Article:
Pozniak, Céline, Hemforth, Barbara, Haendler, Yair et al. (2019) Seeing events vs. entities: The processing advantage of Pseudo Relatives over Relative Clauses. Journal of Memory and Language. pp. 128-151. ISSN: 0749-596X

https://doi.org/10.1016/j.jml.2019.04.001

Seeing events vs. entities: The processing advantage of Pseudo Relatives over Relative Clauses

Céline Pozniak[a], Barbara Hemforth[a], Yair Haendler[a], Andrea Santi[b], Nino Grillo[c,*]

[a] *CNRS-Université Paris 7 Paris Diderot*
[b] *Department of Linguistics, University College London*
[c] *Department of Language and Linguistic Science, University of York*

Abstract

We present the results of three offline questionnaires (one attachment preference study and two acceptability judgments) and two eye-tracking studies in French and English, investigating the resolution of the ambiguity between pseudo relative and relative clause interpretations. This structural and interpretive ambiguity has recently been shown to play a central role in the explanation of apparent cross-linguistic asymmetries in relative clause attachment (Grillo & Costa, 2014; Grillo et al., 2015). This literature has argued that pseudo relatives are preferred to relative clauses because of their structural and interpretive simplicity. This paper adds to this growing body of literature in two ways. First we show that, in contrast to previous findings, French speakers prefer to attach relative clauses to the most local antecedent once pseudo relative availability is controlled for. We then provide direct support for the pseudo relative preference: grammatically forced disambiguation to a relative clause interpretation leads to degraded acceptability and greater processing cost in a pseudo relative environment than maintaining compatibility with a pseudo relative.

Keywords: Universality of Parsing Principles, Ambiguity Resolution, Economy of Computation, Locality, Attachment Preferences, (Pseudo)Relative Clauses

*Corresponding author. Department of Language and Linguistic Science, University of York, Heslington, York, YO10 5DD, UK. Tel:+44 (0)1904 323342

Email addresses: celine.pozniak@gmail.com (Céline Pozniak), bhemforth@gmail.com (Barbara Hemforth), yairhen@gmail.com (Yair Haendler), a.santi@ucl.ac.uk (Andrea Santi), nino.grillo@york.ac.uk (Nino Grillo)

1. Introduction

One strong hypothesis in psycholinguistics is that language processing is governed by mechanisms grounded in universal principles of optimal computation, such as those evidenced in primacy and recency effects. Principles of this sort, which dominate models of language processing and are often observed across cognitive domains, can hardly be construed as acquired, which explains their universal nature.[1]

From this perspective, crosslinguistic variation in parsing preferences is only apparent and ultimately reducible to grammatical variation, i.e. to the interaction of the grammatical properties of a given language with these basic principles of economy of computation.

There is, however, one domain of research in sentence processing where universality was famously called into question: relative clause (RC) attachment (Cuetos & Mitchell, 1988, and much related literature). Cuetos and Mitchell (1988) first observed that speakers of Spanish and English displayed a strikingly different parsing preference in the resolution of syntactic ambiguities involving two potential attachment sites of a RC: while English speakers relied on principles of minimal effort, attaching the RC to the closest potential host (the most local NP, *the actress* in (1-a)); Spanish speakers appeared to violate this principle, showing an overall preference for attachment to the non-local host (*the maid* in (1-b)). The locality principle governing attachment seemed therefore to apply differently across languages.

(1) a. Someone shot the maid$_1$ of the actress$_2$ that$_2$ was$_2$ standing on the balcony

b. Alguien disparó contra la criada$_1$ de la actriz$_2$ que$_1$ estaba$_1$ en el balcón

This asymmetry was particularly striking because of its exceptionality and specificity. Spanish and English speakers, in fact, show the same preferences when disambiguating sentences which involve principles governing structure building and filler-gap dependencies. They also show the same tendency to prefer local attachment when constituents other than RCs are tested (e.g. when attaching temporal modifiers in: *John said that Mary left yesterday*). These findings generated a

[1] A related, though independent, argument from learnability is discussed in detail in Fodor (1998a,b) in support of the universalist perspective: children need to parse the language they hear in order to acquire the grammar of their native language. This will be very hard, if not impossible, if principles of parsing have to be acquired themselves. And principles of parsing can hardly be acquired as long as there is no grammar to base this process on.

vast amount of literature aimed at explaining away this asymmetry. [2]

One such theoretical framework is the *Tuning Hypothesis*. Mitchell & Cuetos (1991) proposed that preferences for high or low attachment of RCs across languages stem from exposure to different statistical regularities that could vary depending on individuals or languages (Mitchell et al., 1995). Empirically, the literature provides conflicting results on this. See Cuetos et al. (1996) for supporting data from corpus analyses of English and Spanish and Mitchell & Brysbaert (1998) for problematic results from Dutch (see also Desmet et al. 2002a; Desmet & Gibson 2003; Gibson & Schütze 1999 for further discussion). One issue with corpus studies on RC attachment, which might in part explain the conflicting results, is that they did not take into account a number of fine-grained properties of both the RC and the complex DP that have been independently shown to strongly affect attachment. These include the type of relation between the two Nouns (Gilboy et al., 1995; Frazier & Clifton, 1996), the type of preposition connecting the two DPs (De Vincenzi & Job, 1993, 1995), the number of NPs (Gibson et al., 1996) and length of the DPs and the RC (on which see Hemforth et al., 2015, and references cited therein).[3] From a theoretical perspective, an important question of *directionality of the effect* is raised by corpus studies, that is, whether a form is less frequent because it is inherently more complex or less favoured by the parser, limiting its explanatory power.

Even though the role of RC-attachment frequency distributions in language specific preferences should not be disregarded, it now appears clear that it can hardly be the only factor explaining all the variation in attachment, particularly across languages. The in-depth investigation of RC attachment across structures and languages thus uncovered a variety of other factors that contribute to the disambiguation of RC attachment and the processing of adjuncts more generally. It is now apparent that semantic, pragmatic and prosodic factors all contribute to the disambiguation of sentences involving multiple potential hosts for an RC, and that these factors apply in substantially the same way across languages (for recent reviews see Grillo & Costa 2014; Hemforth et al. 2015).

