0.46	0.82	[−1.15;2.06]	0.71
$\beta_{\text{Female-author}}$	0.43	0.75	[−1.05;1.86]	0.71
$\beta_{\text{Male-author}}$	−0.88	0.71	[−2.29;0.49]	0.89
$\beta_{\text{Fem-First:Nb-mismatch}}$	−0.22	0.89	[−1.94;1.51]	0.59
$\beta_{\text{Fem-First:Homo phone}}$	0.54	0.87	[−1.1;2.28]	0.73
$\beta_{\text{Fem-First:Syncretic}}$	−0.40	0.93	[−2.22;1.36]	0.66
$\beta_{\text{Fem-First:Pro-subj1}}$	0.02	0.92	[−1.77;1.88]	0.5
$\beta_{\text{Fem-First:Pro-subj2}}$	−0.39	0.90	[−2.14;1.39]	0.66
$\beta_{\text{Fem-First:Human-subj1}}$	0.84	0.90	[−0.94;2.6]	0.82
$\beta_{\text{Fem-First:Human-subj2}}$	1.55	0.90	|−0.21;3.34]	0.96*
$\beta_{\text{Fem-First:Ungen-subj1}}$	−0.06	0.92	[−1.86;1.71]	0.53
$\beta_{\text{Fem-First:Ungen-subj2}}$	−0.59	0.92	[−2.35;1.21]	0.74
$\beta_{\text{Fem-First:Female-author}}$	−0.68	0.91	[−2.44;1.08]	0.78
$\beta_{\text{Fem-First:Male-author}}$	0.57	0.95	[−1.3;2.44]	0.72

Model 2: Nominal Gender mismatch

GEN_MISMATCH~XFIRST*(NUM_MISMATCH+HOMOPHONE+SYNCRETIC+SUBJ1_CAT+SUBJ2_CAT+SUBJ1_HUM+SUBJ2_HUM+SUBJ1_GENDERED+SUBJ2_GENDERED+AUTHOR_GENDER+DETERMINER)+(XFIRST*(NUM_MISMATCH+HOMOPHONE+SYNCRETIC+SUBJ1_CAT+SUBJ2_CAT+SUBJ1_HUM+SUBJ2_HUM+SUBJ1_GENDERED+SUBJ2_GENDERED+AUTHOR_GENDER+DETERMINER)|AUTHOR_ID)+(XFIRST*(NUM_MISMATCH+HOMOPHONE+SYNCRETIC+SUBJ1_CAT+SUBJ2_CAT+SUBJ1_HUM+SUBJ2_HUM+SUBJ1_GENDERED+SUBJ2_GENDERED+AUTHOR_GENDER+DETERMINER)|LEMMA)

Reference levels of predictors are identical to Model 1, with an additional factor DETERMINER ('no').

