Hassen, Amal & Abeillé, Anne. 2025. Preposition omission in French interrogative sluices: empirical findings and theoretical implications. *Glossa: a journal of general linguistics* 10(1). pp. 1–47. DOI: https://doi.org/10.16995/glossa.15197

Preposition omission in French interrogative sluices: empirical findings and theoretical implications

Amal Hassen, Laboratoire de linguistique formelle, Université Paris Cité, FR, amal.hassen@u-paris.fr

Anne Abeillé, Laboratoire de linguistique formelle, Université Paris Cité, FR, anne.abeille@u-paris.fr

Many non-P(reposition)-stranding languages have been shown to allow P-omission under sluicing (Sag & Nykiel 2011; Nykiel 2013; Molimpakis 2019), thus challenging Merchant's (2001) generalization, which derives P-omission in sluices from P-stranding. French is still an open issue: while Merchant (2001) claims it does not allow for P-omission, Rodrigues et al. (2009) provide an example, with a cleft-based derivation. In two experimental studies we show that P-omission in French sluices is acceptable and is sensitive to the remnant type, with both *à* 'to' and *de* 'of'. Through the analysis of a large (written) corpus, we also show that P-omission in French sluices is the preferred option, as in English (Nykiel 2017), and identify the factors favoring it. Our findings confirm the effect of cognitive and information-based cues on P-omission under sluicing (Nykiel & Hawkins 2020) and suggest a non-deletion fragment-based analysis (Ginzburg & Sag 2000; Sag & Nykiel 2011).

1 Introduction

P(reposition)-omission under sluicing is an important point for theories of ellipsis. Ross (1969), who coined the term *sluicing*, derives elliptical interrogatives from verbal ones by a deletion operation. Levin (1982) was the first to point out that the preposition can be omitted in a sluice, unlike a verbal interrogative (1a).

(1) a. I don't remember *(to) who she was talking.
b. Janet was talking to someone, but I don't remember (to) who. (Levin 1982: 592)

This highlights a major mismatch between sluices and verbal interrogatives, which may cast doubts on syntactic theories deriving the former from the latter (cf. Ross 1969). Other mismatches have been noticed in the literature, such as verb form mismatch (Merchant 2001) and person mismatch (Ross 1969).

(2) a. Decorating for the holidays is easy when you know how (to decorate).
(Merchant 2001: 22)
b. Bob knows how to crane his neck but I don't know how (to crane my neck).
(Ross 1969: 33)

1.1 Two structural analyses of P-omission under sluicing

Merchant (2001) derives sluices from verbal counterparts and P-omission under sluicing from P-stranding in regular *wh*-questions. Thus English, which has two options for wh-questions, allows for sluices with (3a) and without (3b) a preposition.

(3) a. Peter was talking to someone, but I don't know to whom ~~he was talking~~.
b. Peter was talking to someone, but I don't know who ~~he was talking to~~.

He claims that non-P-stranding languages ban P-omission under sluicing, which he calls the Preposition Stranding Generalization. However, many non-P-stranding languages have been shown to allow for it, including Indonesian (Fortin 2007; Sato 2011), Brazilian Portuguese (Almeida & Yoshida 2007), Spanish (Rodrigues et al. 2009), Polish (Sag & Nykiel 2011; Nykiel 2013), Russian (Philippova 2014), Bulgarian (Abels 2017), Greek and German (Molimpakis 2019), Mauritian creole (Abeillé & Hassamal 2019) and Saudi Arabic (Alshaalan & Abels 2020).

Following Szczegielniak (2008) and van Craenenbroeck (2004), Rodrigues et al. suggest that, in non-P-stranding languages, the possibility of P-omission in sluices comes from the deletion of a cleft structure, with the wh-remnant functioning as the pivot of the cleft, as in Spanish (4).

(4) *Spanish* (Rodrigues et al. 2009: 178)
Juan ha hablado con una chica pero no sé cuál (es la chica con la
Juan has talked with a girl but not know.1SG which (is the girl with which
que ha hablado Juan).
that has talked Juan)
'Juan has talked with a girl but I don't know which (is the girl with which John has talked).'

Since clefts are used to focus an element in a set, they suggest that P-omission improves when the remnant belongs to such a set, as is the case with a nominal correlate (*una chica* 'a girl') and an anaphoric remnant (*cuál* 'which') (see (4)), compared to a pronominal correlate (*alguien* 'someone') and a non-anaphoric remnant (*quién* 'who') as in (5).

(5) *Spanish* (Rodrigues et al. 2009: 176)
??Juan ha hablado con alguien, pero no sé quién.
Juan has talked with someone, but not know.1SG who
'Juan has talked with someone but I don't know who.'

1.2 Alternative analyses of P-omission under sluicing

Under direct interpretation approaches to ellipsis (Ginzburg & Sag 2000; Culicover & Jackendoff 2005; Sag & Nykiel 2011), an interrogative sluice is not derived from a full interrogative counterpart: a wh-fragment recovers its propositional meaning from an antecedent available in the context, and it inherits its syntactic and semantic properties from a contextual correlate.

This is why, when the correlate is an animate NP, the remnant must be an animate NP, not a PP (6).

(6) Paul has seen [a friend]$_{NP}$. I wonder [who]$_{NP}$ / *[what]$_{NP}$ / *[to who]$_{PP}$.

If the antecedent clause includes a PP, two correlates are in fact available; the whole PP or the NP inside it (7), hence two possible corresponding remnants (an NP or a PP), and the possibility of P-omission (7b).

(7) a. Paul has talked [to a friend $_{PP}$]. I wonder [[to who]$_{PP}$] $_{S}$.
b. Paul has talked to [a friend $_{NP}$]. I wonder[[who] $_{NP}$] $_{S}$.

Under such theories, any language may allow P-omission in sluices, and non-syntactic factors may play a role. Nykiel (2013), working on Polish, a non-P-stranding language, found that P-omission is acceptable in sluices, while the cleft source proposed by Szczegielniak (2008) was as degraded as ungrammatical controls. She also found that P-omission is favored by nominal correlates and remnants (8a), while pronominal correlates (8b) favor PP remnants.

(8) *Polish* (Nykiel 2013: 76)

a. Anna odpowiedziała na jakieś pytanie, ale nie pamiętam (na) które.
Anna answered PREP some question.ACC but not remember.1SG (PREP) which.ACC
'Anna answered to some question but I don't remember which one.'

b. Anna odpowiedziała na coś, ale nie pamiętam (na) co.
Anna answered PREP something.ACC but not remember.1SG (PREP) what.ACC
'Anna answered to something but I don't remember what.'

Based on these results, Nykiel suggested a processing-based theory. Interpreting a sluice comes with a processing cost since an antecedent has to be found. As in anaphora processing, some antecedents are more accessible than others, and some cues are better than others (Lewis & Vasishth 2005; Martin & McElree 2011; Parker 2019). Nominal correlates are more accessible than pronominal ones due to their richer semantic and morphosyntactic content (Hofmeister et al. 2007).

Prepositions also act as cues to retrieve the PP correlate (preventing *who* to be misinterpreted as referring to Paul in (7b), for instance), hence P-omission in sluices comes with a processing cost (Nykiel & Hawkins 2020). Philippova (2014) found that preposition length favors P-omission in Russian, longer prepositions offering a better cue for ellipsis resolution, and Nykiel & Kim (2022b) found a similar effect for Polish.

According to Nykiel & Hawkins (2020)'s theory, two general principles are at play: Minimize Forms (MiF) (Hawkins 2004), favoring ellipsis and P-omission, and Minimize Domains (MiD), which may disfavor P-omission. According to MiF, forms can be reduced (e.g., via ellipsis) when the necessary information can be easily retrieved from the context, with full noun phrases being more accessible than pronouns. For sluices, MiF predicts that P-omission is more likely to occur with nominal correlates, which are easily accessible, and also with nominal remnants, which are better cues than pronominal ones.

According to MiD, on the other hand, processing efficiency increases when semantically and syntactically dependent elements are processed in the smallest possible processing domain. In a PP sluice, P is typically processed within the PP domain when P is independent from the verb as in (9a). However, if the P is dependent on the verb as in (9b), it is interpreted in the VP domain. This domain extension creates an additional processing cost for sluices as P in the remnant must *go back* to re-establish its dependency relation with the V for interpretation. Therefore, MiD predicts that V-dependent Ps are more easily omitted in a sluice.

(9) a. Peter sat on the chair.

b. Peter is waiting for a friend.

In such a theory, the acceptability of P-omission under sluicing is flexible, and favored by contentful remnants and correlates, and also by more dependent prepositions.

1.3 The case of French

As noted by Dagnac (2018: 8), "French sluicing has been poorly studied so far". According to Merchant (2001), P-omission is not possible in French sluices (10a), and (10b) is the only option. According to him, this is explained by the impossibility to strand the preposition as shown in (10c).

(10) *French* (Merchant 2001: 98)

a. *Anne l'a offert à quelqu'un, mais je ne sais pas qui.
Anne it-has offered to someone, but I NEG know NEG who.
'Anne has offered it to someone but I do not know who.'

b. Anne l'a offert à quelqu'un, mais je ne sais pas à qui (elle l'a offert).
Anne it-has offered to someone, but I NEG know NEG to whom.
'Anne has offered it to someone but I do not know who.'

c. *Qui est-ce qu'elle l'a offert à?
*who is-it that-she it-has offered to?
'Who did she offer it to?'

Admittedly, Merchant only asked three French speakers (linguist colleagues), who also consider P-omission as more acceptable with the preposition *avec* 'with' as in (11).

(11) Elle a parlé avec quelqu'un, mais je ne sais pas (avec) qui.
she has talked with someone but I NEG know NEG (with) who.
'She talked with someone but I don't know (with) who.' (Merchant, 2001: 98 fn7)

One may thus hypothesize that the badness of (10a) comes from *à* being a weak non-stand-alone preposition, like *de* 'of' (Abeillé & Kupfermann 2021), while *avec* can stand alone as shown in example (12).

(12) Les allumettes, il ne faut pas jouer avec.
the matches, it NEG have NEG play with
'Matches, you shouldn't play with.'

Rodrigues et al. (2009) suggest that P-omission comes from a cleft source in French, as in Spanish. They assign a question mark to (13a) and consider P-omission to be ameliorated with a nominal correlate (*une des filles* 'one of the girls'), which defines a set, and an anaphoric remnant (*laquelle* 'which.F.SG') (13b).

(13) a. ?Jean a dansé avec quelqu'un, mais je ne sais pas qui.
Jean has danced with someone but I NEG know NEG who
'Jean has danced with someone, but I don't know who.'

b. Jean a dansé avec une des filles, mais je ne sais pas laquelle
Jean has danced with one of.the girls but I NEG know NEG which.F.SG
(c'était la fille avec qui il a dansé).
it.was the girl with who he has danced.
'Jean danced with one of the girls but I do not know which one (it was that he danced with).' (Rodrigues et al. 2009: 198)

They also consider that French clefts have a "weak exhaustivity requirement", like Brazilian Portuguese and unlike Spanish (Destruel 2012), so that P-omission should be possible with an additive sluice (14):

(14) Jean a dansé avec une des filles, mais je ne sais pas (avec) qui d'autre.
Jean has danced with one of.the girls but I NEG know NEG (with) who of-else
'Jean has danced with one of the girls but I don't know who else.'
(Rodrigues et al. 2009: 198)

They thus suggest that French sluices come in two varieties: with a cleft source when P-omission is involved, and with regular wh-fronting in other cases[1]. They do not provide examples of P-omission with other prepositions than *avec*.

Dagnac (2018) finds both (10a) and (13a) 'perfectly acceptable', and Gotowski (2022) reports that "some speakers from France find (10a) perfectly acceptable" (p.13). Following Abeillé & Hassamal (2019), they both challenge the cleft analysis for *quoi* 'what' sluices, since this pronoun cannot be fronted (Obenauer 1976; Abeillé & Godard 2021):

(15) Jean mange quelque chose, mais je ne sais pas quoi (*c'est).
Jean eat.3SG some thing but I NEG know.1SG not what (it.is)
'Jean is eating something, but I do not know what (it is).' (Gotowski 2022: 14)

As noted by Obenauer (1976) and Abeillé & Godard (2021), it is not possible to extract bare *quoi*, which must be in situ, unlike the weak form *que* 'what', which must be extracted as in *Qu'est-ce?* 'What is it?' (Abeillé & Godard 2021).[2]

[1] Alternative structural analyses have been proposed with an in-situ source (Chung et al. 1995; Abe 2015). While in-situ wh-questions are quite common in spoken French (Delaveau 2021), they cannot contribute to a theory of P-omission, since the preposition is obligatory in-situ: *Tu as parlé à qui? /* qui?* ('you have talked to who /*who?). See Gotowski (2022) for other arguments against deriving French sluices from in-situ questions.

[2] *Quoi* can also be part of an extracted prepositional phrase in a verbal interrogative (*À quoi penses-tu ?* 'about what are you thinking?') but this is not relevant for P-omission.

(16) a. *Quoi c'est?
*what it-is
'What is it?'
b. C'est quoi?
It-is what
'What is it?'

Dagnac (2018) provides an example of P-omission with *quoi* (17a), and so do Abeillé & Hassamal (2019) (17b).

