

Chapter 6

Subject inversion in French object relatives: What's your preference?

Céline Pozniak

Université Paris 8, Structures Formelles du Langage, CNRS

Anne Abeillé

Université de Paris, Laboratoire de linguistique formelle

Barbara Hemforth

Université de Paris, Laboratoire de linguistique formelle, CNRS

Subject inversion in French is usually considered to be optional (Le Bidois 1952; Kayne & Pollock 1978) and more costly than variants with preverbal subject. As the result of verb movement (Hulk & Pollock 2001), it is claimed to demand higher processing cost (Holmes & O'Regan 1981). However, some studies suggest that subject inversion in relative clauses may even be favoured by certain semantic or heaviness constraints (Fuchs 2006; Marandin 2011). In this paper, we take an empirical approach to this question. In our corpus study using the French Treebank described in Abeillé et al. (2019), we found that subject inversion in object relatives can be as frequent as cases without inversion. We also found that inversion is preferred with longer subjects and shorter and non-agentive verbs. This pattern was confirmed in an acceptability judgement experiment as well as in a self-paced reading experiment. Thus, object relatives with and without inversion are not merely stylistic variants (i.e. two equivalent syntactic ways of expressing one meaning), but are more or less preferred depending on their properties. Our results are compatible with semantic accounts of relative clause processing (Mak et al. 2006; Traxler et al. 2002).



1 Introduction

French object relative clauses (ORs) are introduced by *que* and may have a preverbal (1) or a postverbal subject (2) (Le Bidois 1952; Kayne & Pollock 1978).

- (1) Object relative with preverbal subject
Le médecin [que l'avocat_{subj} connaît] aime courir.
the physician that the lawyer knows likes run
'The physician [that the lawyer knows] likes running.'
- (2) Object relative with postverbal subject
Le médecin [que connaît l'avocat_{subj}] aime courir.
the physician that knows the lawyer likes run
'The physician [that the lawyer knows] likes running.'

In this paper, we will address the question of the status of these two types of object relatives. Are they just stylistic variants or do they differ with respect to specific properties beyond subject-verb order? We will conclude from our empirical studies that specific properties make each of them more or less felicitous, thus contributing to the many (variants) to one (meaning) aspect of this volume. We will also show that the choice of a (one) particular object relative structure depends on the combination of (possibly many) factors.

Object relatives with a preverbal subject (OR-inv) are considered canonical while object relatives with a postverbal subject (OR+inv) are often seen as marked and as a stylistic variant (especially for written French). According to semantic and pragmatic theories, the postverbal subject is generally thought of as having properties different from a preverbal subject: the postverbal subject is more likely to be indefinite and focal (Lahousse 2011) and/or long and not agentive (Fuchs 2006; Marandin 2011). Syntactic theories usually consider subject inversion as more complex, and the result of verb movement (Déprez 1990; Hulk & Pollock 2001) or specific linearization rules (Bonami & Godard 2001). However, an inversion analysis may be favored in relativized minimality (Rizzi 1990; Friedmann et al. 2009), where a preverbal animate subject may interfere with the filler-gap dependency (3), which can be avoided with a postverbal subject (4).

- (3) OR-inv
le médecin [que l'avocat connaît __]
the physician that the lawyer knows __
'the physician [that the lawyer knows __]'

6 Subject inversion in French object relatives: What's your preference?

(4) OR+inv

le médecin [que connaît __ l'avocat]
the physician that knows __ the lawyer
'the physician [that knows __ the lawyer]'

As for processing theories, differences in linear distance predict that OR+inv should be easier to process than OR–inv. According to dependency locality theory (Gibson 2000), the linear distance between the filler (*que*) and the object gap is shorter in OR+inv (4), leading to a lower storage memory cost than in OR–inv (3). The distance is also shorter between the filler (*que*) and the relative clause verb (*connaît*) if we consider traceless theories of extraction, with a SLASH feature on the verb (Bouma et al. 2001; Sag 2010).

Two (complementary) approaches will be applied in order to put these theories and their conflicting predictions to an empirical test. The first is to look at large corpora, to see how frequent subject inversion is in object relative clauses, and which factors may favour or disfavour it. The second way is to conduct experiments that enable us to test these factors in a controlled environment. In this paper, we will associate corpus studies and experiments in order to provide converging evidence. Previous corpus studies on newspaper texts found a 41% inversion rate for French relative clauses (Fuchs 1997). Fuchs (1997) conducted a corpus study on one issue of the newspaper *Le Monde* and found that relatives with a nominal subject (not only object relative clauses but including those with *dont* 'whose', *où* 'where') were more frequent with a preverbal subject than with a postverbal subject. Based on the frequency distribution in her corpus, she suggests that subject inversion is favoured when the subject is not agentive, inanimate and definite, when it is longer than the verb phrase and when the verb is not agentive. While our own corpus studies will be strongly inspired by Fuchs's analysis, we will restrict our analysis to object relative clauses but also go beyond their approach by testing the different constraints as well as their intercorrelations using state of the art inferential statistics.

On the processing side, previous experimental studies found that object relatives with postverbal subject were more difficult to understand than those with preverbal subject. In an eye-tracking experiment, Holmes & O'Regan (1981) examined participants' eye movements while they read sentences with animate subjects and objects and reversible verbs (*dessiner* 'draw', *voir* 'see', ...). Pozniak & Hemforth (2015), also using animate subjects and reversible verbs, conducted an eye-tracking experiment using the visual world paradigm, where participants listened to sentences including subject relative clauses as well as object relative

clauses with postverbal (6) or with preverbal subject (5) while they saw two pictures on a computer screen, one compatible with a subject relative clause interpretation (where the princess draws the fencer) and the other one compatible with an object relative clause interpretation (where the fencer draws the princess). The participants' task was to look at the "correct" picture, i.e. the picture compatible with the sentence they heard, on a computer screen. More and earlier fixations on the correct picture are interpreted as evidence for easier processing in this paradigm.