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
$\beta_{\text{Fern-First}}$	0.15	1.00	[−1.79;2.04]	0.56
$\beta_{\text{Nb-mismatch}}$	0.12	1.02	[−1.85;2.14]	0.54
$\beta_{\text{Homophone}}$	−0.08	0.98	[−1.98;1.87]	0.53
$\beta_{\text{Syncretic}}$	0.10	0.04	[−1.76;2.01]	0.53
$\beta_{\text{Pro.subj1}}$	0.24	0.97	[−1.65;2.15]	0.6
$\beta_{\text{Pro-subj2}}$	0.28	0.97	[−1.59.2.17]	0.61
$\beta_{\text{Human-subj1}}$	0.18	0.98	[−1.72;2.05]	0.57
$\beta_{\text{Human-subj2}}$	0.03	0.98	[−1.91;2]	0.52
$\beta_{\text{Ungen-subj1}}$	0.22	0.97	[−1.67;2.12]	0.58
$\beta_{\text{Ungen-subj2}}$	0.20	0.94	[−1.66;2.08]	0.58
$\beta_{\text{Female-author}}$	0.01	0.99	[−1.93;1.95]	0.51
$\beta_{\text{Male-author}}$	0.08	0.96	[−1.75;1.99]	0.54
$\beta_{\text{Determiner}}$	0.00	0.95	[−1.85;1.87]	0.5
$\beta_{\text{Fem-First:Nb-mismatch}}$	0.01	1.00	[−1.94;1.98]	0.5
$\beta_{\text{Fem-First:Homophone}}$	0.18	0.98	[−1.8;1.98]	0.57
$\beta_{\text{Fem-First:Syncretic}}$	0.13	1.01	[−1.82;2.07]	0.55
$\beta_{\text{Fem-First:Pro-subj1}}$	−0.21	0.99	[−2.13;1.71]	0.59
$\beta_{\text{Fem-First:Pro-subj2}}$	−0.13	0.97	[−2.06;1.8]	0.55
$\beta_{\text{Fem-First.Human-subj1}}$	0.04	1.02	[−1.95;2.06]	0.51
$\beta_{\text{Fern-First:Human-subj2}}$	0.15	1.00	[−1.81;2.12]	0.55
$\beta_{\text{Fem-First:Ungen-subj1}}$	−0.26	0.98	[−2.15;1.68]	0.61
$\beta_{\text{Fem-First:Ungen-subj2}}$	−0.07	0.97	[−1.97;1.83]	0.53
$\beta_{\text{Fem-First:Female-author}}$	−0.19	1.00	[−2.09;1.74]	0.57
$\beta_{\text{Fem-First:Male-author}}$	0.17	1.01	[−1.85;2.15]	0.57
$\beta_{\text{Fem-First.Determiner}}$	−0.08	1.01	[−2.09;1.86]	0.53

Appendix B

Inferential modelling of experimental data was done within the Bayesian framework: we ran cumulative mixed effects ordinal logistic regression models, adjusted using the Stan language and the 'brms' package (Bürkner 2017) in R. Acceptability ratings call for such models as the categorical predictors are ordered, and we seek to estimate the maximal cumulative probability of observing an ordinal variable Y (participants' judgments), given the fixed effects matrices (predictors) and random effects (participant and item). Each model considers population effects, which are those random effects, as well as the parameters of interest. Models were fit using factored predictors (using centred treatment coding), with intercepts by-participant and by-item, as well as random slopes for random effects.

Posterior predictive checks, priors, and model specifications can be found in this project's OSF.

We present summaries of fixed effect posterior distributions for the models.

Model 3: (Experiment 1)

RATING~ORDER*UTTERANCE+(ORDER*UTTERANCE|PARTICIPANT)+(ORDER*UTTERANCE|ITEM)

Reference levels of predictors are set to 'Masculine-Masculine' ORDER, and 'Full repetition' UTTERANCE.

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
$\beta_{\text{Fem-Fem}}$	−0.18	0.61	[−1.38;1.02]	0.62
$\beta_{\text{Fem-Masc}}$	−1.52	0.84	[−3.1;0.15]	0.96*
$\beta_{\text{Masc-Fem}}$	−1.51	0.60	[−2.68;−0.29]	0.99*
β_{Anaphora}	4.65	0.49	[3.63;5.58]	1*
β_{Ellipsis}	4.99	0.57	[3.83;6.06]	1*
$\beta_{\text{Fem-Fem.Anaphora}}$	−0.54	0.69	[−1.89;0.82]	0.78
$\beta_{\text{Fem-Masc:Anaphora}}$	1.68	0.64	[0.31;2.89]	0.99*
$\beta_{\text{Masc-Fem:Anaphora}}$	0.88	0.73	[−0.56;2.26]	0.89
$\beta_{\text{Fem-Fem.Ellipsis}}$	−0.25	0.71	[−1.64:1.17]	0.63
$\beta_{\text{Fem-Masc:Ellipsis}}$	1.18	0.69	[−0.25:2.48]	0.95*
$\beta_{\text{Masc-Fem:Ellipsis}}$	0.74	0.71	[−0.68;2.1]	0.85

Model 4: (Experiment 1)

RATING~MISMATCH*UTTERANCE*MORPHOLOGY+(MISMATCH*UTTERANCE*MORPHOLOGY|PARTICIPANT)+(ORDER*UTTERANCE*MORPHOLOGY|ITEM)

Reference levels of predictors are set to +MISMATCH, 'consonant' MORPHOLOGY, and 'Full repetition' UTTERANCE.