(17) a. On pourrait le remplir avec quelque chose, mais (je ne sais pas) quoi.
one could it fill with some thing but (I NEG know NEG) what
'One could fill it with something, but (I don't know) what.' (Dagnac 2018)
b. Paul travaille sur quelque chose, je ne sais pas (sur) quoi.
Paul works on some thing, I NEG know NEG (on) what.
'Paul is working on something, I don't know what' (Abeillé & Hassamal 2019: 7)

While Gotowski does not explicitly consider such cases, her movement analysis (+PF insertion of *quoi*) implies that P-omission should not be possible in a *quoi* sluice.

In a corpus of contemporary texts (Frantext after 1980), Smirnova & Abeillé (2021) found two instances of P-omission with *qui ça* ('who that') and *de* ('of'), which is a weak preposition like *à* (18). However, these examples are reprise questions (i.e. clarification requests) with a discourse particle *ça* 'that', which may be more flexible than genuine information seeking questions.

(18) a. L'adresse de Rosine Portinari, tu l'as pas ? – Qui ça? (Thérame 1985)
the-address of Rosine Portinari you it-have NEG ? - who that?
'You don't have the address of Rosine Portinari? – Who?'
b. La ville de Jaufré Rudel ! - Qui ça? (Garat 1984)
the city of Jaufré Rudel ! - who that?
'The city of Jaufré Rudel - Who?' (cited by Smirnova & Abeillé 2021: 247)

Direct interpretation approaches (Ginzburg & Sag 2000; Culicover & Jackendoff 2005; Sag & Nykiel 2011) predict that P-omission should be possible in French sluices. Nykiel's cognitive hypothesis, along with the cue-based retrieval approach, also makes predictions about French. A remnant with a preposition carries more information about the correlate than without, and a nominal remnant carries more information than a pronominal one. Hence, for ellipsis resolution, a preposition is less required with a nominal remnant and more so with a pronominal remnant. French sluices can thus be thought of as following the information hierarchy in **Figure 1**, with more information on the right favoring P-omission : a nominal correlate (e.g. *a friend*) is

more informative than a pronominal one (e.g. *someone*) (hence more P-omission), itself more informative than an implicit correlate. *Quel* ('which') + Noun being more contentful than *lequel* ('which.M.SG'), itself more contentful than *qui/quoi* ('who/what'), would also trigger more P-omission.

	Information hierarchy		
	-	----------	+
Remnant	*qui / quoi*	---------- *lequel* ----------	*quel* + noun
Correlate	implicit	---------- prep + pronoun ----------	prep + noun
Preposition	*à, de*	----------	Other

Figure 1: Information Hierarchy for French sluices.

Similarly, longer and more contentful prepositions may offer a better cue for ellipsis resolution and favor P-omission, compared to monosyllabic ones like *à* 'to' and *de* 'of' (Nykiel & Kim 2022b). On the other hand, taking into account Verb-Prep dependency, Nykiel & Hawkins's (2020) MiD-based theory predicts that more dependent, less informative prepositions (like *à* 'to' and *de* 'of') have more chance to be omitted than more independent, more contentful ones, such as *avec* 'with' or *pour* 'for'.

In this paper, we present two experiments testing the acceptability of P-omission in French interrogative sluices and a (written) corpus study, to examine whether P-omission naturally occurs in French, and which factors have an effect. We also look at the plausibility of a cleft-source explanation for our results. Subsequently, we offer a direct interpretation analysis of French interrogative sluices, based on Ginzburg & Sag (2000), Sag & Nykiel (2011), Ginzburg (2012), Ginzburg & Miller (2018) and Nykiel & Kim (2022a).

2 Two experiments on *à* 'to' and *de* 'of' sluices

We ran two experiments to test whether sluices with P-omission are acceptable compared to verbal interrogatives. We used prepositions *à* 'to' and *de* 'of', because they are the most frequent in French, and because Merchant (2001) claims that P-omission is not possible with *à*.

In line with Rodrigues et al. (2009) and Nykiel (2013)'s hypotheses, we compare pronominal remnants (Experiment 1) and nominal remnants (Experiment 2).

2.1 Experiment 1: *qui* sluices

The first experiment tests French interrrogatives with the pronoun *qui* ('who'). 40 native speakers of French were recruited via Prolific and were paid 1.8GP for a 15-minute experiment. Among the information we collected from the participants, the only selection criterion was being an adult (monolingual) native speaker from France.

2.1.1 Experimental items

The materials for Experiment 1 included 20 items, 10 with the preposition *à* 'to' (19) and 10 with *de* 'of' (20). Both prepositions introduced a complement selected by the verb (*écrire à, parler à, parler de, se souvenir de* 'write to, talk to/about, remember') (see Appendix 1 for the full list of items). The experimental items followed a 2x2 design, crossing two independent variables to create four conditions, as shown in **Tables 1** and **2**.

(19) Speaker A:
J'ai parlé à un ami
I-have talked to a friend
'I have talked to a friend'
Speaker B:

+ PREP + ELLIPSIS	– PREP + ELLIPSIS	+ PREP – ELLIPSIS	– PREP – ELLIPSIS
À qui? 'To whom?'	Qui? 'Who?'	À qui as-tu parlé? 'To whom have you talked?'	Qui as-tu parlé? 'Whom have you talked?'

Table 1: Conditions of experiment 1: *à*.

(20) Speaker A:
J'ai eu des nouvelles d'un collègue.
I-have had INDF.PL news of-a colleague.
'I've had news about a colleague'
Speaker B:

+ PREP + ELLIPSIS	– PREP + ELLIPSIS	+ PREP – ELLIPSIS	– PREP – ELLIPSIS
De qui? 'Of whom?'	Qui? 'Who?'	De qui as-tu eu des nouvelles? 'Of whom have you had news?'	Qui as-tu eu des nouvelles? 'Whom have you had news?'

Table 2: Conditions of Experiment 1: *de*.

Our materials also contained 20 (dialog) filler items (21), coming from an unrelated experiment on fragment answers in French.

(21) A: Il arrive quand ? B: (C'est) demain.
A: he arrives when ? B: (it-is) tomorrow
A: 'When does he arrive?' B: (It's) tomorrow.'

2.1.2 Procedure

The experiment was an acceptability judgment task administered online on Ibex Farm created by Alex Drummond. Participants completed a short questionnaire before starting to read the instructions. They went through several practice trials to ensure full comprehension of the instructions. Items were automatically counterbalanced through 4 lists of participants to ensure that each of them would not encounter more than one condition of the same item. The participants read a short dialog between speaker A and speaker B and were asked to judge the acceptability of B's utterance on a 1 to 5 scale, with 1 being not acceptable at all and 5 fully acceptable.

2.1.3 Hypotheses

According to Merchant's (2001) theory, the lack of a preposition, regardless of ellipsis, makes the sentence ungrammatical and should elicit the lowest acceptability ratings. In this sense, he predicts a main effect of the +Prep condition. In contrast, the direct interpretation approach (Ginzburg & Sag 2000) predicts an interaction: –Prep + Ellipsis sentences should receive higher ratings than their –Prep –Ellipsis counterparts.

2.2 Results of Experiment 1

The results in **Figure 2**[3] indicate that the –Prep –Ellipsis condition elicited the lowest acceptability ratings (mean = 2.07), confirming its ungrammaticality (e.g. **Qui as-tu parlé?* '*Who did you talk?'). In contrast, the –Prep + Ellipsis condition received a much higher rating (mean = 3.12). The +Prep conditions do not show any difference between –Ellipsis (mean = 3.88) and +Ellipsis (mean = 3.86) conditions.

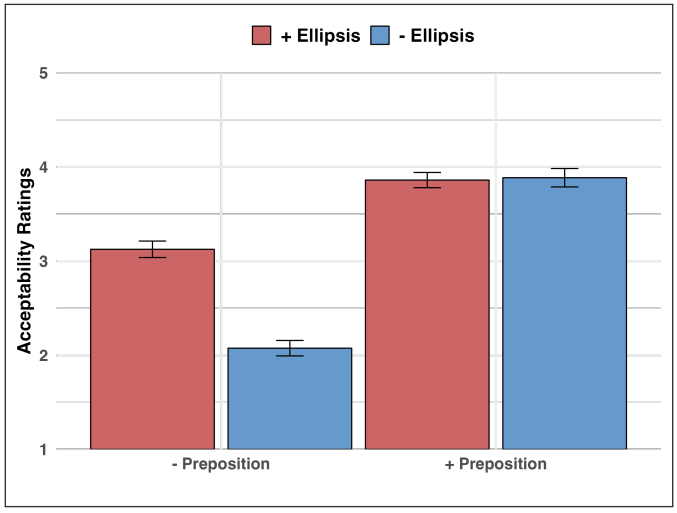

Figure 2: Experiment 1: Acceptability results: Remnant *qui*.

[3] The error bars on the graphs represent the standard error of the mean.

The data were analyzed with a Bayesian ordinal cumulative mixed-effects model with a probit link function (Carpenter et al. 2017) using the brms package (Bürkner 2017) for Bayesian regression models in R[4]. The factors Preposition and Ellipsis were dummy-coded, with +Prep and –Ellipsis serving as reference levels, respectively. Participants and items were included as random factors. Eight sampling chains were run, each consisting of 9000 iterations with a warm-up period of 4500 iterations. We report the parameters' expected values from the posterior distribution and their corresponding 95% credible intervals (CrIs) for the mean values and differences between them.[5] We adopted the default priors provided by the brms package for the analysis. The analysis showed a strong evidence for an effect of Prep ($\widehat{\beta} = -2.53$, 95% CrI = [–3.12, –1.98], $P(\widehat{\beta} < 0) = 1$), suggesting an overall preference for +Prep, and a weak evidence for an overall preference for –Ellipsis ($\widehat{\beta} = -0.24$, 95% CrI = [–0.73, 0.25], $P(\widehat{\beta} < 0) = 0.83$). It also presents strong evidence in support of an interaction of Preposition and Ellipsis ($\widehat{\beta} = 1.67$, 95% CrI = [1.14, 2.22], $P(\widehat{\beta} > 0) = 1$), where P-omission is more acceptable under +Ellipsis, compared to –Ellipsis items. See Appendix 1 for the model.

We also looked at possible differences between the ratings of items with *à* 'to' and items with *de* 'of'. Results (see **Figure 3**) show that in the –Prep+Ellipsis condition, the omission of *de* was rated higher (Mean = 3.39) than that of *à* (Mean = 2.85).

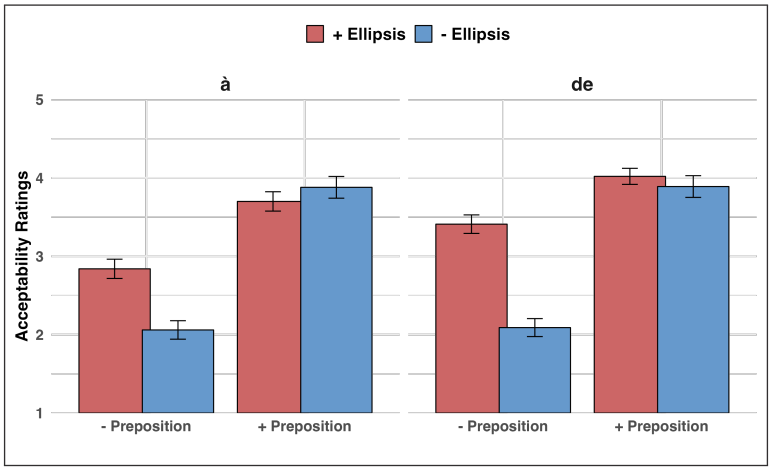

Figure 3: Experiment 1: Acceptability results: à vs. de.

The results of the Bayesian ordinal model (with *à* as the reference level) show an estimated probability of 0.56 with a 95% CrI of [–0.40,0.46]. This suggests no evidence for a difference

[4] For the Bayesian models in this paper, we checked the chains' convergence using Geweke (1992)'s diagnostics and the R-hat statistics for each parameter coefficient. Model fit and assumptions were assessed via pp_check (Gelman et al. 2013).

[5] When $P \geq 0.9$ and the CrIs do not include zero, we interpret it as strong evidence for an effect. If $0.8 \leq P < 0.9$, we consider the evidence weak. Probabilities below 0.8 are interpreted as no evidence for an effect.

between *à* and *de*. The model, on the other hand, showed a weak evidence for a preference for *de* under the +Ellipsis condition, compared to *à* ($\widehat{\beta}$ = 0.27, 95% CrI = [–0.15, 0.71], P($\widehat{\beta} > 0$) = 0.89), suggesting that the omission of *de* is more acceptable than the omission of *à* in sluices.[6] This is not predicted by any theory claiming that P-omission is ungrammatical in French sluicing.