(5) OR–inv

Prière de trouver la princesse correcte, c'est-à-dire la belle
please find the princess correct, that is to say the beautiful
princesse [que l'escrimeur dessine] sur l'image.
princess that the fencer draws on the picture

'Please find the correct princess, that is to say the beautiful princess [that the fencer draws] on the picture.'

(6) OR+inv

Prière de trouver la princesse correcte, c'est-à-dire la belle
please find the princess correct, that is to say the beautiful
princesse [que dessine l'escrimeur] sur l'image.
princess that draws the fencer on the picture

'Please find the correct princess, that is to say the beautiful princess [that the fencer draws] on the picture.'

Subject relative clauses, which were tested in both the Holmes & O'Regan (1981) reading experiment and the visual world eye-tracking experiment by Pozniak & Hemforth (2015), were processed faster and led to more fixations on the correct image than both object relative clause variants. However, both studies also found that OR+inv were more difficult to process than OR–inv, contrary to what is predicted by processing theories like DLT or syntactic theories like relativized minimality. Processing data from experiments also provide evidence for more fine-grained semantic constraints: Frauenfelder et al. (1980) compared OR+inv with reversible verbs (*connaître* 'know') and animate objects (7) vs. with non-reversible verbs (*publier* 'publish') and inanimate objects (8). Using a phoneme monitoring task (where participants have to press a button as soon as they hear a particular phoneme), they found that OR+inv was easier with inanimate objects (8) than with animate objects. However, their data cannot tell us whether these factors are specific to object relatives with a postverbal subject

6 Subject inversion in French object relatives: What's your preference?

(OR+inv) or concern object relatives in general, since this study did not include a comparison of ORs with postverbal and preverbal subject.

- (7) Le savant [que connaît le docteur] travaille dans une université
the scientist that knows the doctor works in a university
moderne.
modern
'The scientist [that the doctor knows] works in a modern university.'
- (8) Les articles [que publie la revue] demandent une lecture attentive.
the articles that publishes the journal ask for a reading careful
'The articles [that the journal publishes] demand a careful reading.'

Using eye-tracking while reading and self-paced reading paradigms, Baudiffier et al. (2011) directly compared OR–inv and OR+inv. They found that OR–inv were generally easier to process than OR+inv with inanimate subjects and animate objects. These experiments mainly focused on the role of animacy for the two types of relatives, but did not test length or other semantic or pragmatic properties that have been suggested to play a role as well (e.g. Fuchs 1997).

In the following sections, we present a new corpus study, based on a syntactically annotated corpus (the French Treebank, Abeillé et al. 2019), and two new experiments. We found that OR+inv can be as frequent as OR–inv, furthermore, they can be as acceptable as OR–inv in two controlled experiments manipulating semantic/pragmatic properties.

2 Inversion in object relatives: a corpus study

We searched for object relatives in the French Treebank (Abeillé et al. 2019)¹, which comprises around 21550 sentences from newspaper texts (*Le Monde* from 1990 to 1993). We extracted object relatives (with *que*) with a nominal subject and obtained 298 ORs in total, 149 of which had a postverbal subject. In order to have a fully parallel comparison for subject inversion, we excluded cleft constructions, appositive relatives, obligatory relatives after demonstratives (*ce que*), relatives with a pronominal subject (*cela* 'this', *certain*s 'some') and some errors (clitic subjects). This leaves 178 object relatives, 90 with subject inversion as in (9), and 88 without as in (10).

¹Available on <http://ftb.linguist.univ-paris-diderot.fr/>.

- (9) ..., le conseil régional de Picardie a pris position sur les problèmes
..., the regional council of Picardy has taken a stance on the problems
[que connaît l'audiovisuel public dans la région].
that encounters the audiovisual media in the region
'..., the regional council of Picardy has taken a stance on the problems
[that encounters the audiovisual media in the region].'
- (10) Le rôle d'intimidation [que l'armée rouge aura après la
The role of intimidation that the army red will have after the
démilitarisation de l'Allemagne]...
demilitarization of Germany
'The role of intimidation [that the Red Army will have after the
demilitarization of Germany] ...'

2.1 Annotation criteria

We annotated our 178 object relatives mainly using criteria from Fuchs (Fuchs 1997; 2006). As illustrated in Table 6.1, we annotated both relatives with animacy of the subject and the object, relative length between the verb and the object as well as length of the relative clause, definiteness of the subject and the object, thematic roles, negation and position of the relative in the sentence.^{2,3}

2.2 Results

To analyze the corpus data, we ran logistic regression models using the `glmer` function in the `lme4` R package from Bates et al. (2015). The dependent variable was subject inversion, coded as 1 for postverbal subject (subject inversion) and 0 for preverbal subject (no subject inversion). The criteria animacy, thematic roles, verb agentivity, verb/subject length, relative clause length, definiteness, negation, and position of the relative clause were included as predictors. They were all coded using mean centering.

We used logistic regressions with simple intercepts (Jaeger 2008) to test whether there is a general difference in frequency between the two types of ORs. No significant difference between ORs with preverbal subject and ORs with postverbal subject could be established ($z = -0.15$, $p > 0.1$).

As for the factors we annotated in our corpus (subject definiteness, negation, animacy, verb agentivity, subject/verb length, relative clause length, position of

²Fuchs (2006) did not mention the position of the relative and the negation as criteria for subject inversion.

³The corpus with the annotated data can be found on <https://osf.io/k97pu/>.