An important recent development in this debate came with the discovery that the previous literature on RC-attachment contained a grammatical confound in the cross-linguistic compar-

[2] Borrowing the words of an anonymous reviewer: "A great deal of psycholinguistic ink has been spilled trying to identify the basis for th(is) purported difference".

[3] Previous corpus studies looking at RC-attachment also did not consider the confounding role of the selective availability of Pseudo Relatives, on which see below, which makes an evaluation of their results highly problematic.

isons (Grillo, 2012). A subset of the languages under study, including Spanish but not English, allow for constructions known as *Pseudo Relatives* (PR). Faithful to their name, these imposters are string identical to RCs (2). The two constructions, however, display very different structural, interpretive and prosodic properties. Crucially, there is no attachment ambiguity under the PR parse (2-b, c), as the first NP of a complex NP is the only accessible subject for the embedded predicate. In other words, with PRs High Attachment is obligatory. [4] PRs are found not only in Spanish but also in a number of so-called High Attachment languages (including French, Dutch, Greek and Serbo-Croatian, among others). PRs, however are not available in Low Attachment languages including English, Basque, Romanian and Chinese.[5] Not recognizing this grammatical distinction necessarily puts the explanatory burden for variation in attachment preferences on the parser, causing the aforementioned crisis.

(2) a. Jean a vu [$_{DP}$ l' [$_{NP}$ homme [$_{CP}$ qui courait.]]]
J. has seen the man that run.
'John saw the man that ran.'

b. Jean a vu [$_{PR}$ [$_{DP}$ l'homme] [$_{CP}$ qui courait.]]
J. has seen the man that ran.
'J. saw the man running.'

c. John saw [$_{PR}$ the man running].

Discovery of this confound led to formulation of the *PR-first Hypothesis*, which suggests that PRs are both interpretively and structurally simpler than RCs and thus should be preferred by the parser (see below for details). Recent results on RC attachment indirectly support this hypothesis by showing a strong effect of PR availability on RC attachment: when the PR confound is eliminated and only unambiguous RCs are presented, there is a strong tendency to attach locally across languages and structures. Non-local / High Attachment is observed across languages when a PR reading is available.

[4] For ease of exposure, throughout this paper, we will call the obligatory interpretation of the NP1 as the subject of the PR in "NP1 of NP2" constructions in PR environments "High Attachment" although this strictly only applies to relative clauses.

[5] For reference on RC attachment in these languages see e.g.: Mitchell et al. (1990); Frenck-Mestre & Pynte (2000); Zagar et al. (1997); Colonna et al. (2000); Colonna & Pynte (2001a) (French), Brysbaert & Mitchell (1996); Mitchell & Brysbaert (1998); Mitchell et al. (2000); Desmet et al. (2002b) (Dutch), Papadopoulou & Clahsen (2003) (Greek), Lovrić (2003) (Serbo-Croatian, Gutierrez-Ziardegi et al. (2004) (Basque), Ehrlich et al. (1999) (Romanian), Shen (2006) (Chinese).

In the present paper we extend these findings in multiple directions. After a brief introduction on the contrast between PRs and RCs and a short summary of previous experiments on the effects of PR availability on the resolution of RC attachment ambiguities (remainder of Section 1) we show in an offline completion study that native speakers of French display a clear preference for Low Attachment when unambiguous RCs are tested and other relevant factors (such as prosody or referentiality) are controlled (Section 2). As predicted, High Attachment is observed with the same complex NP + RC combinations in environments that license PRs. We then directly test the PR-first Hypothesis in two sets of experiments that acquire acceptability judgments (Section 3) and eye-tracking while reading data (Section 4). Each of these experiments was carried out in French, a PR-language, and in English, a non-PR language. Overall, the results further support the claim that parsing principles are universal: previously reported cross-linguistic differences in parsing preferences are strongly grounded in independent grammatical distinctions and are thus epiphenomenal.

1.1. Pseudo Relatives

Relative clauses in the complement position of perceptual verbs in languages like French (3), but not English, are ambiguous between an RC reading (3-a), and a pseudo relative reading (3-b).[6] Despite being string identical, PRs and RCs are structurally and interpretively different.

As shown in (4) and (5), in RCs the CP is embedded within the DP it modifies, but in PRs it stands in a sisterhood relation with the same DP. In (4), the RC *that ran* is embedded in the DP, contrary to the PR *that ran* in (5) which is the sister of the DP and selected by the verb.

(3) a. Jean a vu [$_{DP}$ l' [$_{NP}$ homme [$_{CP}$ qui courait.]]]
J. has seen the man that run.
'John saw the man that ran.'

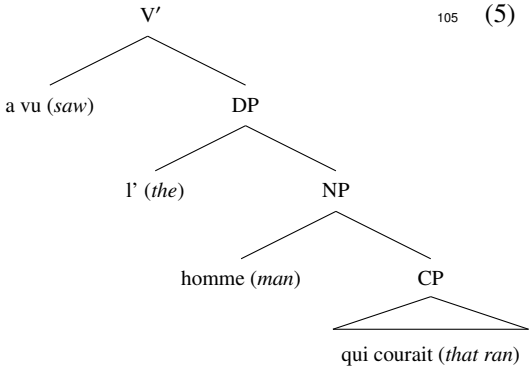
b. Jean a vu [$_{PR}$ [$_{DP}$ l'homme] [$_{CP}$ qui courait.]]
J. has seen the man that ran.
'J. saw the man running.'

c. I saw [$_{PR}$ the man running].

[6] On PRs see e.g. Radford (1975); Kayne (1975); Graffi (1980); Burzio (1986); Cinque (1992); Rizzi (1992); Guasti (1988, 1992); Côté (1999); Rafel (1999); Casalicchio (2013); Moulton & Grillo (2015); Grillo & Moulton (2016) among others. For analysis of PRs as predicative relatives see Koenig & Lambrecht (1999) and Lambrecht (2000).