2.3 Experiment 2: *quel*+ noun

In order to test Rodrigues et al. (2009) and Nykiel (2017)'s hypotheses, we conducted a second experiment with a nominal remnant. The design was the same as the first one. We recruited 48 French native speakers using Prolific and each participant received a compensation of 1.8GP.[7]

2.3.1 Experimental items

The materials for Experiment 2 included 20 items, 10 with *à* 'to' (22) and 10 with *de* 'of' (23), adapted from Experiment 1 (See Appendix 2 for the full list of materials). Similarly to the first experiment, we tested 4 conditions (see **Tables 3** and **4**), but instead of *qui* ('who'), we used *quel* ('which') + noun.[8]

(22) Speaker A:
J'ai écrit à un vieil oncle.
I-have written to an old uncle
'I wrote to an old uncle.'
Speaker B:

+ Prep + Ellipsis	– Prep + Ellipsis	+ Prep – Ellipsis	– Prep – Ellipsis
À quel oncle? 'To which uncle?'	Quel oncle? 'Which uncle?'	À quel oncle as-tu écrit? 'To which uncle have you written?'	Quel oncle as-tu écrit? 'Which uncle have you written?'

Table 3: Conditions of experiment 2: *à*.

[6] As a possible explanation for this (potential) difference, *de* is more frequent than *à* in French: 43% vs. 16.2% of Ps uses (Tseng 2021), which may make *de* a better memory cue. We also checked the dependency between the two prepositions and their predicates in the experimental items, following Nykiel & Hawkins (2020)'s dependency coding: *de* showed slightly more dependency (60% at Level 1 and 40% at Level 2, compared to *à*'s 50% and 30%, see section 3.3.5). Note, however, that we used the same P-Predicate pairs in the items of Experiment 2, and no difference was found between *à* and *de* (see section 2.4).

[7] Data from one participant were excluded since they encountered two lists of the experiment.

[8] For this experiment, we also added simple yes/no comprehension questions after each item, and we kept participants with a minimum accuracy score of 75%. The question after (23) for instance was : *Est-ce-que la personne parle de son patron?* 'Is the person talking about their boss?'.

(23) Speaker A:
Je me souviens d'un ancien voisin.
I REFL remember of-a old neighbor
'I remember an old neighbor'.
Speaker B:

+ PREP + ELLIPSIS	– PREP + ELLIPSIS	+ PREP – ELLIPSIS	– PREP – ELLIPSIS
De quel voisin? 'Of which neighbor?'	Quel voisin? 'Which neighbor?'	De quel voisin te souviens-tu? Lit. 'Of which neighbor do you remember?'	Quel voisin te souviens-tu ? 'Which neighbor do you remember?'

Table 4: Conditions of experiment 2: *de*.

We used 20 filler items, coming from an unrelated experiment on *si* '*if*' conditionals (24).

(24) Speaker A: Léa a confié ses bijoux à son frère.
'Léa entrusted her jewelry to her brother.'
Speaker B: Serait-elle déçue s'il perdait une bague?
'Would she be upset if he lost a ring?'

2.3.2 Hypotheses

As in Experiment 1, Merchant's theory predicts that P-omission, regardless of ellipsis, should receive the lowest ratings, while the direct interpretation approach (Ginzburg & Sag 2000; Culicover & Jackendoff 2005; Nykiel 2013) predicts that elliptical sentences without prepositions would receive higher ratings than their non-elliptical counterparts. In addition, Rodrigues et al. (2009)'s and Nykiel's theory predicts that since the remnant is more informative than in Experiment 1, the –Prep + Ellipsis condition should be rated higher than in Experiment 1.

2.4 Results of Experiment 2

The results (**Figure 4**) show that the –Prep–Ellipsis condition received the lowest ratings (mean = 2.16). Unlike in Experiment 1, the ratings of + Prep + Ellipsis (mean = 4.72) are now higher than those of + Prep–Ellipsis (mean = 4.55) but the difference is quite small. Additionally, the –Prep + Ellipsis sentences now have acceptability ratings (mean = 4.47) comparable to those of the + Prep–Ellipsis sentences.

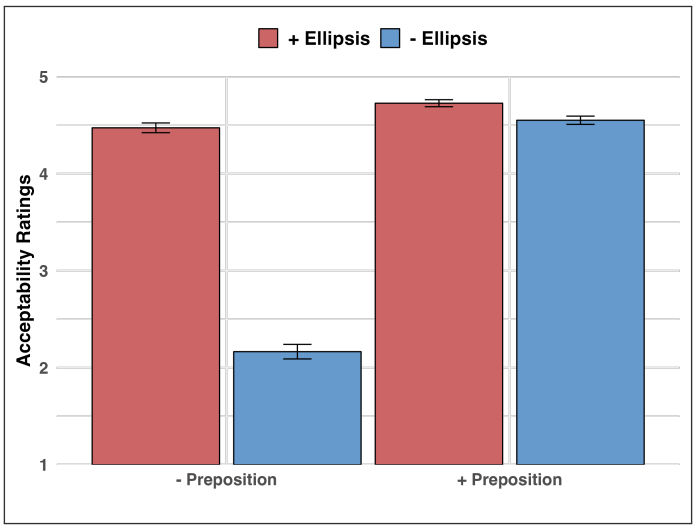

Figure 4: Experiment 2: Acceptability results: Remnant *quel* + noun.

As in Experiment 1, the data were fitted to a Bayesian ordinal cumulative mixed-effects model with a probit link function, with the same independent variables and the same random factors. +Prep and –Ellipsis were used as reference levels. Eight sampling chains were run, each consisting of 9000 iterations with a warm-up period of 4500 iterations. Results show a strong evidence for an overall preference of the +Prep condition ($\widehat{\beta}$ = –2.83, 95% CrI = [–3.44, –2.24], P($\widehat{\beta} < 0$) = 1), as in Experiment 1. They also show that +Ellipsis is preferred ($\widehat{\beta}$ = 0.89, 95% CrI = [0.39, 1.46], P($\widehat{\beta} > 0$) = 0.99), unlike in Experiment 1. We also have strong evidence in support of an interaction, where P-omission is more acceptable in +Ellipsis than in –Ellipsis items ($\widehat{\beta}$ = 2.35, 95% CrI = [1.57, 3.15], P($\widehat{\beta} > 0$) = 1). The estimate of the interaction in Experiment 2 ($\widehat{\beta}$ = 2.35) is higher than that in Experiment 1 ($\widehat{\beta}$ = 1.67), indicating a stronger effect of the interaction in Experiment 2 (See Appendix 2 for the model).

While it is not easy to compare two experiments, with different participants and different fillers, P-omission under ellipsis seems to be more acceptable in Experiment 2 (with a nominal remnant) than in Experiment 1 (with a pronominal remnant).[9]

We compared the ratings of the items with *à* and the items with *de* in this experiment (see **Figure 5**). The data were fitted to a Bayesian model and, similarly to experiment 1, results did not show an evidence for a difference between the two prepositions ($\widehat{\beta}$ = 0.15, 95% CrI = [–0.51, 0.81], P($\widehat{\beta} > 0$) = 0.69), and no interaction with Ellipsis, unlike experiment 1.

[9] The experimental items for both experiments were inspired from our corpus data (see section 3), this is why they included both definite (8 items) and indefinite (12 items) correlates. In both experiments, items with indefinite correlates were rated higher than with definite correlates (mean ratings of 3.34 and 3.09, respectively, in Exp.1; 4.02 and 3.91 in Exp.2) and this holds for the –Prep+Ellipsis condition (3.26 and 2.92 in Exp.1; 4.54 and 4.37 in Exp.2).

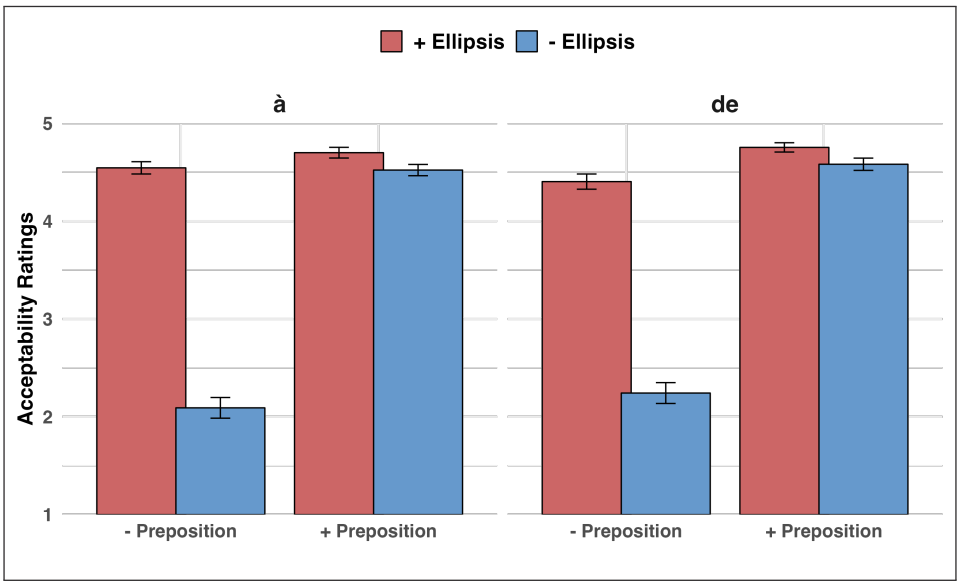

Figure 5: Experiment 2: Acceptability results: *à* vs. *de*.

2.5 Discussion of the experimental part

The main result of both experiments is that French does permit the omission of prepositions *à* and *de* under sluicing, albeit not to the same extent as when the preposition is included. The findings also highlight the distinction between sluicing and verbal interrogatives, as they show mandatory P-inclusion in the latter, in contrast to its optionality in the former. They thus contradict Merchant's generalization. Furthermore, the acceptability of P-omission in elliptical sentences was influenced by the informativeness of the remnant, as seen in the higher acceptability rate found with nominal (*quel* 'which' + noun) remnants, compared to pronominal (*qui* 'who') remnants. Our results thus support Nykiel's hypothesis that P-omission in sluicing is affected by processing-based factors.

The overall preference for Ellipsis was stronger in Experiment 2 than in Experiment 1, especially in the +Prep condition. We consider it to be due to the informativeness of the nominal remnant in Experiment 2 (see Nykiel 2017 for English).[10]

[10] Whether there is a general preference for ellipsis (due to a redundancy penalty) is a matter of debate and may vary across ellipsis types, languages, and genres. In two corpus studies, Lefeuvre (2020) found a 30% sluicing rate for French interrogatives in a spoken corpus (ESLO2), and Reinhardt (2019) a 17% sluicing rate for interrogatives in a detective novel corpus. As shown by Smirnova & Abeillé (2021), the preference for sluicing is sensitive to the presence of question particles, with a 67% sluicing rate for *wh*+*ça* (as in *qui ça* ? 'who that?') interrogatives in a written corpus. In contrast, Ginzburg & Kim (2023) found a high preference for sluices in English exclamatives (84.6% in three spoken corpora), a finding echoed in Hassen et al. (2023), who found a 84.7% sluice rate for French exclamatives. For gapping, on the other hand, Bîlbîie et al. (2022) found a preference for ellipsis in Romance languages but not in English.

3 Corpus study on P-omission in French sluicing

Our corpus study aims to investigate to which extent P-omission naturally occurs in French sluices and to identify the factors that facilitate such omission. According to Merchant (2001), P-omission should only occur with strong prepositions such as *avec* 'with' and not with weak prepositions like *à 'to'*. According to Rodrigues et al. (2009)'s cleft analysis, it should be facilitated by the presence of an anaphoric (D-linked) remnant, and should not occur with *quoi* 'what'. Direct interpretation analyses (Ginzburg & Sag 2000), on the other hand, suggest that P-omission should be possible with all prepositions and all wh-remnants. According to Nykiel's approach, it is more likely to occur when the correlates and remnants convey more informative content, and when the preposition is more dependent on the verb.

3.1 Corpus extraction

We extracted data from the contemporary sub-corpus of the Frantext database, which comprises over 1001 texts published after 1980 for approximately 62.1 million words.[11] We extracted sluices with pronominal remnants (*auquel, duquel, lequel, qui, quoi*, respectively 'to which', 'of which', 'which.M.SG', 'who', 'what'), using two search queries: "interrogative pronoun + ?", for matrix sluices, and "verb + (adv) + interrogative pronoun + punctuation", for embedded sluices. For sluices with nominal remnants, we used the queries: "*quel* ('which') + noun + ?", for matrix sluices, and "verb + (adv) + *quel* + noun + punctuation", for embedded sluices. We filtered out the occurrences without PP antecedents, and the ones that were not sluices (e.g., *n'importe qui/quoi/(le)quel* 'no matter who/what/which' + noun as indefinite NPs). We obtained a data set of 250 sluices, as shown in **Table (5)**.

Sluices	**Matrix**	**Embedded**	**Total**
(Prep) *qui*/*quoi*	32	40	72
(Prep) *quel* + noun	54	18	72
(Prep) *lequel*	81	25	106
Total	167	83	250

Table 5: Sluices with a PP correlate from Frantext subcorpus.

[11] We also searched a spoken corpus (Orfeo) and did not find a sufficient number of examples. Since Frantext data come from published and well edited books, this also reduces the risk of typos or dubious uses.