6 Subject inversion in French object relatives: What's your preference?

Table 6.1: Annotation for object relatives in the corpus

Criteria	Annotation
Relative	With preverbal subject With postverbal subject
Animacy	Animate subject and object Inanimate subject and object Animate subject and inanimate object Inanimate subject and animate object
Thematic roles	Intentional subject, affected theme No intentional subject, no affected theme Intentional subject, no affected theme No intentional subject, affected theme
Verb agentivity	Agentive verb ('to want', 'to fight') Non-agentive verb ('to represent', 'to have')
Verb/subject length (syllables)	Verb longer than subject Verb shorter than subject Verb as long as subject
Relative clause length	Relative with subject and verb only Relative with more constituents than verb and subject
Definiteness	Definiteness of the object of the relative Definiteness of the subject of the relative
Negation	Presence of negation Absence of negation
Position of the relative	Inside the main clause subject Inside the main clause object Inside another constituent

the relative and object definiteness), we decided to run a logistic regression model combining them all as predictors. It makes sense to analyze them in one model because this allows us to establish the independent contribution of highly intercorrelated factors (the factor verb agentivity is linked to thematic roles, for example). The model including only the statistically significant predictors is presented in Table 6.2. Positive estimates correspond to an increase in the number of postverbal subjects such that OR+inv is more likely for short ORs with definite objects, non-intentional subjects and with verbs that are shorter than the subject.

Table 6.2: Significant factors with logistic regression model for subject inversion. The intercept corresponds to indefinite object/short relative/verb longer than subject/non-intentional subject.

Fixed effects	E	SE	<i>z</i>	<i>p</i> <
Intercept	-0.03606	0.19566	-0.184	0.1
Definite object	1.25588	0.47194	2.661	0.01
Long relative	1.04969	0.41009	2.560	0.05
Verb shorter than subject	1.48045	0.63099	2.346	0.05
Intentional subject	-1.45056	0.55517	-2.613	0.01

The factors that did not significantly contribute to the model are the following: position of the relative clause (inside main clause subject or main clause object), animacy of subject or object, affected theme and verb agentivity. However, the fact that verb agentivity did not contribute significantly to the model can be explained by the fact that an agentive verb needs an intentional subject, so these predictors are highly correlated ($r = 0.6, t = 10.08, p < 0.001$). The independent variables that significantly influence subject inversion are shown in Figure 6.1.

2.3 Interim discussion

The corpus study shows that subject inversion can be as frequent as preverbal subjects for object relatives with a nominal subject. It also shows that object relatives with pre- and postverbal subjects have different properties: A preverbal subject is more frequent when the object is indefinite and the subject intentional (implying agentivity of the verb), and when the relative is short (including only the subject and the verb) and with a subject that is shorter than the verb. A postverbal subject is more frequent when it is longer than the verb, when it is non-intentional and has a definite object, and when the relative clause is long. The

6 Subject inversion in French object relatives: What's your preference?

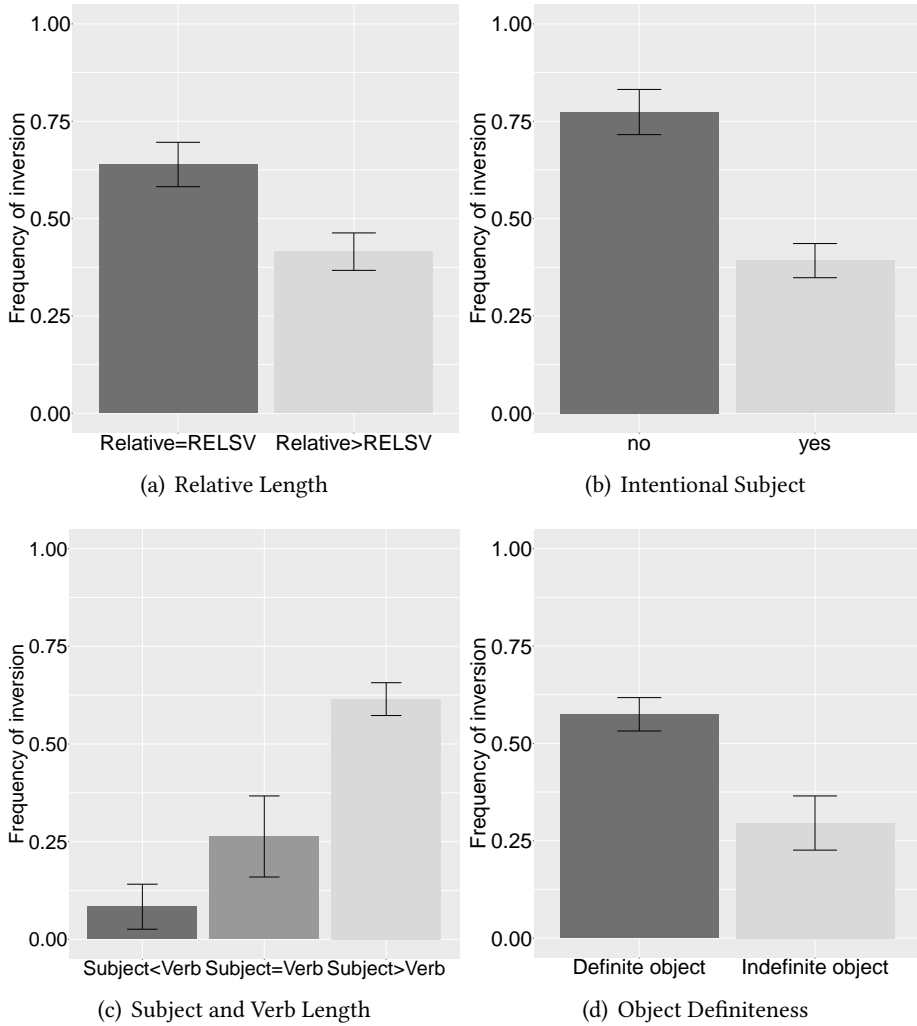


Figure 6.1: Significant factors on subject inversion (corpus study)

relative length effect (verb shorter than the subject) is in line with processing theories like DLT which predict processing difficulty with a long intervening subject between *que* and the verb (or between *que* and the postverbal gap), thus favouring subject inversion. It may also be explained by a more general tendency to put longer (heavier) constituents at the end of the sentence (Behaghel 1909; Wasow 2002).