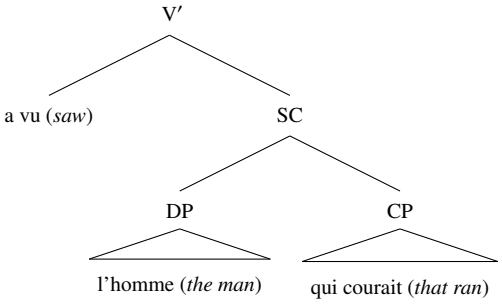
RELATIVE CLAUSE

(4)



PSEUDO RELATIVE

(5)



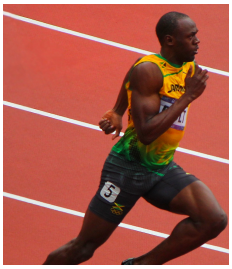
This structural difference is accompanied by a sharp difference in interpretation. RCs, being modifiers of the DP they attach to, denote properties of the entities introduced by those DPs (6). When a perceptual verb takes a DP modified by an RC, it provides a perceptual report of the unique individual introduced by the DP, which possesses the property introduced by the RC (i.e. the unique runner in the example in (6)). PRs, which roughly correspond to so-called eventive small clauses in English (3)[c], are composed by a subject DP and a CP predicate. Like eventive Small Clauses in English, PRs denote events or situations (7). Embedding of a PR under a perceptual verb gives a perceptual report of an event (a *running* event in the present example).[7]

[7]For clarity of presentation, we show simplified semantics for PRs. For a more detailed discussion on the syntax-semantics of PRs see Moulton & Grillo (2015); Grillo & Moulton (2016) and references cited therein. For discussion of how these structural differences are encoded at the prosodic level, see Grillo & Turco (2016).

(6) RC: *John saw the man that runs*

∃e [see(e) & EXPERIENCER(e)(John) & STIMULUS(the unique man that ran)(e)]
There is an event of *seeing* and the experiencer of that event is *John* and the stimulus of the event is *the unique man that ran*.[8]

(7) PR: *John saw the man running*



∃e∃e'[see(e) & EXPERIENCER(e)(John) & STIMULUS(e')(e) & run(e') & AGENT(e')(the man)]
There is an event of *seeing* and the experiencer of that event is *John* and the stimulus of the event is *an event of running* and the agent of running is *the man*.[9]

The structural and interpretive differences between PRs and RCs are responsible for a number of asymmetries in their distribution, which helpfully can be manipulated experimentally to independently investigate their processing. The remainder of this section briefly introduces two asymmetries used in the experiments that follow.

Restriction on Matrix Verb. An important characteristic of PRs is that, just like English eventive small clauses and contrary to run-of-the-mill RCs, they are only available in selected environments. PRs and small clauses are allowed under perceptual verbs but are clearly not available with stative predicates:[10]

(8) a. Marie a vu Bolt qui courait.
M. has seen B. that run.IMPF.
'Marie saw Bolt running.'

[8] *My Hero*, Zwerink (2011), Public Domain.

[9] *Run*, Mackintosh (2012), Public Domain.

[10] PRs and SCs are also licensed under other types of predicates, e.g. *meet, catch, film* a.o. Here and elsewhere we use proper names to disambiguate for the PR reading. Proper names can also head appositive RCs, however it is easy to show that these are also distinct from PRs (e.g., see Radford 1975 and much related work. For example, PRs do not involve the typical *comma intonation* of appositive RCs). The contrast with pronominals in (9) further clarifies that we are not dealing with appositive RCs, which are not licensed with pronouns.

b. *Marie a épousé Bolt qui courait.
M. has married B. that run.IMPF
'*Marie married Bolt that was running.'

The difference in meaning between a PR and RC is responsible for this asymmetry. While perceptual verbs can introduce both events or entities (e.g. Mary can *see (the person) Bolt* or *see (the event of) Bolt running*), other verbs exclusively introduce entities (e.g. Marie can *marry (the person) Bolt*, but not *marry (the event of) Bolt running*).

This contrast is even more striking when pronominal objects are used. A perfectly acceptable result arises under perceptual verbs (9-a) (that is, when a PR reading is licensed), and complete ungrammaticality under stative verbs (when only the RC reading would be available). This asymmetry is due to the fact that RCs, whether restrictive or appositive, can never modify pronominals. Since PRs are not available under stative predicates, the PR-analysis which rescues (9-a), cannot be used for (9-b).

(9) a. Marie l'a vu qui courait.
M. him'has seen that run.IMPF.
'Mary saw him running.'

b. *Marie l'a épousé qui courait.
M. him'has Married that run.IMPF
'*Mary married him that was running.'

Restrictions on Tense. Another characteristic distinguishing PRs from RCs is that the tense of the embedded clause is anaphoric in PRs. This means that the perceptual event introduced in the matrix clause and the perceived event introduced by the PR must happen simultaneously. Simplistically speaking, tense specification of the embedded clause has to match the tense specification of the matrix clause in PRs (10-a). A past under present leads to an ungrammatical structure (10-b). This requirement obviously does not apply to RCs (10-c). This property of PRs is not surprising. As mentioned above, PRs under perceptual verbs involve direct perception of an event. If the event of perception happened in the past, then what was perceived (ie, another event) must have also taken place in the past. Similarly, if I am watching John running now, he must be running now. [11]

[11]There are apparent restrictions to this rule, involving present under future and present under present perfect, which are discussed in Grillo & Moulton (2016). These are irrelevant for the present experiment which used past under present,

(10) a. Jean l'a vu qui courait. PR-only
J. him'has seen that run.IMPF
'John saw him running.'

b. *Jean le voit qui courait.
J. him sees.PRES that ran.IMPF
'John sees him that was running.'

c. Jean voit l'homme qui courait. RC-only
J. sees.PRES the'man that ran.PAST
'John sees the man that was running.'

1.2. Previous studies and the PR-first Hypothesis

Grillo (2012) proposed that RC attachment preference should co-vary with PR availability and thereby explain (in part) the cross-linguistic asymmetry in attachment preferences. Everything else being equal, in Non-PR languages like English (as well as Basque, Romanian and Chinese) speakers display a preference for local attachment of the RC (attachment to *the man* in (11-a)), while in PR languages like French (as well as Italian, Spanish and Greek) speakers prefer non-local, or High Attachment (attachment to *the son* in (11-b)).

(11) a. I saw the son of the man that was running.

b. J'ai vu le fils de l'homme qui courait.
I.have seen the son of the man that ran.
'I saw the son of the man running.'