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
β_{Mismatch}	−2.03	0.51	[−3.02;−1.01]	1*
β_{Anaphora}	4.54	0.56	|3.39;5.61]	1*
β_{Ellipsis}	4.92	0.62	[3.65;6.1]	1*
$\beta_{\text{-eur/-euse}}$	0.10	0.34	[−0.57;0.78]	0.63
$\beta_{\text{-eur/-rice}}$	−0.07	0.34	[−0.77;0.59]	0.58
$\beta_{\text{Mismatch:Anaphora}}$	1.60	0.54	[0.53;2.67]	1*
$\beta_{\text{Mismatch:Ellipsis}}$	1.21	0.57	[0.08;2.35]	0.98*
$\beta_{\text{Mismatch:-eur./-euse}}$	0.16	0.56	[−0.92;1.26]	0.61
$\beta_{\text{Mismatch:-eur/-rice}}$	0.37	0.56	[−0.76;1.48]	0.75
$\beta_{\text{Anaphora:-eur/-euse}}$	−0.15	0.61	[−1.39;1]	0.61
$\beta_{\text{Ellipsis:-eur/-euse}}$	−0.20	0.64	[−1.51;0.98]	0.61
$\beta_{\text{Anaphora:-eur/-rice}}$	−0.29	0.61	[−1.5;0.91]	0.68
$\beta_{\text{Ellipsis:-eur/-rice}}$	−0.50	0.62	[−1.72;0.72]	0.79
$\beta_{\text{Mismatch:Anaphora:-eur/-euse}}$	−0.28	0.77	[−1.78;1.21]	0.64
$\beta_{\text{Mismatch:Ellipsis:-eur/-euse}}$	−0.08	0.80	[−1.63;1.48]	0.54
$\beta_{\text{Mismatch:Anaphora:-eur/-rice}}$	0.36	0.78	[−1.17;1.88]	0.68
$\beta_{\text{Mismatch:Anaphora:-eur/-rice}}$	0.15	0.80	[1.41;1.74]	0.57

Model 5: (Experiment 2)

RATING~ORDER*UTTERANCE+(ORDER*UTTERANCE|PARTICIPANT)+(ORDER*UTTERANCE|ITEM)

Reference levels of predictors are identical to Model 3.

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
$\beta_{\text{Fem-Fem}}$	0.39	0.34	[−0.28;1.06]	0.88
$\beta_{\text{Fem-Masc}}$	−2.85	0.55	[−3.87;−1.73]	1*
$\beta_{\text{Masc-Fem}}$	−2.71	0.44	[−3.57;−1.82]	1*
β_{Anaphora}	3.87	0.35	[3.17;4.56]	1*
β_{Ellipsis}	3.98	0.38	[3.24;4.72]	1*
$\beta_{\text{Fem-Fem:Anaphora}}$	−0.60	0.49	[−1.57;0.37]	0.89
$\beta_{\text{Fem-Masc:Anaphora}}$	0.09	0.52	[−0.91;1.12]	0.57
$\beta_{\text{Masc-Fem:Anaphora}}$	0.40	0.50	[−0.59;1.36]	0.79
$\beta_{\text{Fem-Fem:Ellipsis}}$	−0.40	0.55	[−1.5;0.63]	0.77
$\beta_{\text{Fem-Masc:Ellipsis}}$	−0.19	0.50	[−1.17;0.79]	0.65
$\beta_{\text{Masc-Fem:Ellipsis}}$	−0.13	0.16	[−1.04;0.78]	0.61

Model 6: (Experiment 2)

RATING~MISMATCH*UTTERANCE*MORPHOLOGY+(MISMATCH*UTTERANCE*MORPHOLOGY|PARTICIPANT)+(ORDER*UTTERANCE*MORPHOLOGY|ITEM)

Reference levels of predictors are identical to Model 4.