3.2 Corpus annotation

We manually annotated the following factors:

- the preposition of the correlate and of the remnant (if any)
- the type of preposition (weak/strong, stand-alone/not-stand-alone) (Merchant 2001)
- predicate-preposition dependency (Nykiel & Hawkins 2020)
- whether the correlate is implicit, nominal, or pronominal (Nykiel 2017)
- whether the sluice is direct or reprise (Ginzburg & Sag 2000)
- whether the remnant is pronominal or nominal and whether it is anaphoric or not (Nykiel 2017, Rodrigues et al. 2009)
- whether the sluice is root, coordinate, or embedded

Table 6 summarizes the sluices that we found in the Frantext subcorpus, distinguishing whether they are PPs or NPs, and whether they appeared as matrix (root or coordinate) or embedded phrases. Overall, we found more P-omission (62%) than P-inclusion (38%).

Sluices	**Matrix**	**Embedded**	**Total**
Prep + *qui*/*quoi*	31	35	66
Prep + *quel* + noun	16	9	25
Prep + *lequel*	3	1	4
subtotal (PP remnants)	50	45	95 (38%)
∅ *qui*/*quoi*	1	5	6
∅ *quel* + noun	38	9	47
∅ *lequel*	78	24	102
subtotal (P-omission)	117	38	155 (62%)
Total	167	83	250

Table 6: Annotated sluices from Frantext subcorpus.

The following are examples of sluices with PP correlates and NP remnants from the corpus. Notice that both (25a) and (25b) involve the omission of the weak preposition *à* 'to', and that the sluice in (25a) means *à quelle amie (dois-je) demander?* ('which friend (should I) ask?'), with modal enrichment (Anand et al. 2021).

(25) a. - Je ne sais pas. Demande à ton amie. - Quelle amie? (Labruffe 2019)
- I NEG know NEG. Ask to your friend. - Which friend?
'- I don't know. Ask your friend. - Which friend?'

b. - Il est temps […] de passer à une étape nouvelle. - Laquelle? (Roubaud 2006)
- It is time […] of move to a step new - which.F.SG
'- It's time […] to move on to a new step. - Which step?'

3.3 Corpus analysis

We consider the effect of the correlate and remnant, the effect of the preposition and verb-preposition dependency, the effect of matrix/embedding, then the effect of direct/reprise use[12].

3.3.1 Sluices with implicit correlates

We start by looking at cases with an implicit PP correlate (cf Information hierarchy **Figure 1**). Chung (1995; 2014)'s generalization suggests that P-omission is not possible in this case: *They're jealous, but it's unclear of whom* (Chung 2014: 80), and Dagnac (2018) claims that this holds for French.[13] We found 80 cases with an implicit PP correlate, most of them with P-inclusion, as in (26) with the remnant *par quoi* 'by what' and a correlate that is the implicit agent phrase of the passive.

(26) Et elle était bouleversée, je me demandais par quoi. (d'Ormesson 1987)
and she was upset, I REFL wondered by what.
'And she was upset, I wondered by what.'

We also found 5 examples of P-omission with implicit correlates (**Table 7**), all with nominal remnants. Example (27) shows the omission of *dans* 'in' from *dans quel hôtel* 'in which hotel?', example (28) the omission of *de* 'of' from *de quelle équipe* 'of which team?' and example (29) the omission of *à* 'at' from *à quelle heure* 'at what time?'. Similarly to Anand et al. (2021) for English, they are all adjuncts, and the locative or temporal meaning may be provided by the noun (*heure* 'time', *hôtel* 'hotel').

[12] We provide a statistical model in section 3.4 which says which factors are significant.

[13] However, in the Santa Cruz Sluicing Dataset (US newspapers), Anand et al. (2021) found 17 instances of P-omission with implicit PP correlates, such as: *Decker was weaned in the world of investing by his father, who had also been a mutual fund manager. (Decker won't say which firm)* (Anand et al. 2021: 18), *which firm* meaning 'which firm he has been a mutual fund manager at'

(27) Ils l'avaient perdu à Strasbourg, je ne sais plus quel hôtel.
They him-have lost in Strasbourg, I NEG know anymore which hotel
(Navarre 1988)
(in which hotel)
'They lost him in Strasbourg, I don't remember which hotel.'

(28) - Le gars m'avait taclé la cheville […] - Quelle équipe ? (Palain 2019)
- The guy me-has tackled the ankle […] - Which team
'- The guy had tackled my ankle […] - Which team?'

(29) - […] Nous ne boirons plus jusqu'à demain matin. - Quelle heure? demande
- […] We NEG drink NEG until-at tomorrow morning - Which hour asks
Zins. (Jung 2018)
Zins
'- […] We won't be drinking again until tomorrow morning. - What time? asks Zins.'

Sluices	**–Prep**	**+Prep**	**Total**
qui/quoi	0	54 (100%)	54
quel + noun	5 (20.8%)	19 (79.2%)	24
lequel	0	2 (100%)	2
Total	5 (6.3%)	75 (93.8%)	80

Table 7: P-omission in sluices with implicit correlates.

We conclude that P-inclusion in a sluice with an implicit PP correlate is highly preferred but not compulsory in French. This may be explained by more general cognitive principles governing anaphora production (Parker 2019; Hofmeister et al. 2007): less accessible antecedents (here implicit correlates) favor more explicit anaphors (here PP remnants). This pattern is also predicted by MiF (Nykiel & Hawkins 2020).

In what follows, we only consider the subset of 170 sluices with an explicit PP correlate.

3.3.2 The role of the (explicit) correlates

In examples with explicit correlates, we find that NP remnants (88.2%) outnumber PP remnants (11.8%). The results are summarized in **Table 8**.

Sluices	Matrix	Embedded	Total
Prep + *qui*/*quoi*	9	3	12
Prep + *quel* + noun	5	1	6
Prep + *lequel*	1	1	2
subtotal (PP remnants)	15	5	20 (11.8%)
Ø *qui*/*quoi*	1	5	6
Ø *quel* + noun	36	6	42
Ø *lequel*	78	24	102
subtotal (P-omission)	115	35	150 (88.2%)
Total	130	40	170

Table 8: Annotated sluices with explicit PP correlates (Frantext subcorpus).

As suggested by Nykiel (2013), we compared nominal and pronominal correlates, and found out that the former highly favor P-omission compared to the latter (see **Table 9**).[14]

Correlate	–Prep	+Prep	Total
Nominal	148 (92.5%)	12 (7.5%)	160
Pronominal	2 (20%)	8 (80%)	10
Total	150 (88.2%)	20 (11.8%)	170

Table 9: The effect of pronominal vs. nominal correlates on P-omission in French sluices.

The disproportion between nominal and pronominal correlates, the latter being very few in our corpus, may explain the very high rate of P-omission (88.2%) that we found.

[14] As suggested by a reviewer, we checked for cases where a NP fragment intervenes between the PP correlate and the sluice. We found 5 occurrences, all with *quel* + N remnants, and all were without prepositions (see e.g. (i)).

(i) - Tout est rentré dans l'ordre. - L'ordre ? Quel ordre ? (Page 1982)
- All is entered in the-order - the-order ? which order ?
'- Everything has gotten back in order. - In order? What order?'

If we exclude these instances from the data, we have a 77% P-omission rate with *quel* + N, which is still high compared to the 87.5% rate reported in section 3.3.3. The intervening NP fragment may prime P-omission in the following sluice as in Nykiel & Hawkins (2020), but this is not driving our data.

3.3.3 The role of the remnant

We found three types of remnants: nominal (*quel* 'which' + noun) (as in (25a) and (30)), or pronominal: *lequel* 'which.M.SG' (as in (25b) and (31)), and *qui/quoi* 'who/what' (32).

(30) - C'est que j'ai vu votre affichette sur un banc. - Quel banc?
- This.is that I-have seen your flyer on a bench. - Which bench?
'- It's just that I saw your flyer on a bench. - Which bench?' (Bouherrafa 2019)

(31) - Obéir au gouvernement, messieurs. - Lequel?
- Obey.INF to.DEF government, gentlemen. - which.M.SG?
'- Obeying to the government, gentlemen. - Which one?' (Jenni 2011)

(32) - Il avait fait mal à quelqu'un. - À qui ?
- He has done pain to someone - To who
'- He hurt someone - Who?' (Bienne 1990)

We first checked the correlation between correlates and remnants, see **Table 10**. In (spoken) English corpora, (Nykiel 2017) found a high frequency of information matching correlates and remnants (both pronominal or both nominal). Here we found very few cases with pronominal correlates and remnants: the more frequent case is a nominal correlate with a pronominal (*lequel*) remnant.

Correlates	**Remnants**			**Total**
	qui/quoi	***quel* + *noun***	***lequel***	
Nominal	10	46	104	160
Pronominal	8	2	0	10
Total	18	48	104	170

Table 10: Number of occurrences per correlate-remnant combination.

We also found (**Table 11**) that while *lequel* and *quel*+noun highly favor P-omission, bare wh-pronouns (*qui, quoi*) disfavor it. Since *qui* and *quoi* are quite rare in our corpus, this may explain the very high rate of P-omission (88.2%) we found overall.

Remnant	–Prep	+Prep	Total
qui/quoi	6 (33.3%)	12 (66.7%)	18
quel + noun	42 (87.5%)	6 (12.5%)	48
lequel	102 (98.1%)	2 (1.9%)	104
Total	150 (88.2%)	20 (11.8%)	170

Table 11: P-omission with different remnant types.

First, we looked at the difference between pronominal *lequel, qui, quoi* versus nominal *quel*+noun. **Table 12** shows that this factor does not have an effect.

Remnant	–Prep	+Prep	Total
Nominal	42 (87.5%)	6 (12.5%)	48
Pronominal (*qui*, *quoi*, *lequel*)	108 (88.5%)	14 (11.5%)	122
Total	150 (88.2%)	20 (11.8%)	170

Table 12: P-omission with nominal/pronominal remnants.

Second, grouping together *quel*+noun and *lequel* as anaphoric remnants, we obtain another partition. Anaphoric remnants led to a higher frequency of P-omission (94.7%) compared to non-anaphoric *qui/quoi* (33.3%) (see **Table (13)**).

Remnant	–Prep	+Prep	Total
Anaphoric (*quel*+N, *lequel*)	144 (94.7%)	8 (5.3%)	152
Non-anaphoric (*qui/quoi*)	6 (33.3%)	12 (66.7%)	18
Total	150 (88.2%)	20 (11.8%)	170

Table 13: P-omission with anaphoric and non-anaphoric remnants.

Looking for a possible correlation between remnants and correlates, we found (**Table 14**) a preference for P-omission with a nominal correlate and an anaphoric remnant, and a dispreference with pronominal correlates and non-anaphoric remnants. This preference may be driven by nominal correlates, since non-anaphoric remnants are rare in our corpus.

correlate	Remnant	–Prep	+Prep	Total
Nominal	Anaphoric	143 (95.3%)	7	150
Pronominal	Anaphoric	1	1	2
Nominal	Non-anaphoric	5	5	10
Pronominal	Non-anaphoric	1 (12.5%)	7	8
Total		150 (88.2%)	20 (11.8%)	170

Table 14: Interaction between the correlate and remnant types on P-omission.

As shown by the general model in section 3.4, Anaphoric remnant is a significant factor for P-omission and its interaction with a nominal correlate is significant as well (see **Table 21** below).

3.3.4 The role of the preposition

According to Merchant's contrast (between 10a and 11 above), *à* 'to', as well as other weak prepositions *de* 'of' and *en* 'in' (Abeillé & Kupfermann 2021), should not allow for P-omission, unlike strong prepositions like *avec 'with', avant 'before', après 'after'*. According to the information hierarchy (**Figure 1**), weak prepositions being less informative should allow for more P-omission, and the same prediction holds for Nykiel's theory, since weak prepositions tend to be more verb-dependent than strong ones (Bonami 2021). In light of these hypotheses, we looked at the effect of these two preposition types on P-omission in the corpus. **Table 15** shows that P-omission is highly preferred regardless of the preposition type.

Preposition	–Prep	+Prep	Total
à/de/en	97 (87.4%)	14 (12.6%)	111
Other prepositions	53 (89.8%)	6 (10.2%)	59
Total	150 (88.2%)	20 (11.8%)	170

Table 15: Effect of weak/strong preposition on P-omission.

Although P-stranding is not possible in French, many prepositions can stand alone, and they are a subtype of strong prepositions (33a). The omission of the complement is impossible after a weak preposition (*à, de, en*) but also after some strong prepositions (*par, sur, sous*) (Tseng 2021) (33b).

(33) a. Cette proposition, je vote pour/contre.
This proposition I vote for/against
'This proposition, I am for/against it.'
b. *Cette proposition, je pense à / parle de / compte sur.
This proposition I think of / talk of / count on

We thus tested another bipartition of prepositions: stand-alone (e.g. *pour, contre, avec, pendant* 'for, against, with, during') versus non-stand-alone (e.g. *à, chez, dans, de, en, jusqu'à, sur, sous, par, vers* 'to, at, in, of, in, until, on, under, by, towards'). **Table 16** shows that P-omission is almost equally high with both types of prepositions.

Preposition type	**–Prep**	**+Prep**	**Total**
Stand-alone	19 (90.5%)	2 (9.5%)	21
Not-stand-alone	131 (87.9%)	18 (11.9%)	149
Total	150 (88.2%)	20 (11.8%)	170

Table 16: P-omission with stand-alone and not-stand-alone prepositions.