Under the PR reading, the attachment ambiguity disappears because of standard structural restrictions (c-command), the only accessible subject for the embedded predicate is the non-local NP (*the son*).[12]

(12) J'ai vu $[_{PR}\ [_{DP}$ le fils$_i$ $[_{PP}$ de $[_{DP}$ l'homme$_j$]]] $[_{CP}$ qui$_{i,*j}$ courait.]]
I.have seen the son of the man that ran.
'I saw the son of the man running.'

To account for this pattern, Grillo (2012) and Grillo & Costa (2014) proposed the *PR-first Hypothesis*: PRs are easier to parse than RCs for structural, semantic and pragmatic reasons. For

which in no way can be construed as a PR.

[12] Structurally, this is the same scenario we find when the complex DP in a subject position, as in e.g.: *[[the son [of [the man]]] ran]*. In this configuration, the second NP (*man*) is too deeply embedded within the subject DP to be an accessible subject.

instance, PRs have impoverished syntax and semantics with respect to RCs. As discussed above, Tense is anaphoric/dependent in PRs but referential in RCs. Moreover, PRs stand in a sisterhood relation with the head NP, while RCs are embedded within the same NP, making the RC an arguably more complex configuration. Another reason for the smaller difficulty with PRs is that they convey information relevant for the main assertion of the clause (Frazier, 1990); in fact they can be projected as *arguments* of the main clause (I saw an event). By contrast, RCs are always *adjuncts* (I saw an entity, which has a certain property, introduced by the RC itself). Finally, PRs involve less presuppositions than RCs. PRs do not require selection from a pre-established set of entities in the discourse (Crain & Steedman, 1985; Altmann & Steedman, 1988). [13]

Further to the correlation observed in the previous literature, a number of novel experiments directly manipulated PR availability across languages and demonstrated a strong effect of PR availability on attachment preference.

Grillo & Costa (2014) report a significant effect of verb type on RC attachment in Italian, with High Attachment observed under perceptual verbs such as *see* (78.6% High Attachment preference) and Low Attachment under stative verbs like *live with* (24.2% High Attachment preference). Comparable results were obtained from other PR-languages: Greek (Grillo & Spathas, 2014), Portuguese (Grillo et al. 2012a,b, 2013; Fernandes 2012; Tomaz et al. 2014) and Spanish (Grillo et al., 2012b; Aguilar & Grillo, 2016). These are all languages that were previously classified as High Attachment languages. However, these studies showed that this classification is epiphenomenal: as predicted by the *PR-first Hypothesis*, Low Attachment preference was observed consistently in each of these languages in unambiguous RC environments, whereas High Attachment preference was only observed in PR-compatible environments.

Importantly, Grillo et al. (2015) show that predicate semantics/plausibility alone does not

[13] It is important to clarify that the name *PR-first* does not imply that this hypothesis is tied to a serial model of language processing. The hypothesis is perfectly compatible with (ranked) parallel models of sentence processing and *first* can be interpreted as most highly ranked. *PR-first* should also not be taken to be an independent parsing principle, it simply captures the application of independently motivated parsing principles (e.g. Relativized Relevance (Frazier, 1990) or Principle of Parsimony (Crain & Steedman, 1985)) to the PR-RC ambiguity. Preference for PR over RCs, finally, is certainly not the sole factor determining RC attachment; previous work has shown that this is strongly modulated by a number of other factors both within and across languages, including pragmatics (Gilboy et al., 1995; Frazier & Clifton, 1996), prosody (Fodor, 2002; Hemforth et al., 2015) as well as independent grammatical properties of the languages under scrutiny (e.g., RCs introduced by a complementizer vs. obligatory relative pronouns, like in German, Russian and Bulgarian; see Hemforth et al. 2000 and Grillo & Costa 2014).

account for these results, as the same verb type manipulation (perceptual vs stative) in English (a non-PR language) did not effect preferences (ie, lead to overall High Attachment).

In sum, RC attachment preferences in complex NPs are modulated by the availability of an eventive small clause (be it a PR, a prepositional infinitive construction or the garden variety accusative+progressive small clause in English), adding evidence to the idea that cross-linguistic differences in RC attachment are rooted in grammatical differences. These results are also consistent with the idea that there exists a preference for the PR reading over the RC reading, as claimed by the *PR-first Hypothesis*.

A straightforward prediction of the *PR-first Hypothesis* that has not been tested yet is that tense mismatch in the environment of PR-compatible verbs should lead to reanalysis of the initial PR preference, with observable processing costs. By contrast, tense (mis)match should play no role in the interpretation of the embedded clauses in globally unambiguous RCs, for instance, RCs in the environment of stative verbs. We therefore predict a qualitatively different effect of tense manipulation in globally unambiguous RC environments.

Before we describe the experiments that directly tested this *PR-first* prediction (Section 3 and Section 4), we briefly show that PR availability modulates RC attachment preferences in French (Section 2). This would serve as a pre-test of the effects of PR availability and add to the literature on RC attachment in French.

2. Experiment 1: RC-attachment in French

In the first experiment we wanted to make sure that verb type manipulation (that is, PR availability) modulates RC attachment in French in the same way as it does in other PR-compatible languages. As discussed for Italian, under perceptual verbs (cf. example (13-a)) the embedded clause might be attached as sister of the non-local NP (*the son*), leading to a PR-reading with no attachment ambiguity. The alternative RC-reading can either be attached to the local NP (*the policeman*) or the non-local one (*the son*). Under stative verbs, however, only the latter ambiguity is present. The embedded clause can only be construed as a RC, attached either high or low. We tested the availability of these PR- and RC-readings in an offline completion study in French.

Participants

Sixty-nine native speakers of French (mean age 41.4 years) volunteered to participate in the experiment. They were recruited on the RISC platform (https://www.risc.cnrs.fr). In this as well as all following experiments, participants provided informed consent before starting the first trials.

Material

The critical trials consisted of 24 ambiguous target sentences containing complex NPs of the form NP1 of NP2 followed by a finite CP. These complex NPs were placed in object position of either perceptual verbs (13-a) or stative verbs (13-b). Sixty unrelated fillers were added. All items were presented to participants in a standard Latin square design: verb types were counterbalanced across the items. The order of items was randomized individually for each participant.

(13) a. Marie écoute le fils du policier qui parle. PERCEPTUAL
Marie hears the son of the policeman that is speaking.

b. Marie est employée par le fils du policier qui parle. STATIVE
Mary is employed by the son of the policeman that is speaking.