Prosodic factors may also play a role in explaining why a short relative (subject and verb only) favours inversion. If one considers the general tendency to have balanced prosodic constituents, and that a prosodic boundary usually occurs between the subject and the verb (Di Cristo 2016), (11) has a less natural prosodic structure than (12) (with a longer relative) or (13) (with subject inversion).

- (11) les problèmes (que l'audiovisuel) (connaît).
the problems that the audiovisual knows
'the problems that the audiovisual knows...'
- (12) les problèmes (que l'audiovisuel) (connaît dans la région)
the problems that the audiovisual knows in the region
'the problems that the audiovisual knows in the region'
- (13) les problèmes (que connaît) (l'audiovisuel)
the problems that knows the audiovisual
'the problems that the audiovisual knows'

The definiteness effect can be explained by the relative discourse status of the subject and the object: in the context of an object relative clause, a definite object is more topical than the subject, and a less topical subject is more likely to be postverbal (Kampers-Manhe et al. 2004). The effect of subject intentionality or agentivity of the verb is in line with Marandin (2011) and Bonami & Godard (2001), suggesting that postverbal subjects lose their dynamic and agentive properties.

The corpus analysis thus shows that semantic/pragmatic features differ for object relatives with a preverbal subject and object relatives with a postverbal subject. Verb semantics, length and definiteness seem to play an important role, meaning that subject inversion in object relative clauses is not merely a stylistic variant. However, corpus studies suffer from the problem that the factors of interest are often intercorrelated, as we have seen for the intentionality of the subject and agentivity of the verb. Also, the constraints we annotated may be affected by some other variables co-varying in the corpus that we have not taken into account. In order to have a more controlled picture of the usage difference, we therefore tested these factors with two experimental studies.

3 Subject inversion in object relatives: An acceptability judgement task

In order to better understand the use of object relatives with preverbal and postverbal subjects, we ran an acceptability judgement task manipulating some of the factors found in the corpus study.

3.1 Material

We manipulated three variables: subject position (preverbal/postverbal), verb semantics (agentive/non-agentive) and subject length (long/short). As for verb semantics, pairs of agentive and non-agentive verbs were created with the same number of syllables. Verb agentivity is highly correlated with subject intentionality, as we saw in the preceding section, and easier to control in the experimental materials. Concerning subject length, the subject was treated as short when it was only composed of the article and the noun, whereas it was considered as long when a noun complement and/or an adjective was included.

Thirty-two items were created with four items per condition (Latin square design), as shown in Table 6.3. Forty-four fillers were added as distractors. The subject was always animate (humans, human groups or nouns symbolizing a collective group like a firm or a country) and the object inanimate, which favors object relative processing across the two variants (Frauenfelder et al. 1980; Mak et al. 2006). The experimental materials were inspired by the sentences from the corpus study. All relatives were short (relativizer, subject, verb), the object and the subject were definite, and all relatives modified the main clause subject. Subject length was manipulated by adding a modifier or a complement to the subject noun; the agentivity condition was an alternation between two related verbs, one non-agentive like *cost* making the subject non-intentional, and one agentive like *pay* making the subject intentional.⁴

3.2 Participants

Eighty French native speakers (56 women, mean age: 36 years, $\sigma = 18$) volunteered to participate in the experiment, which was run on IbexFarm (Drummond 2013). They were recruited via the RISC (<http://www.risc.cnrs.fr>) platform.

⁴The materials and the entire analysis can be found on <https://osf.io/k97pu/>.

Table 6.3: Examples of sentences used in the acceptability judgement task: subject inversion (Inv \pm), verb agentivity (Ag \pm), subject length (Long \pm)

Inv	Ag	Long	Example
-	-	-	Le prix astronomique que la firme coûte énerve considérablement les dirigeants. the price astronomical that the company costs irritates considerably the managers
+	-	-	'The astronomical price that the company costs considerably irritates the managers.'
-	-	+	Le prix astronomique que coûte la firme énerve considérablement les dirigeants. the price astronomical that costs the company irritates considerably the managers
+	-	+	'The astronomical price that the company costs considerably irritates the managers.'
-	-	+	Le prix astronomique que la firme agroalimentaire coûte énerve considérablement les dirigeants. the price astronomical that the company agri-food costs irritates considerably the managers
+	-	+	'The astronomical price that the agri-food company costs considerably irritates the managers.'
-	+	-	Le prix astronomique que coûte la firme agroalimentaire énerve considérablement les dirigeants. the price astronomical that costs the company agri-food irritates considerably the managers
+	+	-	'The astronomical price that the agri-food company costs considerably irritates the managers.'
-	+	+	Le prix astronomique que paie la firme énerve considérablement les dirigeants. the price astronomical that pays the company pays irritates considerably the managers
+	+	+	'The astronomical price that the company pays considerably irritates the managers.'
-	+	+	Le prix astronomique que la firme agroalimentaire paie énerve considérablement les dirigeants. the price astronomical that the company agri-food pays irritates considerably the managers
+	+	+	'The astronomical price that the agri-food company pays considerably irritates the managers.'

3.3 Procedure

Participants read sentences on a computer screen at a location of their choice. They had to judge the acceptability of each sentence on a scale from 1 (not at all acceptable) to 10 (fully acceptable). The experiment lasted about 15 minutes.