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
β_{Mismatch}	−3.13	0.52	[−4.12;−2.05]	1*
β_{Anaphora}	4.00	0.40	[3.18;4.75	1*
β_{Ellipsis}	4.05	0.44	[3.14;4.89]	1*
$\beta_{\text{-eur/-euse}}$	−0.10	0.41	[−0.88;0.73]	0.6
$\beta_{\text{-eur/-rice}}$	0.30	0.34	[−0.36;0.96]	0.82
$\beta_{\text{Mismatch:Anaphora}}$	0.32	0.53	[−0.71;1.35]	0.73
$\beta_{\text{Mismatch:Ellipsis}}$	−0.25	0.48	[−1.19;0.7]	0.7
$\beta_{\text{Mismatch:-eur/-euse}}$	−0.43	0.58	[−1.52;0.73]	0.78
$\beta_{\text{Mismatch:-eur/-rice}}$	−0.55	0.52	[−1.55;0.48]	0.86
$\beta_{\text{Anaphora:-eur/-euse}}$	−0.17	0.56	[−1.3;0.89]	0.6
$\beta_{\text{Ellipsis:-eur/-euse}}$	0.03	0.63	[−1.28;1.23]	0.53
$\beta_{\text{Anaphora:-eur/-rice}}$	0.13	0.51	[−0.89;1.11]	0.62
$\beta_{\text{Ellipsis:-eur/-rice}}$	0.16	0.57	[−0.98;1.29]	0.62
$\beta_{\text{Mismatch:Anaphora:-eur/-euse}}$	0.23	0.76	[−1.29;1.74]	0.62
$\beta_{\text{Mismatch:Anaphora:-eur/-euse}}$	0.20	0.75	[−1.25;1.72]	0.6
$\beta_{\text{Mismatch:Anaphora:-eur/-rice}}$	−0.27	0.69	[−1.58;1.12]	0.65
$\beta_{\text{Mismatch:Ellipsis:-eur/-rice}}$	0.28	0.71	[−1.11;1.71]	0.66

Model 7: (Experiment 3)
RATING~ORDER*UTTERANCE+(ORDER*UTTERANCE|PARTICIPANT)+(ORDER*UTTERANCE|ITEM)

Reference levels of predictors are identical to Model 3.

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
$\beta_{\text{Fem-Fem}}$	0.39	0.34	[−0.28;1.06]	0.88
$\beta_{\text{Fem-Masc}}$	−2.85	0.55	[−3.87;−1.73]	1*
$\beta_{\text{Masc-Fem}}$	−2.71	0.44	[−3.57;−1.82]	1*
β_{Anaphora}	3.87	0.35	[3.17;4.56]	1*
$\beta_{\text{Ellipisis}}$	3.98	0.38	[3.24;4.72]	1*
$\beta_{\text{Fem-Fem:Anaphora}}$	−0.60	0.49	[−1.57;0.37]	0.89
$\beta_{\text{Fem-Masc:Anaphora}}$	0.09	0.52	[−0.91;1.12]	0.57
$\beta_{\text{Masc-Fem:Anaphora}}$	0.40	0.50	[−0.59;1.36]	0.79
$\beta_{\text{Fem-Fem:Ellipsis}}$	−0.40	0.55	[−1.5;0.63]	0.77
$\beta_{\text{Fem-Masc:Ellipsis}}$	−0.19	0.50	[−1.17;0.79]	0.65
$\beta_{\text{Masc-Fem:Ellipsis}}$	−0.13	0.46	[−1.04;0.78]	0.61

Model 8: (Experiment 3)

RATING~MISMATCH*UTTERANCE*MORPHOLOGY+(MISMATCH*UTTERANCE*MORPHOLOGY|PARTICIPANT)+(ORDER*UTTERANCE*MORPHOLOGY|ITEM)

Reference levels of predictors are identical to Model 4.