We also tested whether preposition length had an effect, as observed for Russian (Philippova 2014) and Polish (Nykiel & Kim 2022b). We found only 15 cases of polysyllabic prepositions (e.g. *pendant* 'during'), with 93.3% P-omission, which is not very different from the 87.7% found with monosyllabic ones (e.g. *sur* 'on').

Overall, we did not find any effect of preposition type on P-omission.

3.3.5 The role of preposition dependency

We annotated predicate-preposition dependency and eliminated three irrelevant examples since they are not introduced by an overt predicate (e.g. (34)). We retained cases where the preposition was dependent on a non-verbal predicate (e.g. (35)).

(34) - Tout, sauf un détail. - Lequel ? (Navarre 1988)
- All except a detail - which.M.SG ?
'- All except a detail. - Which one?'

(35) Un enfant qui dort paraît si sage, […] plus proche de la Vérité (mais laquelle?) (Demoulin 2016)
A kid who sleeps seems so wise […] more close to the truth (but which.F.SG)
'A sleeping child appears so wise, closer to the Truth (but which one?).'

We considered the preposition as dependent when the PP was obligatory or when the preposition could not be changed.[15] For example, the preposition *sur* 'on' in (30) was annotated as non-dependent, and *au* 'to.DEF' in (31) as dependent. We found that dependent prepositions outnumbered independent ones but did not favor P-omission (see **Table 17**).

Predicate dependency	**–Prep**	**+Prep**	**Total**
Non-dependent	59 (90.8%)	6 (9.2%)	65
Dependent	88 (86.3%)	14 (13.7%)	102
Total	147	20	167

Table 17: Predicate dependency effect on P-omission.

Limiting ourselves to verb-governed prepositions (121 cases), we did not find an effect on P-omission: the omission rate was 88.2% (45/51) with independent prepositions, and 85.7% (60/70) with verb-dependent ones.

We also compared our annotation with Nykiel and Hawkins (2020)'s finer coding with three dependency levels, using Hawkins's (1999) entailment tests: PP omission (A test: X PP entails X), testing X dependency on P, and X pronominalization (B test: X PP entails Pro-V PP), testing P dependency on X, where X can be a verb, noun, or adjective within the VP. Using these tests, they define three levels:

- Level 0: No dependency between X and P (both tests succeed (30), or X depends on P (only test B succeeds);
- Level 1: P depends on X (only test A succeeds (36));
- Level 2: Both P and X depend on each other (both tests fail (25b)).

A: (30) entails *J'avais vu votre affichette* ('I saw your flyer') hence V is independent from the P *sur*.
B: (30) entails *Je l'avais vue sur un banc* 'I saw it on a bench', hence P is independent from V.

(36) Parlons d'autre chose, dit-il, cassant. Elle hésita. - De quoi ?
talk.IMP.1PL of-another thing, says-he, breaking-off. She hesitated. - Of what?
'Let's talk about something else, he said, breaking off. She hesitated. - About what?'
(Labro 1982)

[15] These tests are used by Bonami (2021) to distinguish between complement and adjunct PPs in French.

A: (36) entails *Parlons* ('let's talk'), hence V is independent from P.
B: (36) does not entail *Faisons-le d'autre chose* (lit. 'Let's do it of something else'), hence P is dependent on V.

A: (25b) does not entail *Il est temps de passer* 'It is time to pass', hence V is dependent on P.
B: (25b) does not entail *Il est temps de le faire à une étape nouvelle* (lit. 'It is time to do it to a new step'), hence P is dependent on V.

We excluded 26 examples from our analysis because they did not involve a VP[16].

Some of our cases of obligatory Prep (**Table 17**) are now considered as level 0 (both tests succeed (32)), this is why we now have more independent Ps (75/141) with this new coding.

Table 18 shows that P-omission is consistently preferred in our dataset, regardless of the dependency level.[17]

Semantic dependency level	**–Prep**	**+Prep**	**Total**
Level 0	66 (88%)	9 (12%)	75
Level 1	33 (86.8%)	5 (13.2%)	38
Level 2	23 (82.1%)	5 (17.9%)	28
Total	122 (86.5%)	19 (13.5%)	141

Table 18: Semantic dependency effect on P-omission (Nykiel & Hawkins coding).

We did not find an effect of predicate-preposition dependency on P-omission, unlike Nykiel and Hawkins (2020), who found for English an increasing preference for P-omission as follows: (58%) at level 0, (77.5%) at level 1 and (88.5%) at level 2.

3.3.6 The role of embedding

Since Merchant (2001) and Rodrigues et al. (2009) only discuss embedded sluices, we hypothesize that embedding may have an effect. We found instances of P-omission both with embedded sluices (37) and matrix sluices.

---

[16] As in example (i):

(i) - Voix profonde d'un homme (mais lequel?) (Ernaux 1997)
- Voice deep of-a man (but which.M.SG)
'- A deep voice of a man (but which one?)'.

[17] We ran a model to test the interaction between semantic dependencies and correlate content and found no significant interaction.

(37) - Cela devait donner matière à un livre, je n'ai jamais trouvé
- This should give matter to a book I NEG-have never found
lequel. (François 2016)
which.M.SG
Lit. '- It would be a good story for a book, I never found which one.'

We found 16 coordinated sluices (e.g. (38)), which we grouped together with the root ones, as matrix sluices.

(38) - Il est effectivement atteint d'un mal incurable mais lequel? (Pontalis 1996)
- He is indeed affected of-a disease incurable but which.M.SG?
'- He is indeed suffering from an incurable disease but which disease.'

Table 19 indicates that P-omission is almost equally frequent in both constructions.

Construction type	**–Prep**	**+Prep**	**Total**
Embedded	35 (87.5%)	5 (12.5%)	40
Matrix	115 (88.5%)	15 (11.5%)	130
Total	150 (88.2%)	20 (11.8%)	170

Table 19: P-omission in embedded vs. matrix sluices with explicit correlates.

Our corpus data show that P-omission is highly preferred in both embedded and matrix (root and coordinate) sluices.

3.3.7 The role of direct/reprise use

Ginzburg & Sag (2000) identified two types of interpretation: reprise sluices, used for clarification requests, as in (39), which are always root, and direct sluices, used for information seeking (40), which can be embedded.

(39) - Jo phoned. - WHO? (Ginzburg & Sag 2000: 408)

(40) - Someone called. - Who?

We annotated root sluices (leaving aside embedded and coordinate sluices) into reprise (41) and direct uses (see example (34)). The former usually have a definite correlate, which can be nominal (e.g. *mon ami* 'my friend', proper names) or pronominal (e.g. *lui* 'him'), and the latter an indefinite correlate, which can be nominal (*un ami* 'a friend') or pronominal (*quelqu'un* 'someone').

(41) - Césari était aviateur et “a descendu plein d’Allemands pendant la première
- Césari was pilot and had shot.down plenty of-Germans during the first
guerre”. - Quelle guerre ? (d’Ormesson 1987)
was - Which war
‘- Césari was an aviator and “shot down a lot of Germans during the First War”. - Which war?’

The results in **Table (20**) show that reprise use does not seem to have an effect and that both sluices exhibit a very high rate of P-omission.

Sluice	**–Prep**	**+Prep**	**Total**
Reprise	17 (85%)	3 (15%)	20
Direct	83 (88.3%)	11 (11.7%)	94
Total	100 (87.7%)	14 (12.3%)	114

Table 20: The effect of reprise use on P-omission in French matrix sluices.

3.4 General corpus model

We fit our corpus data in a Generalized Linear Mixed Model (GLMM) (McCulloch & Neuhaus 2005). We used the glmer function from the lme4 package in R, specifying a binomial family, and using the Nelder_Mead optimizer within the glmerControl function. In our best model, the fixed factors are P-omission (reference level P-inclusion) and two predictors: pronominal correlates versus nominal correlates (the latter being the reference level[18]), and anaphoric remnants versus non-anaphoric remnants (the latter being the reference level). Author names were included as a random factor. The model in **Table 21** highlights that, among the predictors, only the anaphoric remnant and its interaction with pronominal correlates are significant[19]. Anaphoric remnants favor P-omission and disfavor it with pronominal correlates. We note, however, that only two examples with these criteria were found in the dataset, cf. **Table 14** (section 3.3.3).

The model achieved approximately 98.24% accuracy, 98.68% precision and 99.33% recall, outperforming simpler models that included only correlate or remnant as predictors, which showed significant effects for each factor. Models with more predictors (embedding, stand-alone P, direct/reprise, etc.) did not show other significant effects. Attempts to compute the interactions between all the variables resulted in convergence issues.

[18] Treatment contrasts were applied for the predictors in the model.

[19] We adjusted for multiple comparisons using the Bonferroni correction. The initially found statistically significant results ($p < 0.05$) persisted, with the adjusted p-values remaining below the corrected significance threshold.

Predictor	Estimate	Std. Error	Z-value	P-value
(Intercept)	–6.522	3.099	–2.105	0.0353 *
Anaphoric remnants	17.043	3.810	4.473	7.71e-06 ***
Pronominal correlates	–3.887	4.298	–0.904	0.3659
Anaphoric remnants: Pronominal correlates	–15.490	7.160	–2.164	0.0305 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 21: Model with anaphoric/non-anaphoric remnants and pronominal/nominal correlates (GLMM Results).

3.5 Discussion of the corpus study

Our corpus study shows that P-omission in French interrogative sluices is highly preferred. Among the factors we tested, only correlate and remnant types had significant effects. As predicted by Rodrigues et al. (2009), anaphoric (*lequel*) remnants favored P-omission, and so did nominal (*quel* + noun) remnants, as predicted by Nykiel.

The dispreference for pronominal correlates and non-anaphoric remnants that we found may be a general property of discourse cohesion and anaphora resolution. Pronominal correlates have weaker memory representations compared to nominal ones, making them harder to retrieve and necessitating more complex anaphors as cues (Hofmeister et al. 2007, Parker 2019). Sluices with pronominal correlates (*quelqu'un*) and non-anaphoric remnants (*Qui*?) may increase processing difficulty, which can explain their lower frequency in our corpus, independently of P-omission, as MiF predicts (Nykiel & Hawkins 2020).

Contrary to Merchant (2001)'s contrast for French, the type of preposition did not have an effect in our data.

In **Table 22**, we compare our results with those from Nykiel (2017), on (spoken) English (Corpus of Contemporary American English, Switchboard and Santa Barbara corpus) and Polish (National Corpus of Polish).

Language	P-omission rate
(written) French	88.2%
(spoken) English	67.2%
(spoken) Polish	18.3%

Table 22: P-omission in sluices with explicit PP correlates.

Nykiel's hypothesis that in a non-P-stranding language like Polish, P-omission is possible but not preferred, unlike in a P-stranding language like English, is not confirmed for French, which shows a higher P-omission rate than English.

Modality may play a role since her corpora were spoken while ours is written. The corpus samples may play a role as well. In Polish and English, she extracted sluices with bare pronominal (*who, what*) and nominal remnants (*which/what/whose* + noun) but not *which* or *which one*, which would parallel our *lequel* sluices. Leaving aside *lequel* remnants, we found a 72.7% P-omission rate in our French data, which is still higher than what she found for English.

4 General discussion

We studied preposition omission under interrogative sluicing in French in two ways: at the comprehension level with two acceptability judgment experiments, and at the production level through a corpus-based analysis.

The experiments showed that P-omission is possible in French and favored by a nominal remnant (*quel* + noun) in Experiment 2, compared to a pronominal remnant (*qui*) in Experiment 1. The experiments also confirmed that the preposition is required in verbal interrogatives, and this is challenging for theories deriving sluices from verbal interrogatives. Our corpus study shows that P-omission is highly frequent in French sluices, and is favored by anaphoric remnants, in addition to nominal correlates.[20]

4.1 Issues for a P-stranding analysis

Our experiments show that P-omission is acceptable in French sluices with *à* and *de*, unlike in full interrogatives. Our corpus study shows that P-omission is possible with various prepositions and is preferred to P-inclusion, despite French being a non-P-stranding language. This contradicts Merchant's hypothesis that P-omission would be restricted to prepositions like *avec* 'with' or other stand-alone prepositions (with a null pronoun), such as *pour* 'for' and *contre* 'against'. In our corpus data, the P-omission rate with *de* 'of', *à* 'to' and *en* 'in' (87.4%), was not significantly lower than with strong prepositions (89.8%) or stand-alone ones (90.5%).

Since Merchant's illusory contrast between *à* and *avec* was based on the intuition of three linguist colleagues, we may also highlight the importance of quantitative data and controlled experiments in linguistics (Gibson & Fedorenko 2013).

Given our corpus data, we are in a better position to understand why Merchant's informants may have been mislead: the example he used was with a pronominal correlate (*quelqu'un*) and a

[20] As suggested by a reviewer, a difference between the experiments and the corpus study was that the former included only different speaker sluices and the latter included both different- and same-speaker sluices. In our corpus, we have 100 same-speaker and 70 different-speaker sluices, with a similar P-omission rate (88% in the former and 88.6% in the latter).

bare wh-remnant (the non-anaphoric pronominal *qui*), which we found to be the least frequent in our corpus study. Only 8/170 of the analyzed data were with pronominal correlates and pronominal remnants, out of which only one example is with P-omission (42). This also highlights the failure of the one-example-one-speaker intuitionist method, and the importance of taking into account naturally occurring data.