3.4 Results

We analyzed the acceptability judgements with generalized linear mixed models (Baayen et al. 2008), using the `lmer` function in R with the `lme4` package from Bates et al. (2015). As predictors, we included subject length (short, long), verb semantics (agentive, non-agentive) and subject position (postverbal, preverbal). We applied mean centered coding for all predictors. Acceptability judgements are the dependent variable in the model. Participants and items were included as random variables. We used a “maximal model”, by including by-participants and by-items random intercepts as well as random slopes for all the relevant fixed factors (Barr et al. 2013). We enforced zero correlations between random effects in order to avoid overparameterization or false convergence (Bates et al. 2015). Figure 6.2 illustrates the effects of the three independent variables on judgements.

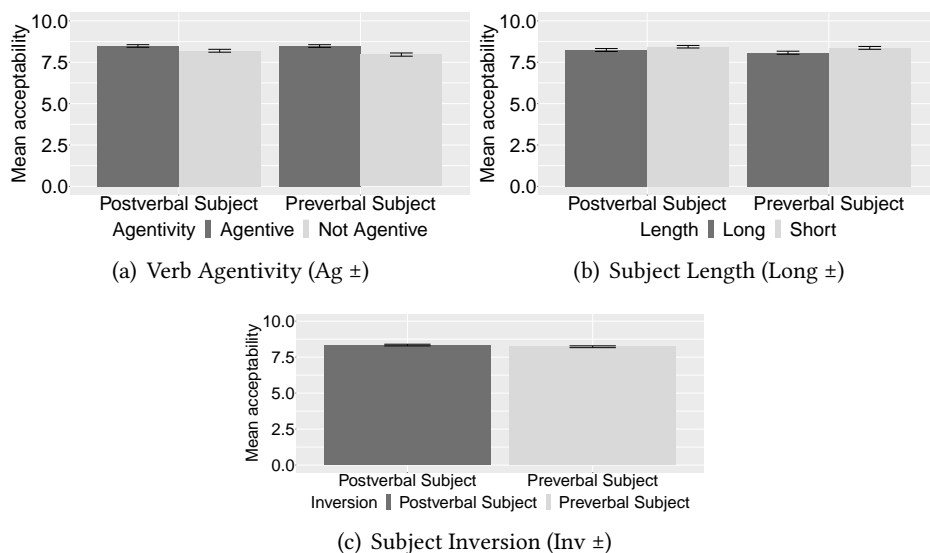


Figure 6.2: Influence of verb agentivity, subject length and subject inversion on acceptability

When looking at subject position, the model shows that relatives both with and without inversion are rated well (higher than 8/10). Object relatives with postverbal subject are considered marginally more acceptable than object relatives with preverbal subject (8.34 vs. 8.22, $t = 1.751$, $p = 0.08$).

Main effects of agentivity ($t = 2.925$, $p < 0.01$) and subject length ($t = 3.322$, $p < 0.01$) were found, meaning that sentences are rated better when the verb is agentive and the subject short. We also found an interaction between those two variables ($t = -2.571$, $p < 0.05$): sentences with short subjects received higher ratings when the verb is not agentive. Otherwise, no significant interaction between the three factors was found.

3.5 Interim discussion

The acceptability judgements showed that relatives both with preverbal subject and with postverbal subject are well acceptable, which is in line with what was found in the corpus study (the two possibilities were used about equally often). The results showed that object relatives with postverbal subject are in fact judged slightly better, contrary to the results from previous experiments (Holmes & O'Regan 1981; Pozniak & Hemforth 2015), which only considered ORs with animate objects and reversible verbs. This can be explained by the fact that all objects were definite in our material and all relatives were short, meaning that, as shown in the corpus study, all our materials already realized two of the constraints that make object relatives with postverbal subject favored, compared to object relatives with preverbal subject.

No interaction was found, however, between subject position, agentivity, and subject length. One reason for this lack of an effect could be that both relatives are perfectly grammatical and that participants chose a rather conscious and metalinguistic approach to the task, which may have obscured subtle differences between object relatives with preverbal subject and object relatives with postverbal subject. In order to have a more fine-grained analysis of processing at every point in the sentence as well as more spontaneous data, we decided to run a self-paced reading experiment with the same material.

4 Subject inversion in object relatives: A self-paced reading experiment

Our acceptability judgement task only showed global acceptability ratings of object relatives with preverbal and postverbal subject. This paradigm cannot differentiate which part of the sentence makes a relative clause more or less acceptable and natural. That is why we ran a self-paced reading experiment as well.

4.1 Material

The items used were the same as in the acceptability study and the conditions were the same as well. 16 fillers were added, as well as 24 comprehension questions: 14 questions for experimental items and 10 for fillers, around 50% of all the trials.

4.2 Participants

Forty-nine French native speakers (36 women, mean age: 29 years, $\sigma = 10$) participated online in the experiment via the IbexFarm platform. They were recruited on the RISC platform.

4.3 Procedure

Participants read sentences on a computer screen at a place of their choice. Sentences appeared one word at a time in a moving window paradigm (participants had to press the spacebar each time to make the following word appear). After reading each sentence, they had to judge its acceptability on a scale from 1 (not at all acceptable) to 10 (fully acceptable). They had to answer a question about the previous sentence in around 50% of the trials. The experiment lasted about 15 minutes.

4.4 Results

Results were analyzed with generalized linear mixed models using the lmer function. Independent variables were again subject length, verb semantics, and subject position, with mean centered coding applied for all predictors. Random variables were participants and items. The dependent variable was the mean reading time on every region of the sentence. Models take into account log-transformations of reading times as well as general length.