Parameter thresholds	Estimate	Std. dev.	95% Credible interval	Posterior probability P(β<0|β>0)
β_{Mismatch}	−3.13	0.52	[−4.12;−2.05]	1*
β_{Anaphora}	4.00	0.40	[3.18;4.75]	1*
β_{Ellipsis}	4.05	0.44	[3.14;4.89]	1*
$\beta_{\text{-eur/-euse}}$	−0.10	0.41	[−0.88;0.73]	0.6
$\beta_{\text{-eur/-rice}}$	0.30	0.34	[−0.36;0.96]	0.82
$\beta_{\text{Mismatch:Anaphora}}$	0.32	0.53	[−0.71;1.35]	0.73
$\beta_{\text{Mismatch:Ellipsis}}$	−0.25	0.48	[−1.19;0.7]	0.7
$\beta_{\text{Mismatch:-eur/-euse}}$	−0.43	0.58	[−1.52;0.73]	0.78
$\beta_{\text{Mismatch:-eur/-rice}}$	−0.55	0.52	[−1.55;0.48]	0.86
$\beta_{\text{Anaphora:-eur/-euse}}$	−0.17	0.56	[−1.3;0.89]	0.6
$\beta_{\text{Ellipsis:-eur/-euse}}$	0.03	0.63	[−1.28;1.23]	0.53
$\beta_{\text{Anaphora:-eur/-rice}}$	0.13	0.51	[−0.89;1.11]	0.62
$\beta_{\text{Ellipsis:-eur/-rice}}$	0.16	0.57	[−0.98;1.29]	0.62
$\beta_{\text{Mismatch:Anaphora:-eur/-euse}}$	0.23	0.76	[−1.29;1.74]	0.62
$\beta_{\text{Mismatch:Anaphora:-eur/-euse}}$	0.20	0.75	[−1.25;1.72]	0.6
$\beta_{\text{Mismatch:Anaphora:-eur/-rice}}$	−0.27	0.69	[−1.58;1.12]	0.65
$\beta_{\text{Mismatch:Ellipsis:-eur/-rice}}$	0.28	0.71	[−1.11;1.71]	0.66

References

Abeillé, Anne, Aixiu An & Aoi Shiraïshi. 2018. L'accord de proximité du déterminant en français. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics* 22. 1–25.

Abeillé, Anne, Gabriela Bîlbîie & François Mouret. 2014. A Romance perspective on gapping constructions. In Boas, Hans C & Gonzálvez-García, Francisco, *Constructional approaches to language* 15, 227–265. John Benjamins Publishing Company.

Abeillé, Anne & Jong-Bok Kim. 2022. 'Me too' fragments in English and French: a direct interpretation approach. *The Linguistic Review* 39(3). 495–524.

Amsili, Pascal, Ellsiepen, Emilia & Grégoire Winterstein. 2016. Optionality in the use of 'too': The role of reduction and similarity. *Revista da ABRALIN* 15(1).

Aparicio, Helena, Franich, Kathryn & Ming Xiang. 2015. Morphological feature mismatches under ellipsis: An eye-tracking study. *Proceedings of NELS* 45.

Benzitoun, Christophe, Jeanne-Marie Debaisieux & Henri-José Delofeu. 2016. The ORFEO project: A study corpus for contemporary French. *Corpus* 15. 1–18.

Bîlbîie, Gabriela. 2021. L'ellipse dans les comparatives: le cas du gapping en roumain. *Quand le syntagme nominal prend ses marques: du prédicat à l'argument*, 475–497. Éditions et presses universitaires de Reims.

Bîlbîie, Gabriela. 2023. Gender mismatches on adjectival predicates: An experimental perspective on stripping-like constructions in Romanian. *ECBAE 2023*. Amherst.

Bobaljik, Jonathan & Cynthia Zocca (2011). Gender markedness: The anatomy of a counterexample. *Morphology* 21(2). 141–166.

Bonami, Olivier & Gilles Boyé. 2019. Paradigm uniformity and the French gender system. In Baerman, Matthew et al. (eds.) *Morphological Perspectives*, 169–192. Edinburgh University Press.

Chung, Sandra. 2006. Sluicing and the lexicon: The point of no return. *Proceedings of the annual meeting of the Berkeley Linguistics Society* 31. 73–91.

Corbett, Greville. 1991. *Gender*. Cambridge University Press.

Corbett, Greville. 2023. The Agreement Hierarchy and (generalized) semantic agreement. *Glossa: A Journal of General Linguistics* 8(1).

Culicover, Peter & Ray Jackendoff. 2005. *Simpler syntax*. Oxford University Press.

Da Cunha, Yanis. 2024. *Le genre comme trait de proéminence en syntaxe*. Université Paris Cité PhD thesis.

Fiengo, Robert & Robert May. 1994. *Indices and identity*. MIT Press.