(42) - Préserver Carville, à tout prix : j'avais une dette à payer, contractée par
- Preserve.INF Carville at all cost : I-had a debt to pay contracted by
d'autres, je ne sais qui. (Pontalis 1980)
of-others, I NEG know who
'- To preserve Carville, at all costs : I had a debt to pay, contracted by others, I don't know who.'

French thus adds to the list of non-P-stranding languages, allowing for P-omission under sluicing.[21]

4.2 Issues for a cleft analysis

Rodrigues et al. (2009) suggest that in non-P-stranding languages, on top of a simple derivation with a PP remnant, an alternative derivation with a cleft is available for NP remnants. Given the more complex derivation for sluices with P-omission, they would expect that they would be less frequent than prepositional ones, with a PP correlate, contrary to our corpus findings. We also found examples of P-omission with the strong pronoun *quoi* 'what', which are not compatible with a cleft source, as noted by Dagnac (2018), Abeillé and Hassamal (2019) and Gotowski (2022).[22]

(43) - [...] Je suis tellement bluffé par ce truc-là - Quoi? Le dessin de
- [...] I am so shaken by this thing-here - What? The drawing of
Nina? (Gavalda, 2000)
Nina?
'- [...] I'm so shaken by this thing - What? Nina's drawing?'

The P-omission rate of *quoi* (4/12 = 33.3%) is lower than that of *lequel* (98.1%), but higher than that of *qui* (2/7 = 28.6%).

Similar findings challenging a cleft source have been reported in other non-P-stranding languages, such as Polish (Nykiel 2013). Molimpakis (2019) experimentally showed that P-omission is possible in Greek and German sluices, but not in clefts. Alshaalan and Abels (2020) presented similar results for Saudi Arabic.

---

[21] Even for English, it has been shown (Levin 1982, Chung et al. 1995, Ginzburg & Sag 2000, Sag & Nykiel 2011), that P-omission may apply independently of P-stranding : *They will report me under some circumstances, but I forget (under) which circumstances. I forget under which circumstances they'll report me / *which circumstances they'll report me under.* (Levin 1982: 607–608)

[22] We only kept examples where it is clear that *quoi* is a sluice with a PP correlate and not the discourse particle meaning 'pardon'.

4.3 Further issues for syntactic reconstruction

Further issues arise with syntactic reconstruction theories. Any analysis based on wh-movement predicts that sluices should be sensitive to locality constraints. In our corpus, we found examples which violate island constraints, such as the adjunct island. These instances occur with a correlate inside a relative clause as in (44), and inside an adjunct PP as in (45).

(44) - Je voudrais voir le laboratoire où vous préparez les plats de nos clients.
- I want see the laboratory where you prepare the dishes of our clients.
[…] - Mais quels clients (*c'est de qui vous voudriez voir le laboratoire où
[…] - But which clients (it-is of who you want see the laboratory where
vous préparez les plats)? (Lovey, 2014)
you prepare the dishes)?
'- I would like to see the laboratory where you prepare the dishes for our clients. […] - But which clients (*it is that you would like to see the laboratory where you prepare the dishes for)?'

(45) - Il y avait donc là-bas… cachée au fond de la forêt, une petite
- There LOC had thus there… hidden at.the heart of the forest a little
princesse… - Quelle forêt? (*c'était de laquelle il y avait cachée au
princess… - Which forest? (*it-is of which.F.SG there LOC had hidden at.the
fond une petite princesse)? (Sarraute 1982)
heart a little princess)
'- There was thus over there… hidden deep in the forest, a little princess… - Which forest? (*it was that there was a little princess hidden deep in)?'

Some correlates also occur in a subject island, as in (46).

(46) - Deux soldats de deux camps ennemis qui s'affrontent. - Quels camps? (*c'est
- Two soldiers of two camps ennemies who REFL-confront. - Which camps (it-is
dont deux soldats qui s'affrontent?) (Sarraute 1982)
of.which two soldiers who REFL-confront?)
'- Two soldiers of two ennemy camps who confronted each other. - Which camps? (*it was that two soldiers of confronted each other?)'

While various proposals have tried to explain island insensitivity under deletion theories (Merchant 2001, a.o.), it seems fair to say that direct interpretation approaches predict such insensitivity (if there is no syntactic structure at the ellipsis site).

In addition, we found instances of sluices without a verbal antecedent in the corpus, as in (47) and (48). In such cases, reconstruction of a verbal structure is challenging.

(47) - Marre de tes conneries. - Quelles conneries? (Page 1982)
- sick of your bullshits. - Which bullshits?
'- I'm fed up with your bullshit. - What bulshit?'

(48) - Rien à fiche de la ligne d'arrivée. - Quelle ligne? (Bouillier 2018)
- Nothing to care of the line of-arrival. - Which line?
'- Who cares about the finish line? - Which line?'

4.4 Processing-based factors

Our corpus study showed that (written) French sluices (with an explicit PP correlate) have a very high rate of P-omission (88.2%), driven by nominal correlates and anaphoric remnants: *quel* + noun (87.5%) and *lequel* (98.1%), which are also more frequent than the non-anaphoric remnants *qui/quoi* (33.3% P-omission rate). This confirms the effect of processing-based factors.

In our experiments, with nominal correlates, P-omission was found acceptable, and (anaphoric) nominal remnants favored P-omission (Exp.2), while (non-anaphoric) pronominal ones disfavored it (Exp.1). This aligns with cue-based retrieval mechanisms, given that more informative nominal remnants are better cues than pronominal ones.

In (spoken) English, Nykiel (2017) found that contentful (nominal) correlates exhibited a greater tendency to allow for P-omission (i.e., NP remnants). Conversely, non-contentful correlates (non-anaphoric pronouns), were predominantly followed by prepositional remnants. In our corpus, the effect of the correlate type was even stronger: with 20% P-omission rate with pronominal correlates and 92.5% P-omission rate with nominal ones. This goes in line with Nykiel (2017) and Nykiel and Hawkins (2020)'s theory since nominal correlates are more accessible in memory than pronominal ones.

Correlates	**French**		**English**	
	PP REMNANT	**NP REMNANT**	**PP REMNANT**	**NP REMNANT**
Contentful (Nominal)	12 (7.5%)	148 (92.5%)	42 (24,1%)	132 (75,9%)
Less contentful (Pronominal)	8 (80%)	2 (20%)	53 (39%)	83 (61%)

Table 23: Sluices with PP correlates in (written) French vs. (spoken) English.

4.5 An HPSG analysis

Our results go in-line with context-based direct interpretation approaches to ellipsis (Ginzburg & Sag 2000; Sag & Nykiel 2011, Ginzburg 2012). According to these analyses, there is no syntactic structure at the ellipsis site. Sluicing, then, yields a structure in which a clause is reduced to a *wh*-phrase, which derives its meaning from the context. In this section, we sketch an analysis

based on Head-driven Phrase Structure Grammar (HPSG) (Pollard & Sag 1994; Müller et al. 2021), following Sag & Nykiel (2011). In this framework, linguistic signs can be lexical or phrasal, and their description is represented with a feature structure including SYN (for syntax), SEM (for semantics) and CTXT (for context) attributes. Well-formed signs inherit their properties from a hierarchy of types. An example of a hierarchy of phrases is in **Figure 6**, where phrasal types are distinguished according to the syntactic function of their immediate constituents (head, subject, complement, etc.) and to their clause type (declarative, interrogative, etc.).

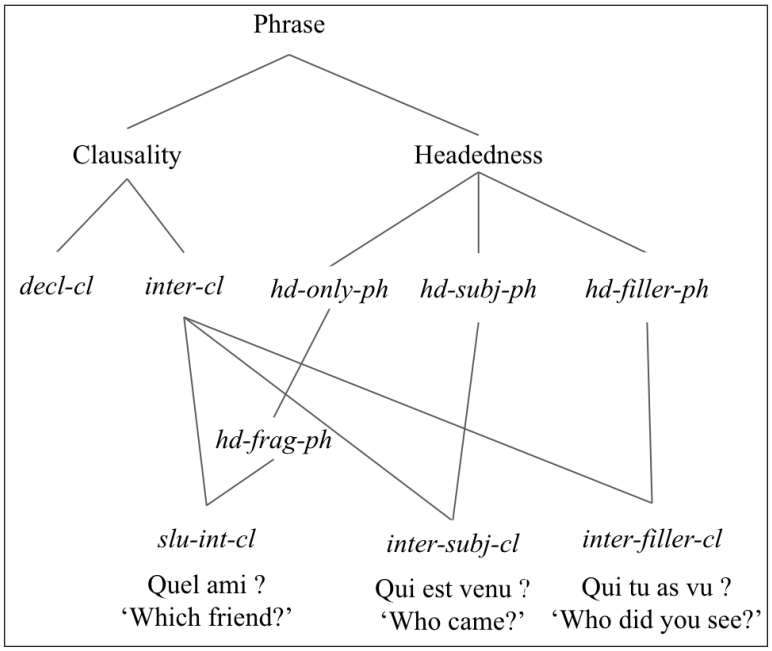

Figure 6: Clausal types hierarchy.

We analyze sluices as clausal fragments (head-only phrases) with a propositional meaning (SEM) computed from the context. We use Ginzburg & Fernandez (2010)'s Dialogue Game Board (DGB), which records information that emerges through dialogical interaction, keeping track *inter alia* of the interlocutor roles (speaker, addressee) and discourse topic(s), and includes a Maximal Question under Discussion (MAX-QUD), updated throughout the interaction, which serves as the antecedent of the sluice. We rely on a *sluice-phrase* type, which inherits from a more general *head-fragment-phrase* type. As suggested by Ginzburg & Sag (2000) and Sag & Nykiel (2011), the content of the antecedent clause (MAX-QUD) is analyzed as a polar question, i.e. lambda abstraction over an empty set of parameters, with a propositional content (p(j)). The SEM of the sluice is a wh-question, with a non-empty set of parameters, and the same propositional content (p(j)).[23]

[23] We do not distinguish here between direct and reprise sluices, although the former require a positive existential quantifier in the correlate (SAL-UTT), such as *someone* or *a friend*.

(49) Sluice rule (simplified)

$$\left[\begin{array}{ll} \textit{sluice-clause} & \\ \textsc{syn} \mid \textsc{cat} & S \\ \textsc{sem} & \lambda\{j\}p(j) \\ \textsc{ctxt} \mid \textsc{dgb} & \left[\begin{array}{ll} \textsc{max-qud} & \lambda\{\}p(j) \\ \textsc{sal-utt} & \left\{\left[\begin{array}{l} \textsc{syn} \mid \textsc{cat}\ \boxed{1} \\ \textsc{sem}\left[\textsc{index}\ j\right] \end{array}\right]\right\} \end{array}\right] \end{array}\right] \rightarrow \left[\begin{array}{ll} \textsc{syn} & \left[\textsc{cat}\ \boxed{1}\ \textit{nonverbal}\right] \\ \textsc{sem} & \left[\textsc{index}\ j\right] \end{array}\right]$$

The remnant is coindexed with a correlate, which is a salient utterance (SAL-UTT) in the DGB, with which it shares its syntactic properties (CAT for syntactic category).[24]

Applying this rule to example (19), repeated in (50) for convenience, the sluice retrieves its propositional content (p(j)) from the MAX-QUD in the antecedent clause; the remnant, the only daughter of the S node, is coindexed (INDEX *j*) with a contextual correlate (SAL-UTT) and matches its category (CAT 1): it is thus either a PP as in (50a) (see **Figure 7**) or a NP as in (50b) (see **Figure 8**). The RELS feature specifies the semantic relation expressed by the noun in the remnant.

(50) Speaker A: J'ai parlé à un ami.
'I have talked to a friend.'

a. Speaker B: À quel ami ?
'To which friend?'

b. Speaker B: Quel ami ?
'Which friend ?'

In both examples, the antecedent clause is speaker A's utterance: *I talked to a friend*. So the MAX-QUD is a polar question (whether speaker A talked to a friend), and the sluice inherits the same propositional content to denote a wh-question (which friend A talked to). In both **Figures 7** and **8**, the SAL-UTT is the correlate, either the PP *à un ami* or the NP *un ami*, respectively. It shares its category and index with the remnant's node: a PP (*À quel ami*) or a NP (*Quel ami*).

[24] Fragments may come in different subtypes. Recent studies (Nykiel & Kim 2022a; Abeillé & Kim 2022) suggest that some fragments require no CAT-feature identity between the remnant and the correlate (SAL-UTT), but only valence (VAL) identity. This accounts for possible case mismatch between correlate and remnant in Bulgarian, Korean, or English (-Who is coming? -Me/*I). This does not seem to be needed for French interrogative sluices.