Procedure

Participants read the sentences on a computer screen through the IBEX platform (Drummond 2013). After reading each sentence, participants were asked to complete a sentence describing the event in the embedded clause by filling in the blank space in the subject position with either the local or non-local NP. This is illustrated in (14).

(14) Léa est fiancée au voisin du coordonnier qui danse.
Le ..... danse.
'Lea is engaged to the neighbour of the shoemaker that is dancing.'
The ..... is dancing.'

Analysis

The attachment preference data were analyzed with a Bayesian linear mixed-effects model that assumes a Bernoulli distribution of the dependent variable, with a logit link function. The

model was fitted using the *brms* package (Bürkner, 2018) in R (R Core Team, 2018). The binary dependent variable of attachment preference was coded as 1 (High Attachment) or 0 (Low Attachment). The levels of the factor Verb Type were coded as 1 (perceptual) and -1 (stative). The Bayesian model generates a posterior distribution for each of the model parameters. The parameters of interest are typically the slopes of the group-level (or fixed) effects. In this model, it is the slope parameter for the fixed effect of Verb Type.

Due to the 1/-1 coding of the factor, zero represents the point of "no difference" between the two verb types. Therefore, we will calculate the probability that the parameter of interest is greater or smaller than zero. A high probability that the Verb Type parameter is greater than zero would suggest there is evidence that High Attachment is preferred with perceptual verbs (coded as 1). By contrast, high probability that the parameter is smaller than zero would mean that High Attachment is preferred with stative verbs (coded as -1).

For the interpretation of the data, we will look at the estimated mean ($\hat{\beta}$) and range of the posterior distribution for the parameter of interest. The distribution's 85% credible intervals will mark the range of which which we can be certain with probability 0.85 that it includes the true value of the parameter of interest, given the data and the model at hand. Similarly, the 95% credible intervals will mark the range of which we can be certain with a probability of 0.95 that it includes the true value of the parameter (Hoekstra et al., 2014; Morey et al., 2016). It should be emphasized that, in this kind of analysis, we refrain from calling an effect "significant" or "not significant". These terms, used in Frequentist analyses, reflect the binary nature of statistical inference in such models (an effect can be either significant or not). By contrast, inference in the Bayesian analysis used here is not categorical. Evidence for a particular effect may be more or less strong, as expressed by whether zero is excluded from the posterior distribution altogether, whether it lies outside the 95% credible intervals, outside the 85% credible intervals, and so on. In the appendix we provide a more detailed motivation for preferring the Bayesian analysis over a Frequentist one.

In the model for this experiment, the fixed effects part included an intercept and the main effect of Verb Type. The random effects part included adjustments for subjects and items of an intercept, the slope for this main effect and the correlations between intercepts and slopes (Baayen et al., 2008; Barr et al., 2013). We placed weakly informative priors on the model parameters (Nicenboim & Vasishth, 2016). The model ran with 4 MCMC chains and 3000

iterations each, of which the first 1500 iterations, the so-called warm-up phase, were discarded. Model convergence was verified by checking visually that the chains converged, by making sure that the $\hat{R}$ statistics for each parameter coefficient was equal to 1 and by checking that the number of effective samples per iteration was reasonably high for the parameters of interest (Gelman et al., 2013).

Results

As can be seen in Figure 1 The results show an overall High Attachment preference (61%) in PR-compatible environments (under perceptual verbs) and a strong preference for Low Attachment (72%) with unambiguous RCs (under stative verbs).[14] The statistical model confirms that the difference between High Attachment preference in the two verb types is reliable ($\hat{\beta}$ = 1.11, 95% CrI = [0.75, 1.51], P($\hat{\beta}$) > 0 = 1). The posterior distribution of the main effect of Verb Type has a range of positive numbers only (cf. Figure 2). We thus can be certain with probability 1 that the true parameter value for this effect is greater than zero. In other words, given the data, we find strong evidence for a High Attachment preference in sentences with perceptual verbs.

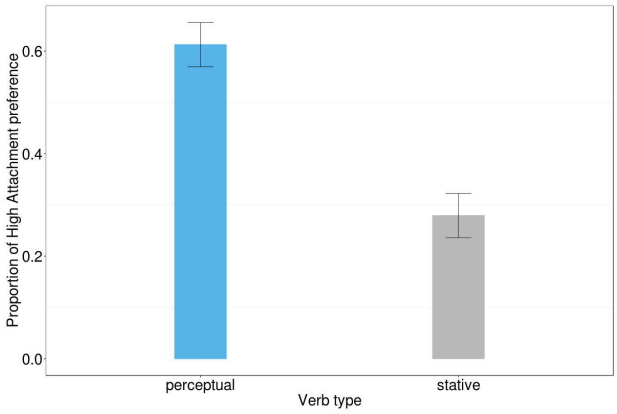


Figure 1: Mean High Attachment preference under the two verb types (with 95% confidence intervals).

[14] On a few trials, participants left the space blank or filled in an irrelevant word (37 out of 1656 trials, 2.2% of the data). These trials were excluded from the analysis.

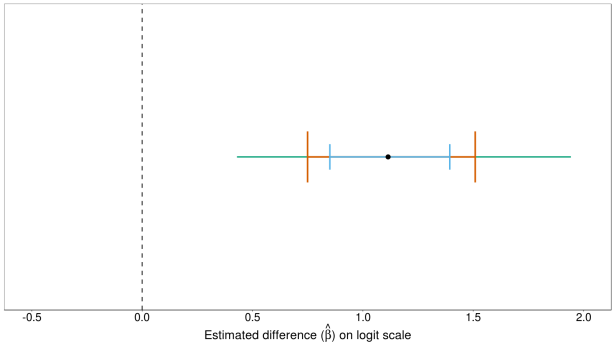


Figure 2: Posterior distribution for the effect of High Attachment preference in perceptual vs. stative verbs. The posterior's mean is represented by the black dot. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

311

312 *Intermediate discussion*

313 As in previously reported results for other PR-compatible languages, we find that PR avail-
314 ability strongly modulates RC attachment also in French. A strong preference for local attach-
315 ment is observed with unambiguous RCs (under stative predicates), supporting the idea that
316 locality principles play a central role across languages. Conversely, a High Attachment prefer-
317 ence emerges when PRs are available (with perceptual verbs). Since only the non-local NP is an
318 accessible subject in the PR reading, we can explain the strong preference for High Attachment
319 in PR-compatible environments as a preference for the PR reading over the RC reading. Exper-
320 iment 1 thus provides further support for the idea that PR-compatible structures are universally
321 preferred by the parser over RCs.