Frazier, Lyn, Lori Taft, Tom Roeper, Charles Clifton & Kate Ehrlich. 1984. Parallel structure: A source of facilitation in sentence comprehension. *Memory and Cognition* 12.

Gandón-Chapela, Evelyn & Francisco Gallardo-del-Puerto. 2025. The effect of gender (mis)match on the interpretation of English reflexive pronouns in cases of VP-ellipsis by native and non-native speakers of English. *Glossa: A Journal of General Linguistics* 10(1). 1–33.

Ginzburg, Jonathan & Ivan Sag. 2000. *Interrogative investigations*. Stanford: CSLI publications.

Hankamer, Jorge & Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7(3). 391–428.

Haspelmath, Martin. 2002. On understanding word order asymmetries. (Comments on John A. Hawkins, "Symmetries and asymmetries: their grammar, typology and parsing"). *Theoretical Linguistics* 28. 159–170.

Haspelmath, Martin. 2020. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174.

Hough, Julian, Ye Tian, Laura de Ruiter, Simon Betz, Spyros Kousidis, David Schlangen & Jonathan Ginzburg. 2016. Duel: A multi-lingual multimodal dialogue corpus for disfluency, exclamations and laughter. *Proceedings of LREC'16.* 1784–1788.

Kaplan, David. 1989. Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals.

Merchant, Jason. 1999. *The syntax of silence: Sluicing, islands, and identity in ellipsis.* UCSC.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis.* Oxford University Press.

Merchant, Jason. 2014. Gender mismatches under nominal ellipsis. *Lingua. Structural approaches to ellipsis* 151. 9–32.

Mickus, Timothee, Olivier Bonami & Denis Paperno. 2019. Distributional effects of gender contrasts across categories. *Proceedings of SCL* 2. 174–184.

Miller, Phillip & Geoffrey K. Pullum. 2014. Exophoric VP ellipsis. *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag 5.* 32.

Misersky, Julia, Pascal M. Gygax, Ute Gabriel, Alan Garnham, Friederike Braun, Tania Chiarini, Kjellrun Englund, Adriana Hanulikova, Anton Öttl, Jana Valdrova, Lisa Von Stockhausen & Sabine Sczesny. 2014. Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior research methods* 46. 841–871.

Noailly, Michèle. 2021. L'adjectif et le syntagme adjectival. In Abeillé, Anne & Godard, Danièle (eds.), *La Grande Grammaire du français.* Arles: Actes Sud.

Nunes, Jairo. & Zocca, C. 2009. Lack of morphological identity and ellipsis resolution in Brazilian Portuguese. In Nunes, Jairo (Ed.), *Minimalist essays on Brazilian Portuguese syntax*, 142, 215–236. John Benjamins Publishing Company.

Nykiel, Joanna & Jong-Bok Kim. 2021. Ellipsis. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *HPSG: The handbook*, 847–888. Berlin: Language Science Press.

Parker, D. 2019. Cue combinatorics in memory retrieval for anaphora. *Cognitive science* 43(3). e12715.

Poppels, Till. 2022. Explaining ellipsis without identity. *The Linguistic Review* 39(3). 341–400.

Richy, Célia & Heather Burnett. 2019. Jean does the dishes while Marie fixes the car: A qualitative and quantitative study of social gender in French syntax articles. *Journal of French Language Studies* 30(1). 47–72.

Ross, John R. et al. 1969. Guess who? *Proceedings from the annual meeting Chicago Linguistic Society* 5(1). 14–39.

Spencer, A. 2002. Gender as an inflectional category. *Journal of Linguistics* 38(2). 279–312.

Sprouse, Jon, Troy Messick & Jonathan D. Bobaljik. 2022. Gender asymmetries in ellipsis: An experimental comparison of markedness and frequency accounts in English. *Journal of Linguistics* 58(2). 345–379.

Sudo, Yasutada & Giorgos Spathas. 2020. Gender and interpretation in Greek: Comments on Merchant (2014). *Glossa: A Journal of General Linguistics* 5(1).

Zwanenburg, Wiecher. 1988. 'Flexion et derivation: le féminin en français.' Ronald Landheer (ed.), *Aspects de linguistique française, Hommage à Q.I.*M Mok, 191–208. Amsterdam: Rodopi.