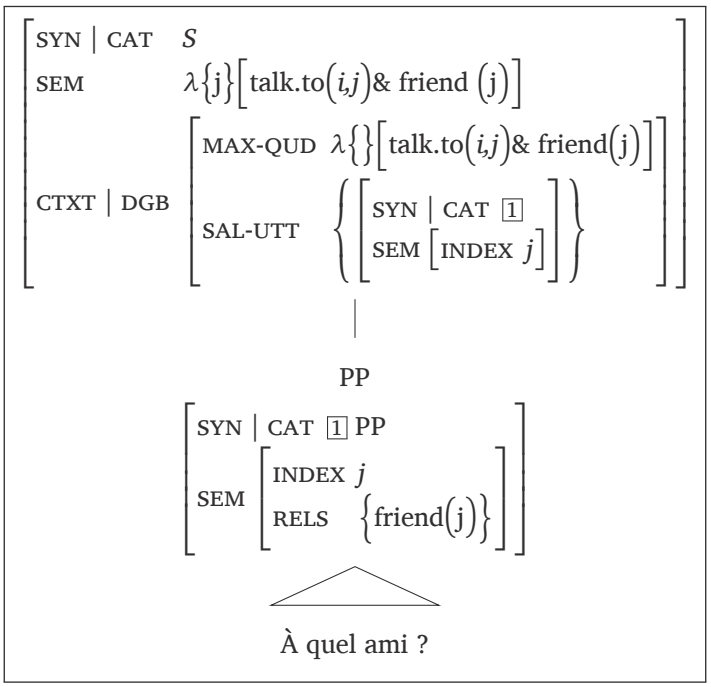

Figure 7: Analysis of a PP sluice with an explicit PP correlate.

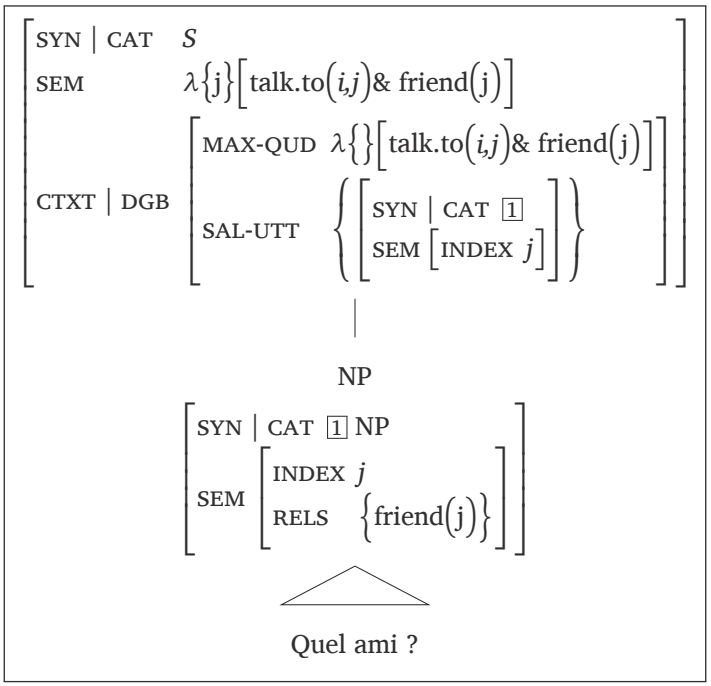

Figure 8: Analysis of an NP sluice with an explicit PP correlate.

5 Conclusion

We show that P-omission is possible in French sluices, which was considered ungrammatical (or limited to certain prepositions) by Merchant (2001), or restricted to specific remnants by Rodrigues et al. (2009).

In two online acceptability judgment tasks, we tested French matrix sluices with weak prepositions (*à* 'to' and *de* 'of'), and with pronominal (*qui* 'who') and nominal (*quel* 'which' + noun) remnants. Results showed that P-omission is acceptable in French sluices, unlike in verbal interrogatives, and preferred with more contentful remnants (i.e. nominal) than with less contentful ones (i.e. pronominal non-anaphoric).

In a corpus analysis of written French (contemporary subpart of Frantext), we found a very high P-omission rate (88.2% in sluices with explicit PP correlates), with nominal correlates and anaphoric remnants as favoring factors. As French is not a P-stranding language, Merchant's deletion-based approach to sluicing cannot account for these results. We also show that our results challenge Rodrigues et al.'s cleft analysis, and sketch an HPSG analysis based on a direct interpretation approach (Ginzburg & Sag 2000; Sag & Nykiel 2011).

Our findings also lend support to Nykiel (2013; 2017) and Nykiel & Hawkins (2020)'s theory that P-omission is possible across a wide variety of languages and is mostly sensitive to processing-based factors: as the preposition is a cue for ellipsis resolution, more informative remnants and correlates help dispense with it. But unlike Nykiel & Hawkins (2020), we cannot conclude that English is "typologically unusual" nor "exceptional", since French also shows a higher frequency of P-omission in sluices.

Appendices

Appendix 1. Materials of experiment 1

(1) A: J'ai un coup de fil à donner à un voisin. 'I have a call to give to a neighbor.'
B:
+Prep+Ellipsis: À qui? 'To who?'
–Prep+Ellipsis: Qui ? 'Who ?'
+Prep–Ellipsis: À qui as-tu un coup de fil à donner? 'To whom do you have a call to give?'
–Prep–Ellipsis: Qui as-tu un coup de fil à donner? 'Who do you have a call to give?'

(2) A: J'essaie de faire attention aux autres. B: À qui ?
'I try to pay attention to others. B: - To who?'

(3) A: Je crois que j'ai fait de la peine à un élève. B: À qui ?
'I think that I-had given pain to a pupil. B: - To who?'

(4) A: J'ai parlé à un ami. B: À qui ?
'I have talked to a friend. B: - To who?'

(5) A: J'ai écrit à un vieil oncle. B: À qui ?
'A: I have written to an old uncle. B- To who?'

(6) A: Cette histoire a fait peur à un enfant. B: À qui ?
'A: This story has given fear to a child. B: 'To who?'

(7) A: Je dois tout dire au responsable. B: À qui ?
'A: I must say everything to the manager. - To who?'

(8) A: J'ai une dette à payer à mon cousin. B: À qui ?
'A: I have a debt to pay to my cousin. B: - To who?'

(9) A: J'ai eu des nouvelles d'un collègue. B: De qui ?
'A: I have had news of a colleague. B: - Of who?'

(10) A: Je me souviens d'un ancien voisin. B: De qui ?
'A: I remember an old neighbor. B: - Who?'

(11) A: C'est l'anniversaire d'un de mes enfants. B: De qui ?
'A: It's the birthday of one of my kids. B: Of who?'

(12) A: Je me suis éloigné de certains parents. B: De qui ?
'A: I have distanced myself from certain parents. B: - From who?'

(13) A: Je parle du nouveau candidat. B: De qui ?
'A: I'm talking about the new candidate. B: - About who?'

(14) A: Ça fera l'affaire d'un autre. B: De qui ?
'A: That will do for someone else. B: - For who?'

(15) A: Je me méfie de cet artiste. B: De qui ?
'A: I'm wary of this artist. B: - Of who?'

(16) A: Ce livre plait à beaucoup de gens. B: À qui ?
'A: This book pleases many people. B: - Who?'

(17) A: Je pense souvent à mon père. B: À qui ?
'A: I think often about my father. B: - About who?'

(18) A: J'ai peur de la bande dans le quartier. B: De qui ?
'A: I'm scared of the gang in the neighborhood. B: - Of who?'

(19) A: Je viens d'apprendre la mort de quelqu'un. B: De qui ?
'A: I just learned the death of someone. B: - Of who?'

(20) A: Je voulais m'approcher de lui. B: De qui ?
'A: I wanted to get closer to him. B: - To who?'

Models for Experiment 1

Parameters	Estimate	Est.Error	95% CrI	Post. Prob.
Intercept[1]	–2.82	0.30	[–3.43, –2.25]	1
Intercept[2]	–1.65	0.29	[–2.24, –1.09]	1
Intercept[3]	–0.82	0.29	[–1.40, –0.26]	1
Intercept[4]	0.21	0.28	[–0.35, 0.77]	0.78
No preposition	–2.53	0.29	[–3.12, –1.98]	1
Ellipsis	–0.24	0.25	[–0.73, 0.25]	0.83
No preposition:Ellipsis	1.67	0.27	[1.14, 2.22]	1

Table 24: Bayesian ordinal model: Exp. 1.

Parameters	Estimate	Est.Error	95% CrI	Post. Prob.
Intercept[1]	–2.16	0.21	[–2.57, –1.75]	1
Intercept[2]	–1.21	0.20	[–1.61, –0.82]	1
Intercept[3]	–0.57	0.20	[–0.96, –0.18]	1
Intercept[4]	0.27	0.20	[–0.12, 0.66]	0.92
No preposition	–1.90	0.16	[–2.22, –1.57]	1
Ellipsis	–0.23	0.15	[–0.53, 0.07]	0.93
de	0.03	0.22	[–0.40, 0.46]	0.56
No preposition:Ellipsis	1.11	0.22	[0.69, 1.54]	1
No preposition:*de*	0.05	0.22	[–0.38, 0.49]	0.6
Ellipsis:*de*	0.27	0.22	[–0.15, 0.71]	0.89
No preposition:Ellipsis:*de*	0.16	0.31	[–0.44, 0.76]	0.7

Table 25: Bayesian ordinal model: *à* vs. *de* - Exp. 1.

Appendix 2. Materials of experiment 2

(1) A: J'ai un coup de fil à donner à un voisin. 'I have a call to give to a neighbor.'
B:
+Prep+Ellipsis: À quel voisin? 'To which neighbor?'
–Prep+Ellipsis: Quel voisin ? 'Which neighbor ?'
+Prep–Ellipsis: À quel voisin as-tu un coup de fil à donner? 'To which neighbor do you have a call to give?'
–Prep–Ellipsis: Quel voisin as-tu un coup de fil à donner? 'Which neighbor do you have a call to give?'

(2) A: J'essaie de faire attention aux autres. B: À quel autres ?
'A: I try to pay attention to others. B: To which others?'

(3) A: Je crois que j'ai fait de la peine à un élève. B: À quel élève ?
'I think that I've given pain to a pupil. B: To which pupil?'

(4) A: J'ai parlé à un ami. B: À quel ami?
'A: I've talked to a friend. B: To which friend?'

(5) A: J'ai écrit à un vieil oncle. B: À quel oncle?
'A: I wrote to an old uncle. B: To which uncle?'

(6) A: Cette histoire a fait peur à un enfant. B: À quel enfant?
'A: This story has given fear to a child. B: 'To which child?'

(7) A: Je dois tout dire au responsable. B: À quel responsable?
'A: I must tell eveything to the manager. B: To which manager?'

(8) A: J'ai une dette à payer à mon cousin. B: À quel cousin?
'I have a debt to pay to my cousin. B: To which cousin?'

(9) A: J'ai eu des nouvelles d'un collègue. B: De quel collègue?
'A: I got news of a colleague. B: Of which colleague?'

(10) A: Je me souviens d'un ancien voisin. B: De quel voisin?
'A: I remember an old neighbor. B: Which neighbor?'

(11) A: C'est l'anniversaire d'un de mes enfants. B: De quel enfants?
'A: It's the birthday of-one of my kids. B: Of which kid?'

(12) A: Je me suis éloigné de certains parents. B: De quels parents?
'I have distanced myself from certain parents. B: From which parents?'

(13) A: Je parle du nouveau candidat. B: De quel candidat?
'A: I'm talking about the new candidate. B: About which candidate?'

(14) A: Ça fera l'affaire d'un autre. B: De quel autre?
'A: That will do for someone else. B: For which else?'

(15) A: Je me méfie de cet artiste. B: De quel artiste?
'A: I'm wary of this artist. B: - Of which artist?'

(16) A: Ce livre plait à beaucoup de jeunes. B. À quels jeunes?
'This book pleases many young people? B: Which young people?'

(17) Je pense souvent à ce chien. B. À quel chien?
'A: I think often about this dog. B: About which dog?'

(18) A: J'ai peur des voyous dans le quartier. B: De quels voyoux?
'A: I'm afraid of the thugs in the neighborhood. B: Of which thugs?'

(19) A: Je viens d'apprendre la mort d'une tante. B: De quelle tante ?
'A: I just learned the death of an aunt. B: Of which aunt?'

(20) A: Je voulais m'approcher du groupe. B: De quel groupe?
'A: I wanted to get closer to.the group. B: To which group?'

Models for Experiment 2

Parameters	Estimate	Est.Error	95% CrI	Post. Prob.
Intercept[1]	–2.82	0.23	[–3.28, –2.38]	1
Intercept[2]	–2.14	0.22	[–2.59, –1.71]	1
Intercept[3]	–1.50	0.22	[–1.94, –1.09]	1
Intercept[4]	–0.67	0.21	[–1.09, –0.26]	1
No Preposition	–2.83	0.30	[–3.44, –2.24]	1
Ellipsis	0.89	0.27	[0.39, 1.46]	0.99
No Preposition:Ellipsis	2.35	0.40	[1.57, 3.15]	1

Table 26: Bayesian ordinal model: Exp. 2.

Parameters	Estimate	Est.Error	95% CrI	Post. Prob.
Intercept[1]	–2.75	0.29	[–3.33, –2.20]	1
Intercept[2]	–2.06	0.28	[–2.63, –1.52]	1
Intercept[3]	–1.43	0.28	[–1.98, –0.89]	1
Intercept[4]	–0.59	0.27	[–1.14, –0.06]	0.98
No preposition	–2.91	0.35	[–3.61, –2.24]	1
Ellipsis	0.92	0.34	[0.28, 1.62]	1
de	0.15	0.33	[–0.51, 0.81]	0.69
No preposition:Ellipsis	2.62	0.53	[1.60, 3.67]	1
No preposition:*de*	0.13	0.32	[–0.51, 0.78]	0.67
Ellipsis:*de*	–0.03	0.40	[–0.85, 0.77]	0.53
No preposition:Ellipsis:*de*	–0.53	0.67	[–1.86, 0.79]	0.8

Table 27: Bayesian ordinal model: *à* vs. *de* - Exp. 2.