322 The set of results discussed so far provides merely indirect support for a parsing preference
323 of PRs over RCs. Our goal here is to provide a test capable to directly falsify the *PR-first Hy-*
324 *pothesis*. If PRs are indeed preferred to RCs, a RC disambiguation of otherwise PR-compatible
325 structures should come with an observable cost.

326 Several factors can be manipulated to force a RC reading in otherwise PR-compatible envi-
327 ronments. One case in point is tense. As mentioned, tense is anaphoric in PRs, but not in RCs.
328 In other words, the tense specification of the embedded clause must match the tense specification
329 in the matrix clause in PRs but not in RCs.

This allows us to construct minimal pairs which are locally ambiguous between a PR and RC reading up to the point of the tense specification of the embedded predicate which disambiguates the structure: tense match (15-a) is compatible with a PR reading; tense mismatch (15-b) will force a RC reading.

(15) a. Jean **a vu** la fille qui **poussait** la femme. PR/RC
J. **has seen.PAST** the girl that **pushed.PAST** the woman.
'J. saw the girl that pushed the woman / pushing the woman.'

b. Jean **voit** la fille qui **poussait** la femme. RC-only
J. **sees.PRES** the girl that **pushed.PAST** the woman.
'J. sees the girl that pushed the woman / *pushing the woman.'

Tense manipulation constitutes an ideal type of disambiguation in that it allows us to keep the entire structure of 'DP + embedded clause' identical across conditions, as in example (15). The only difference across conditions is the tense specification on the matrix clause: PAST in the PR-compatible condition vs. PRESENT in the globally unambiguous RC-condition.

In the following sections we present the results of four experiments that were designed to test the effects of tense (mis)match in French (a PR-language) and in English (a non-PR language), while manipulating the environment of PR-compatible sentences (perceptual verbs) and RC-only sentences (stative verbs). The first two experiments used acceptability judgments as a proxy for processing complexity (Section 3). The second set of studies used eye-tracking while reading with the same stimuli, in order to further investigate the time course of the potential processing difficulty (Section 4).

3. Experiment 2: Acceptability of tense (mis)match in French and English

The interaction of PR-availability (verb type) and tense (mis)match was tested in an acceptability rating experiment in French, a PR-language, and English, a non-PR language, as a control.

Participants

Fifty-eight native French speakers (mean age 29) and 103 native English speakers (mean age 31) participated in the experiments.

French participants were volunteers recruited on the RISC platform; English speakers were recruited via Amazon Mechanical Turk. English speakers received a monetary compensation for

their participation.

Material

Both experiments, in French and in English, realized a 2*x*2 design. As shown in Table 1, we manipulated two variables: VERB TYPE (perceptual / stative) and TENSE (match / mismatch). The rest of the sentence, i.e. the region of interest (NP + embedded clause), including the critical region (embedded verb), were kept identical across conditions. All relative clauses in the critical items were subject relative clauses, both in French and in English.[15] We created 24 items, with 6 items per condition, arranged in 4 lists in a standard Latin square design. Twenty-six fillers in French and twenty-nine fillers in English were added to each list and three practice trials preceded the experiment for each list.[16] The items and the fillers were fully randomized, so that each participant saw a different order of the sentences. The experimental items and the fillers in French and English were close translations.

[15]The French relative clauses used here cannot be confused with object relative clauses with a post-verbal subject. The complementizer *qui* unambiguously marks the sentence as a subject relative clause, object RCs are introduced by the alternative complementizer *que*.

[16]The fillers consisted of ambiguous sentences in English and in French with either a collective or a distributive reading. Sentences like "The children built a sandcastle. The sandcastle(s) were beautiful" tested whether participants thought all the children built only one castle together, or whether they thought each child built one castle of their own). All the fillers were grammatical sentences, and were rated in the range between 6.5 and 9.3 on the 1-10 scale (Dobrovie-Sorin et al. 2016).

Verb Type	Tense	Example item
Perceptual	Match	*Jean **a vu** la fille qui **poussait** la femme.* John **saw** the girl that **pushed** the lady.
	Mismatch	*Jean **voit** la fille qui **poussait** la femme.* John **sees** the girl that **pushed** the lady.
Stative	Match	*Jean **était marié à** la fille qui **poussait** la femme.* John **was married to** the girl that **pushed** the lady.
	Mismatch	*Jean **est marié à** la fille qui **poussait** la femme.* John **is married to** the girl that **pushed** the lady.

Table 1: Example of an item in the four conditions

Procedure

The procedure was the same in English and French. Participants had to judge the acceptability of each sentence on a scale from 1 (completely unacceptable) to 10 (completely acceptable). The sentences appeared one at a time on the computer screen with the acceptability scale below it. Both experiments were run on the Ibex Platform (Drummond 2013 Ibex Farm, http://spellout.net/ibexfarm/). Participants did the experiment at a place of their choice. They were however asked to do it in a quiet environment and to not take breaks.

Predictions

Based on the *PR-first Hypothesis* we should expect to see a three-way interaction of Tense, Verb Type and Language. More specifically, the *PR-first Hypothesis* predicts higher acceptability for French sentences with tense match (which are PR-compatible) than with tense mismatch under perceptual verbs. No such effect is predicted in sentences with stative verbs. Also, no interaction is expected in English, as a PR interpretation is excluded in this languages and all items describe unambiguous RCs.

Analysis

The acceptability judgment data were analyzed with a Bayesian cumulative mixed-effects model with a logit link function (Agresti 2012; Christensen & Brockhoff 2013) using the *brms* package

in R. The levels of the factor Tense were coded as 1 (match) and -1 (mismatch); the levels of the factor Verb Type were coded as 1 (perceptual) and as -1 (stative); the levels of the factor Language were coded as 1 (French) and -1 (English). The interpretation of the results will follow the same principles as in Experiment 1.