Again, we used a “maximal model”, by including by-participants and by-items random intercepts as well as random slopes for all the relevant fixed factors (Barr et al. 2013). We enforced zero correlations between random effects in order to avoid overparameterization or false convergence (Bates et al. 2015).⁵

4.4.1 Comprehension questions

The percentage of correct answers to comprehension questions was above 90% in all eight conditions. Logistic regression models do not show a significant difference between the conditions.

⁵The entire analysis can be found on <https://osf.io/k97pu/>.

4.4.2 Mean reading times

For the statistical analysis, we divided the items into six regions of interest: antecedent of the relative (object), relativizer, subject/verb, verb/subject, main clause verb, end of the sentence. This is illustrated in Table 6.4 for conditions without subject inversion and in Table 6.5 with subject inversion. Figure 6.3 represents the results for all regions and all conditions.

Table 6.4: Regions without subject inversion

1	Le prix astronomique	The price astronomical
2	que	that
3	la firme (agroalimentaire)	the company (agrifood)
4	coûte/paie	costs/pays
5	irrite	irritates
6	considérablement les dirigeants.	considerably the managers.

Table 6.5: Regions with subject inversion

1	Le prix astronomique	The price astronomical
2	que	that
3	coûte/paie	costs/pays
4	la firme (agroalimentaire)	the company (agrifood)
5	irrite	irritates
6	considérablement les dirigeants.	considerably the managers.

In this paper, we focus on region 5 (main clause verb) since this region is identical across conditions. It appears after the relative clause and may show differences in processing. Figure 6.4 represents mean residual reading times for the main clause verb (region 5).

In region 5 (Figure 6.4), we found a general effect of verb semantics ($t = 2.26$, $p < 0.05$): reading non-agentive verbs took longer than reading agentive verbs. An interaction between subject length and verb semantics is also observed ($t = 2.59$, $p < 0.05$): reading times are longer with short subjects and non-agentive verbs compared to long subjects and agentive verbs. We also found a marginal effect of subject position ($t = -1.81$, $p = 0.08$): relatives with a postverbal subject seem to be read faster than relatives with a preverbal subject.

6 Subject inversion in French object relatives: What's your preference?

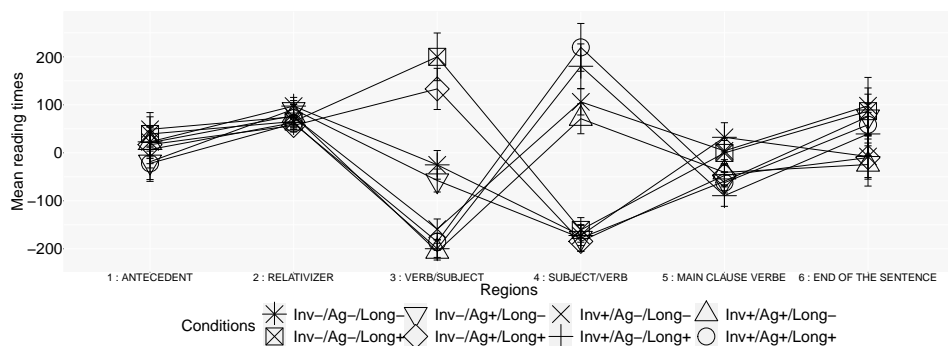


Figure 6.3: Residual reading times for the eight conditions in each region of the sentences

Interestingly, when subsetting relatives with postverbal subject, we found an effect of subject length ($t = -2.06, p < 0.05$) as well as an interaction between verb agentivity and subject length ($t = -3.08, p < 0.01$).⁶ This means that when the verb is not agentive, relatives are read faster with a long subject rather than with a short subject, whereas there is no difference in relatives with agentive verbs. Non-agentive verbs with long subjects correspond to the most felicitous context in the corpus study for relatives with postverbal subject. As for relatives with preverbal subject, an effect of verb semantics is found ($t = -2.70, p < 0.05$): relatives with preverbal subject are read faster when the verb is agentive.

4.5 Interim discussion

As in the acceptability judgements, object relatives with a postverbal subject are no more difficult to process than object relatives with a preverbal subject in our study – contra Holmes & O'Regan (1981), Pozniak & Hemforth (2015). The self-paced reading study also showed an effect of length and semantics in the main clause verb region (just after the relative): relatives with a short, preverbal subject and an agentive verb are read faster than when the verb is not agentive. Reading times in the same region showed that relatives with a long, postverbal subject and a non-agentive verb are easier to process than when the subject is short. Both of these combinations echo the specific conditions for pre- and postverbal subjects

⁶We had to remove the interaction of the fixed factors in the random variables to make the model converge: `m1=lmer(log(reaction) ~ Sémantique + Longueur + length + (Sémantique*Longueur + 1||sujet) + (Sémantique + Longueur + 1||item), data=inversion, control = lmerControl(optimizer = "optimx", calc.derivs = FALSE,optCtrl = list(method = "nlnmb", starttests = FALSE, kkt = FALSE)))`