Abbreviations

NP = noun phrase, PP = prepositional phrase, P / PREP = preposition, ACC = accusative, NEG = negation, VP = verb phrase, M = masculin, F = feminine, SG = singular, REFL = reflexive, INF = infinitive, LOC = locative, adv = adverb

Data availability

The annotated corpus data, experimental items, and analysis scripts are available in an OSF repository: DOI: 10.17605/OSF.IO/AWG8U.

Funding information

The Excellence Center: Empirical Foundations of Linguistics (Labex EFL) and the PhD grant awarded to the first author from Université Paris Cité.

Acknowledgements

We would like to thank Barbara Hemforth and Joanna Nykiel, as well as the anonymous reviewers, for their helpful comments.

Part of this work was presented at the ECBAE Workshop in 2022 at the University of Bucharest, where we benefited from the feedback of the audience.

Competing interests

The authors have no competing interests to declare.

References

Abe, Jun. 2015. *The in-situ approach to sluicing*, vol. 222 (Linguistik aktuell). Amsterdam: Benjamins. DOI: https://doi.org/10.1075/la.222

Abeillé, Anne & Godard, Danièle. 2021. Les mots interrogatifs, relatifs et exclamatifs. In Abeillé, Anne & Godard, Danièle (eds.), *La grande grammaire du français*, 1105–1124.

Abeillé, Anne & Hassamal, Shrita. 2019. Sluicing in Mauritian: a fragment-based analysis. In Pinon, Christopher (ed.), *Empirical issues in syntax and semantics 12*, 1–30. Paris: CSSP.

Abeillé, Anne & Kim, Jong-Bok. 2022. *Me too* fragments in English and French: a direct interpretation approach. *The Linguistic Review* 495. 495–524. DOI: https://doi.org/10.1515/tlr-2022-2095

Abeillé, Anne & Kupfermann, Lucien. 2021. Les prépositions *à* et *de*. In Abeillé, Anne & Godard, Danièle (eds.), *La grande grammaire du français*, 820–840.

Abels, Klaus. 2017. On the interaction of P-stranding and sluicing in Bulgarian. In Mueller-Reichau, Olav & Guhl, Marcel (eds.), *Aspects of Slavic linguistics: Formal grammar,*

lexicon and communication, 1–28. Berlin, Boston: De Gruyter. DOI: https://doi.org/10.1515/9783110517873-002

Almeida, Diogo & Yoshida, Masaya. 2007. A problem for the preposition stranding generalization. *Linguistic Inquiry* 38(2). 349–362. DOI: https://doi.org/10.1162/ling.2007.38.2.349

Alshaalan, Yara & Abels, Klaus. 2020. Resumption as a sluicing source in Saudi Arabic: Evidence from sluicing with prepositional phrases. *Glossa: A Journal of General Linguistics* 5(1). 8. DOI: https://doi.org/10.5334/gjgl.841

Anand, Pranav & Hardt, Daniel & McCloskey, James. 2021. The Santa Cruz sluicing data set. *Language* 97(1). e68–e88. DOI: https://doi.org/10.1353/lan.2021.0009

Bonami, Olivier. 2021. Les fonctions du syntagme prépositionnel. In Abeillé, Anne & Godard, Danièle (eds.), *La grande grammaire du français*, 800–820.

Bürkner, Paul-Christian. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). 1–28. DOI: https://doi.org/10.18637/jss.v080.i01

Bîlbîie, Gabriela & De La Fuente, Israel & Abeillé, Anne. 2022. Factivity and complementizer omission in English embedded gapping. *Journal of Linguistics* 59(2). 389–426. DOI: https://doi.org/10.1017/S0022226722000445

Carpenter, Bob & Gelman, Andrew & Hoffman, Matthew D. & Lee, Daniel & Goodrich, Ben & Betancourt, Michael & Brubaker, Marcus & Guo, Jing & Li, Peter & Riddell, Allen. 2017. Stan: A probabilistic programming language. *Journal of Statistical Software* 76(1). 1–32. DOI: https://doi.org/10.18637/jss.v076.i01

Chung, Sandra. 2014. Sluicing and the lexicon: The point of no return. In *Annual Meeting of the Berkeley Linguistics Society*, vol. 31. DOI: https://doi.org/10.3765/bls.v31i1.896

Chung, Sandra & Ladusaw, William A. & McCloskey, James. 1995. Sluicing and logical form. *Natural Language Semantics* 3. 239–282. DOI: https://doi.org/10.1007/BF01248819

Culicover, Peter W. & Jackendoff, Ray. 2005. *Simpler syntax*. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199271092.001.0001

Dagnac, Anne. 2018. French. In van Craenenbroeck, Jeroen & Timmermann, Tanja (eds.), *The Oxford handbook of ellipsis*, chap. 30, 785–814. Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198712398.013.35

Delaveau, Annie. 2021. Les phrases interrogatives. In Abeillé, Anne & Godard, Danièle (eds.), *La grande grammaire du français*, chap. XII, 1402–1437.

Destruel, Emilie. 2012. The French c'est-cleft: empirical studies of its meaning and use. In Pinon, Christopher (ed.), *Empirical issues in syntax and semantics*, vol. 9, 95–112.

Fortin, Catherine. 2007. *Indonesian sluicing and verb phrase ellipsis: Description and explanation in a minimalist framework*: University of Michigan dissertation.

Gelman, Andrew & Carlin, John B. & Stern, Hal S. & Dunson, David B. & Vehtari, Aki & Rubin, Donald B. 2013. *Bayesian data analysis*. London: Chapman and Hall/CRC 3rd edn. DOI: https://doi.org/10.1201/b16018

Geweke, John. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian statistics*, 169–193. Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198522669.003.0010

Gibson, Edward & Fedorenko, Evelina. 2013. The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes* 28(1–2). 88–124. DOI: https://doi.org/10.1080/01690965.2010.515080

Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.

Ginzburg, Jonathan & Fernández, Raquel. 2010. Computational models of dialogue. In Clark, Alexander & Fox, Chris & Lappin, Shalom (eds.), *Handbook of computational linguistics and natural language processing*, 429–481. Chichester: Wiley Blackwell. DOI: https://doi.org/10.1002/9781444324044.ch16

Ginzburg, Jonathan & Kim, Jong-Bok. 2023. Exclaiming non-sententially: exclamative sluices and the (non-elliptical) nature of ellipsis constructions. *Glossa: A Journal of General Linguistics* 8(1). DOI: https://doi.org/10.16995/glossa.9639

Ginzburg, Jonathan & Miller, Philip. 2018. Ellipsis in Head-Driven Phrase Structure Grammar. In van Craenenbroeck, Jeroen & Temmerman, Tanja (eds.), *The Oxford handbook of ellipsis*, 75–121. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198712398.013.4

Ginzburg, Jonathan & Sag, Ivan A. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives*. Stanford, CA: CSLI Publications.

Gotowski, Megan. 2022. Quoi-sluices in French. *Glossa: A Journal of General Linguistics* 7(1). DOI: https://doi.org/10.16995/glossa.5793

Hassen, Amal & Ginzburg, Jonathan & Abeillé, Anne. 2023. French exclamative sluices: a corpus study. In *Experimental and corpus-based approaches to ellipsis*. University of Massachusetts, Amherst, USA. Presentation.

Hawkins, John A. 1999. The relative order of prepositional phrases in English: Going beyond manner–place–time. *Language Variation and Change* 11. 231 – 266. DOI: https://doi.org/10.1017/S0954394599113012

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199252695.001.0001

Hofmeister, Philip & Jaeger, T. Florian & Sag, Ivan A. & Arnon, Inbal & Snider, Neal. 2007. *Locality and accessibility in wh-questions* 185–206. Berlin, New York: De Gruyter Mouton. DOI: https://doi.org/doi:10.1515/9783110198621.185

Lefeuvre, Florence. 2020. Les interrogatives partielles dans un corpus de théâtre contemporain. *Langages* 217(1). 23–38. DOI: https://doi.org/10.3917/lang.217.0023

Levin, Lori S. 1982. Sluicing: A lexical interpretation procedure. In Bresnan, Joan (ed.), *The mental representation of grammatical relations*, 590–654. Cambridge, MA: MIT Press.

Lewis, Richard L. & Vasishth, Shravan. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29(3). 375–419. DOI: https://doi.org/10.1207/s15516709cog0000_25

Martin, Andrea E. & McElree, Brian. 2011. Direct-access retrieval during sentence comprehension: evidence from sluicing. *Journal of Memory and Language* 64(4). 327–343. DOI: https://doi.org/10.1016/j.jml.2010.12.006

McCulloch, Charles & Neuhaus, John. 2005. Generalized linear mixed models. *Encyclopedia of Biostatistics* 4. DOI: https://doi.org/10.1002/0470011815.b2a10021

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and identity in ellipsis*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780199243730.001.0001

Molimpakis, Emilia. 2019. *Accepting preposition-stranding under sluicing cross-linguistically; A noisy-channel approach*. London: UCL dissertation.

Müller, Stefan & Abeillé, Anne & Borsley, Robert D. & Koenig, Jean-Pierre (eds.). 2021. *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax 9). Berlin: Language Science Press. DOI: https://doi.org/10.5281/zenodo.5543318

Nykiel, Joanna. 2013. Clefts and preposition omission under sluicing. *Lingua* 123. 74–117. DOI: https://doi.org/10.1016/j.lingua.2012.10.011

Nykiel, Joanna. 2017. Preposition stranding and ellipsis alternation. *English Language and Linguistics* 21(1). 27–45. DOI: https://doi.org/10.1017/S1360674315000477

Nykiel, Joanna & Hawkins, John. 2020. English fragments, minimize domains, and minimize forms. *Language and Cognition* 12. 1–33. DOI: https://doi.org/10.1017/langcog.2020.6

Nykiel, Joanna & Kim, Jong-Bok. 2022a. Fragments and structural identity on a direct interpretation approach. *Journal of Linguistics* 73–109. DOI: https://doi.org/10.1017/S0022226720000420

Nykiel, Joanna & Kim, Jong-Bok. 2022b. On the grammaticality of morphosyntactically reduced remnants in Polish sluicing. *Linguistics* 60(1). 177–213. DOI: https://doi.org/10.1515/ling-2021-0071

Obenauer, Hans-Georg. 1976. *Etudes de syntaxe interrogative du français*. Berlin, Boston: Max Niemeyer Verlag. DOI: https://doi.org/doi:10.1515/9783111340364

Parker, Dan. 2019. Cue combinatorics in memory retrieval for anaphora. *Cognitive Science* 43(3). e12715. DOI: https://doi.org/10.1111/cogs.12715

Philippova, Tatiana. 2014. P-omission under sluicing, [P clitic] and the nature of P-stranding. In *Proceedings of ConSOLE XXII*. 133–155.

Pollard, Carl & Sag, Ivan A. 1994. *Head-driven phrase structure grammar*. University of Chicago Press.

Reinhardt, Janina. 2019. *Regularity and variation in French direct interrogatives*: University of Konstanz dissertation.

Rodrigues, Cilene & Nevins, Andrew & Vicente, Luis. 2009. Cleaving the interactions between sluicing and preposition stranding. In *Romance languages and linguistic theory (2006)*, 175–198. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.303.11rod

Ross, John. 1969. Guess who? In *Chicago Linguistic Society, vol. 5*, 252–286.

Sag, Ivan A. & Nykiel, Joanna. 2011. Remarks on sluicing. In Müller, Stefan (ed.), *Proceedings of the 18th international conference on Head-Driven Phrase Structure Grammar*, 188–208. Stanford, CA: CSLI Publications. DOI: https://doi.org/10.21248/hpsg.2011.11

Sato, Yosuke. 2011. P-stranding under sluicing and repair by ellipsis: Why is Indonesian (not) special? *Journal of East Asian Linguistics* 20. 339–382. DOI: https://doi.org/10.1007/s10831-011-9082-3

Smirnova, Anastasiia & Abeillé, Anne. 2021. The French question particles *ça* and *donc*: A corpus study. *Linguistic Research* 38(2). 239–269. DOI: https://doi.org/10.17250/khisli.38.2.202106.003

Szczegielniak, Adam. 2008. Islands in sluicing in Polish. In Abner, Natasha & Bishop, Jason (eds.), *Proceedings of the 27th West Coast conference on formal linguistics*, 404–412. Somerville, MA, USA: Cascadilla Proceedings Project.

Tseng, Jesse. 2021. La préposition et le syntagme prépositionnel. In Abeillé, Anne & Godard, Danièle (eds.), *La grande grammaire du français*, chap. VII, 747–860.

van Craenenbroeck, Jeroen. 2004. *Ellipsis in Dutch dialects*, vol. 096 (LOT Dissertation Series). Leiden: Universiteit van Leiden (Doctoral dissertation).