The fixed effects part included an intercept and the main effects of Tense, Verb Type and Language, as well as all their possible interactions. The random effects part included adjustments for subjects and items of an intercept, of slopes for the effects and interaction of Tense and Verb Type, and the correlations between intercepts and slopes. We ran the model with 4 chains and 6000 iterations each. The first 3000 warm-up iterations in each chain were discarded. Model convergence was verified by checking the chains' convergence, the $\hat{R}$ statistics for each parameter coefficient and the number of effective samples per iteration.

Results

Sentences with tense match were rated as more acceptable than sentences with a tense mismatch under perceptual verbs in French (Figure 3). The effect did not show up under stative verbs, for which there was a similar acceptability level in both tense mismatch and tense match trials. Moreover, the different effect of tense on the two verb types did not show up in English.

Further the analysis showed that the posterior distribution for the crucial three-way interaction of Tense by Verb Type by Language excludes zero and has a range of only positive values ($\hat{\beta} = 0.17$, 95% CrI = [0.10, 0.24], P($\hat{\beta}$) > 0 = 1), presenting strong evidence in support of an interaction (Figure 4). To break down this interaction and see in which language we find evidence for the interaction of Tense and Verb Type, we ran two additional models, one for each language separately. Each model estimated fixed effects intercept and slopes for the main effects of Tense, Verb Type and their interaction, and random intercepts for subjects and items. [17]

For French (left panel of Figure 5) we found evidence for the effect of Verb Type ($\hat{\beta} = 0.20$, 95% CrI = [0.09, 0.31], P($\hat{\beta}$) > 0 = 0.99), suggesting that sentences with perceptual verbs were rated as more acceptable than sentences containing stative verbs. There was similarly reliable evidence for the effect of Tense ($\hat{\beta} = 0.19$, 95% CrI = [0.09, 0.29], P($\hat{\beta}$) > 0 = 0.99), meaning that sentences

[17] Note that in a Bayesian analysis it is not necessary to apply a correction for multiple testing, such as a Bonferroni-correction for p-values in a Frequentist analysis. This is because there are no Type I or Type II errors in a Bayesian analysis, where inference does not depend on assumptions made concerning the replicability of the experiment and its results (Nicenboim et al. 2018).

with tense match were rated higher than sentences with tense mismatch. Most importantly, there was strong evidence for the interaction of Verb Type by Tense ($\hat{\beta}$ = 0.27, 95% CrI = [0.17, 0.38], P($\hat{\beta}$) > 0 = 1). This interaction means there was no difference in rating between tense match and mismatch under stative verbs, but under perceptual verbs tense mismatch sentences were rated lower than tense match sentences.

Crucially, these results were different in English (right panel of Figure 5). In this language, there was strong evidence for the effect of Verb Type ($\hat{\beta}$ = 0.11, 95% CrI = [0.04, 0.19], P($\hat{\beta}$) > 0 = 0.99). There was no evidence for the main effect of Tense ($\hat{\beta}$ = 0.04, 95% CrI = [-0.04, 0.11], P($\hat{\beta}$) > 0 = 0.84) and, as predicted, no evidence for the interaction of Verb Type and Tense ($\hat{\beta}$ = 0.03, 95% CrI = [-0.04, 0.11], P($\hat{\beta}$) > 0 = 0.78).

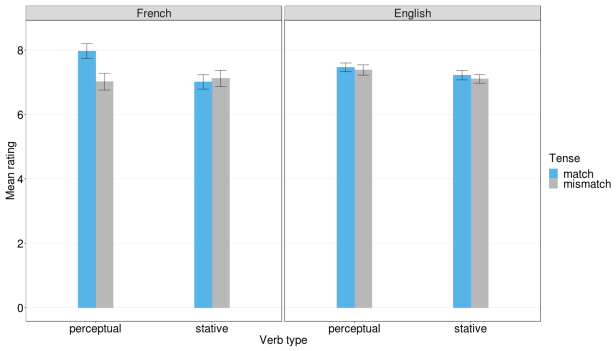


Figure 3: Mean acceptability rate (with 95% confidence intervals) as a function of tense and verb type, in French (left panel) and English (right panel).

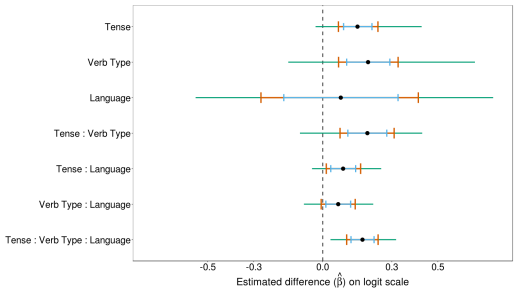


Figure 4: Acceptability rating – posterior distributions of the fixed-effects parameters in the model that includes both languages together. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

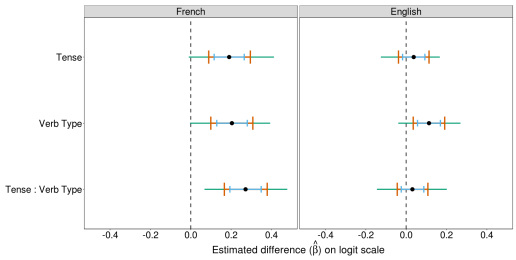


Figure 5: Acceptability rating – posterior distributions of the fixed-effects parameters in the models fitted separately for French (left panel) and English (right panel). The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

Intermediate Discussion

As predicted by the *PR-first Hypothesis*, tense mismatch negatively affected acceptability when it forced a RC reading in an otherwise PR-compatible environment, that is, under perceptual verbs and only in French. The tense manipulation, however, did not affect acceptability in globally unambiguous RC-only environments. Importantly, this interaction was only observed in French. No effects of tense and no interaction between TENSE and VERB TYPE were found in English, where only the RC-parse was available across the two types of verbs. The results thus fully support the *PR-first* predictions. The parser does appear to favour a PR over a RC interpretation, when the former is available. Along with previous results from RC attachment studies, these results point to a preference for secondary predication over restrictive interpretation or, to put it differently, a preference for *events* over *entities* in the complement of perceptual verbs.