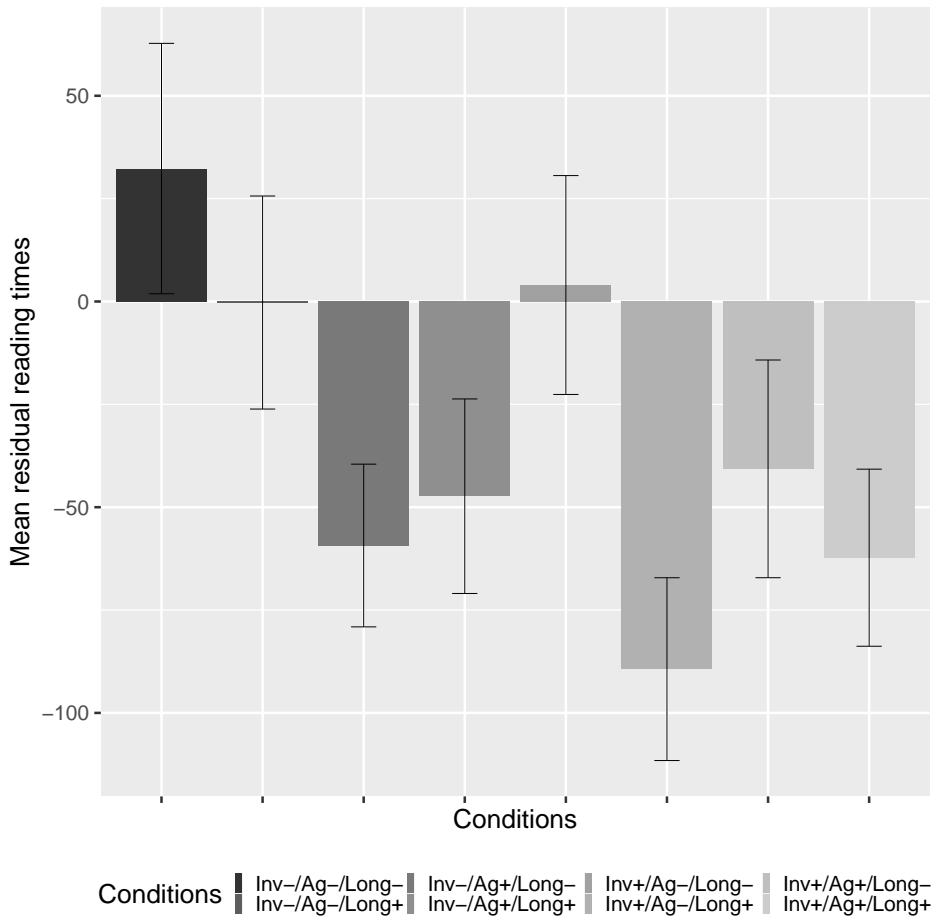


Figure 6.4: Mean residual reading times in the main verb region. Error bars represent standard errors.

that we identified in the corpus analysis. Overall, the experiment suggests that subject length and verb semantics play a role in the position of the subject in object relative clauses.

5 Discussion and conclusions

Predictions and data on the usage of object relative clauses with postverbal subject in French are inconsistent in the linguistic as well as in the psycholinguistic literature. There seems to be some general understanding that they are marked, more complex, less frequent and harder to understand than object relatives with preverbal subject. This general understanding, however, goes against predictions of some syntactic approaches (e.g. relativized minimality, Rizzi 1990) as well as some psycholinguistic processing theories (e.g. DLT, Gibson 2000). Previous qualitative corpus studies (Fuchs 1997) as well as psycholinguistic experiments point to an even more complex picture where a variety of constraints has to be taken into account.

This inconsistency in the literature led us to the hypothesis that treating object relative clauses with pre- or postverbal subject as just two more or less complex or marked variants may be the wrong approach. What if these two variants are not basically different in acceptability or processing complexity but just favored by different sets of properties? Increased processing complexity of ORs with postverbal subject would then be the consequence of using materials more adapted to ORs with preverbal subject. Testing this hypothesis requires an approach based on controlled empirical data. Therefore, we decided to run a written corpus study, followed by acceptability judgements and a self-paced reading experiment.

Contrary to previous corpus studies (Fuchs 1997), who claimed a slight advantage for preverbal subject overall (all relatives confounded), we found that subject inversion can be as frequent as preverbal subjects in French object relatives under fully controlled conditions. Thus, frequency per se does not predict a preference for one or the other as was found in Frauenfelder et al. (1980) or Pozniak & Hemforth (2015).

Corpus annotation on the French Treebank (Abeillé et al. 2019) also shows that object relatives with preverbal and postverbal subjects have different properties and are used in different contexts. Logistic regression models (Baayen et al. 2008) show that semantic factors (agentivity and intentionality) as well as length play a significant role in subject inversion.

In order to see whether these properties differentiate object relatives with preverbal and postverbal subject, we manipulated them in two experimental studies:

an acceptability judgement study and a self-paced reading experiment. The acceptability judgement experiment shows that subject inversion is rated highly acceptable and might even be preferred in object relative clauses under the right circumstances, contrary to previous experimental studies (Holmes & O'Regan 1981; Pozniak & Hemforth 2015). The self-paced reading experiment shows that verb agentivity and subject length both play a role in the use of object relatives with preverbal subject and with postverbal subject. A non-agentive verb and a long subject make OR+inv easier to process. However, we did not test other semantic factors such as object definiteness. More experiments examining semantic and discourse factors are needed to complete the picture. The results of our experiments were also not as clear cut as we might have wished. This may be due to the fact that they were run on an internet platform, where the experimental environment is much less controlled than in the lab.

Overall, our results cannot be explained by theories which would consider postverbal subjects generally more complex than preverbal subjects as proposed by some of the syntactic theories mentioned in the introduction. They cannot be explained either by processing theories such as DLT, which predicts a systematic advantage for subject inversion, or by syntactic theories like relativized minimality that may similarly predict an advantage for inversion. Depending on semantic properties, object relatives with a postverbal subject are not always easier or harder to understand than object relatives with a preverbal subject as suggested in the psycholinguistic literature, which is mainly focused on reversible relative clauses with animate subjects and objects, mostly using agentive verbs (Holmes & O'Regan 1981; Pozniak & Hemforth 2015).

To conclude, our three empirical studies emphasize the role of length and semantic/pragmatic factors (Mak et al. 2006; Traxler et al. 2002). The role of subject length could be explained by dependency locality theory (Gibson 2000) or by a more general tendency to put longer constituents at the end of the sentence (Behaghel 1909; Wasow 2002). The role of verb agentivity could be explained by semantic theories (Fuchs 2006; Marandin 2011).

Our studies also show that subject inversion is not marked and is no less frequent than preverbal subjects in French object relatives. Relatives with postverbal subject, which existed in Ancient French (Buridant 1999; Fuchs & Le Goffic 2006) coexist now with relatives with preverbal subject and can be felicitous depending on their semantic/pragmatic properties. Thus, we propose that subject inversion is not just a stylistic variant. Object relatives with preverbal or postverbal subject can be seen as two variants of a grammatical construction with different usage profiles and each can be more appropriate than the other in the right context.

Abbreviations

OR	Object relative
OR+inv	Object relative with subject inversion
OR–inv	Object relative without subject inversion
DLT	Dependency locality theory

Acknowledgements

We thank Clément Plancq for extracting the relative clauses from the French Treebank as well as Jean-Marie Marandin for very helpful discussions. We also thank the Procope Paris–Frankfurt project *One2Many* for its help. This work was partially supported by a public grant overseen by the French National Research Agency (ANR) as part of the program “Investissements d’Avenir” (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris – ANR-18-IDEX-0001.

References

- Abeillé, Anne, Lionel Clément & Loïc Liégeois. 2019. Un corpus annoté pour le français: le French Treebank. *TAL* 60(2). 19–43.
- Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4). 390–412.
- Barr, Dale J, Roger Levy, Christoph Scheepers & Harry J Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278.
- Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & R. Harald Baayen. 2015. Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.
- Baudiffier, Vanessa, David Caplan, Daniel Gaonac’h & David Chesnet. 2011. The effect of noun animacy on the processing of unambiguous sentences: Evidence from French relative clauses. *Quarterly Journal of Experimental Psychology* 64(10). 1896–1905.
- Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.
- Bonami, Olivier & Danièle Godard. 2001. Inversion du sujet, constituance et ordre des mots. In Jean-Marie Marandin (ed.), *Cahier Jean-Claude Milner*, 117–174. Paris: Verdier.

- Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language & Linguistic Theory* 19(1). 1–65.
- Buridant, Claude. 1999. L'ordre verbe-sujet en ancien français et son évolution vers le français moderne. Esquisse de comparaison avec les langues romanes. *Linx. Revue des linguistes de l'université Paris X Nanterre* (11). 167–188.
- Déprez, Viviane. 1990. Two ways of moving the verb in French. *MIT working papers in linguistics* 13. 47–85.
- Di Cristo, Albert. 2016. *Les musiques du français parlé: Essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*. Vol. 1. Berlin: Walter de Gruyter.
- Drummond, Alex. 2013. *Ibex farm (free hosting for Internet-based experiments)*. Version 0.3. <https://spellout.net/ibexfarm/> (23 February, 2021).
- Frauenfelder, Ulrich, Juan Segui & Jacques Mehler. 1980. Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior* 19(3). 328–337.
- Friedmann, Naama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1). 67–88.
- Fuchs, Catherine. 1997. La place du sujet nominal dans les relatives. In Catherine Fuchs (ed.), *La place du sujet en français contemporain*, 135–178. Louvain: Duculot.
- Fuchs, Catherine. 2006. La place du sujet nominal en français: de la syntaxe à l'énonciation. In F. Hrubaru & A. Velicu (eds.), *Enonciation et syntaxe*, 9–25. Editura Echinox.
- Fuchs, Catherine & Pierre Le Goffic. 2006. Le français moderne, entre V2 et SVO? In Olivier Bertrand, Sophie Prévost, Michel Charolles & Jacques François (eds.), *Discours, diachronie, stylistique du français*, 17–35. Peter Lang.
- Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec P. Marantz, Yasushi Miyashita & Wayne O'Neil (eds.), *Image, language, brain: Papers from the first mind articulation project symposium*, 95–126. Cambridge: MIT Press.
- Holmes, Virginia M. & J. Kevin O'Regan. 1981. Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior* 20(4). 417–430.
- Hulk, Aafke C.J. & Jean-Yves Pollock. 2001. *Subject inversion in Romance and the theory of Universal Grammar*. Oxford: Oxford University Press.
- Jaeger, T Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4). 434–446.

6 Subject inversion in French object relatives: What's your preference?

- Kampers-Manhe, Brigitte, Jean-Marie Marandin, Frank Drijkoningen, Jenny Doetjes & Aafke C.J. Hulk. 2004. Subject NP inversion. In Francis Corblin & Henriëtte de Swart (eds.), *Handbook of French semantics*, 553–579. Stanford: CSLI Publications.
- Kayne, Richard & Jean-Yves Pollock. 1978. Stylistic inversion, successive cyclicity, and Move NP in French. *Linguistic Inquiry*. 595–621.
- Lahousse, Karen. 2011. *Quand passent les cigognes. Le sujet nominal postverbal en français moderne*. Presses Universitaires de Vincennes; Paris.
- Le Bidois, Robert. 1952. *L'inversion du sujet dans la prose contemporaine (1900-1950): étudiée plus spécialement dans l'oeuvre de Marcel Proust*. Artrey.
- Mak, Willem M, Wietske Vonk & Herbert Schriefers. 2006. Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language* 54(4). 466–490.
- Marandin, Jean-Marie. 2011. Subject inversion in French. The limits of information structure. In *Proceedings of the 18th hpsg conference*, 327–347. CSLI Publications.
- Pozniak, Céline & Barbara Hemforth. 2015. Processing subject and object relative clauses in French and Mandarin Chinese. In Poster presentation, CUNY, Los Angeles.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge: MIT Press.
- Sag, Ivan A. 2010. English filler-gap constructions. *Language*. 486–545.
- Traxler, Matthew J, Robin K. Morris & Rachel E. Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47(1). 69–90.
- Wasow, Thomas. 2002. *Postverbal behavior*. Stanford: CSLI Publications.