An anonymous reviewer of a previous version of this paper correctly pointed out that the rating of the matching condition under perceptual verbs is higher than all the other conditions and raises the question of whether this might depend on the fact that in this condition the parser does not have to make a choice among alternative parses, while this choice is imposed in the tense mismatch condition, i.e. when an RC reading is imposed on the locally ambiguous structure. Notice, however, that an interpretation of the results in terms of competition does not explain the low ratings of the unambiguous RC conditions. Given that matrix stative predicates only license an RC parse of the embedded clause, the positive effect of lack of competition should be also

observed here, contrary to what we see. Low rates in all the RC conditions seem to align with an inherent higher complexity of the RC. The reviewer's observation, nevertheless, raises a very important point, as it is important to decide whether a single factor (i.e. complexity of RCs) underlies the low ratings of RCs across verb type, or whether the effect is qualitatively different in the two environments. To address this question we need to complement the current offline results with online data.

4. Experiment 3: Tracking eye-movements while reading tense (mis)match across languages

We now proceed to the following question: how does the apparent preference for PRs over RCs unfold online? To address this question, we ran two eye-tracking studies, one in French and one in English, using the same stimuli from the acceptability judgment experiments.

Participants

We had two separate groups of participants, a French-speaking one and an English-speaking one. In the first group, 62 French native speakers living in Paris participated in the French experiment (mean age 28). For the experiment in English, 50 English native speakers participated in the experiment, with 26 living in London, 20 in Glasgow and 4 temporarily in Paris (mean age 26). All participants gave their informed consent and received either monetary compensation or course credits to participate in the experiment, and all were naive as to the purpose of the study. All participants had normal or corrected-to-normal vision.

Materials

The linguistic material in each language was the same as in the corresponding acceptability judgment experiment. Differently from the rating experiments, though, comprehension questions were added to verify that participants were paying attention and concentrated on reading the sentences. The experiment in French and the one in English each included 16 simple questions for each list (around 35% of all the trials). For example, for a sentence like *John saw the girl that pushed the lady*, the question was *Did the girl push a lady?* These questions were identical across conditions.

Procedure

Eye fixations were recorded with Eyelink II in the French experiment. As for English, eye fixations were recorded with Eyelink 1000 for the experiments in London, and Eyelink II for the experiments in Glasgow and Paris. The system recorded each participant's dominant eye movements while they were reading sentences using the Miles test (Miles, 1930). Sentences appeared in 20-point font on the screen on a single line for the target items. Participants had to read the sentence at a natural pace and press the space bar on the keyboard when they were done. To answer the comprehension questions, they had to press yes/no-buttons on the keyboard. Each session started with the same three practice items and lasted less than 30 minutes.

Exclusion of participants

In the French experiment, 10 participants had an accuracy rate of less than 85% on the comprehension questions. Since the questions were very easy, we assumed these participants were not sufficiently concentrated on understanding the sentences. We therefore decided to exclude them from the analysis. The rest of the French participants had an accuracy rate of above 95%. In the English experiment, 13 participants with an accuracy rate of less than 85% were also excluded from the analysis; the rest had an accuracy rate of above 90%. Following these exclusions, we analyzed the data of 52 French-speaking participants and 37 English-speaking participants.

Analysis

The items were divided into four regions (see Table 2). The critical region included the embedded verb. According to our hypothesis, and following the results from the acceptability judgment study, tense mismatch should generate longer reading times at this disambiguating region only in PR-compatible environments (with perceptual verbs) in French, but not in English. Besides the embedded verb region we also analyzed the pre-critical region with the second noun and the complementizer, to make sure that the relevant effects do not emerge prior to the disambiguation point. Also the post-critical end-of-sentence region was analyzed, to check for possible spill-over effects. The complementizer did not constitute a separate analysis region because there were hardly any fixations on it alone. We merged it into the second region and not into the third one because, unlike the embedded verb, the complementizer does not carry any disambiguating information.

Conditions	Regions			
	First Noun + Verb	Second Noun + Complementizer	Verb	End of Sentence
Perception-match	*Jean a vu*	*la fille qui*	*poussait*	*la femme*.
	John saw	the girl that	pushed	the lady.
Perception-mismatch	*Jean voit*	*la fille qui*	*poussait*	*la femme*.
	John sees	the girl that	pushed	the lady.
Stative-match	*Jean était marié à*	*la fille qui*	*poussait*	*la femme*.
	John was married to	the girl that	pushed	the lady.
Stative-mismatch	*Jean est marié à*	*la fille qui*	*poussait*	*la femme*.
	John is married to	the girl that	pushed	the lady.

Table 2: Eye-tracking experiment: example of an item divided into the four analysis regions

In the analysis we were interested mainly in two dependent variables: regression path duration, which reflects the time readers fixate the region of interest for the first time until they move on to fixate the following region (Konieczny et al. 1997; Liversedge et al. 1998), and the proportion of regressions out of the region (Clifton et al. 2007). We analyzed other eye-tracking measures as well, specifically first pass reading times and total reading times, but we did not necessarily expect to find effects in all of them. The reason for this assumption is that all the sentences in this experiment are not only perfectly grammatical, but also short and involve a relatively simple semantic revision when reanalysis is required, as compared to that required by other well-known cases of reanalysis (as in e.g. classical garden path sentences or long-distance dependencies). In classical garden path sentences (e.g. *the defendant examined by the lawyer turned out to be unreliable*), revision of the initial thematic role assignment (from agent to patient for *the defendant*) is forced at the disambiguation point. Similarly, filled-gap sentences (e.g. *the cat that the dog worried* ***about***__) force revision of the initial thematic assignment to the head of the RC (from experiencer to theme for *the cat*). Our manipulation does not require such a significant revision to thematic role assignment: the head NP of the RC/subject of the PR is in both readings the subject (and agent) of the embedded predicate.[18] Similarly, the present manipulation does not

[18]Notice that a change of argument structure is needed for the matrix predicate, as the matrix verb takes the modified

involve any major syntactic modification of the critical region (in terms of e.g. categorial status or argument structure). Garden path sentences typically involve either voice change (e.g. for the verb *examined* in the previous example) or changes to argument structure, e.g. from transitive to intransitive interpretation (for the verb *mended* in e.g. *while the woman mended the sock fell off her lap*). The same is true of filled-gap sentences, which involve revision from argument to adjunct interpretation. In the present manipulation, the verb stays in the active voice and no changes to argument structure are required.

Finally, the present manipulation does not engender any complexity effects related to locality/ similarity based interference, as those observed in the processing of long-distance dependencies (Gibson, 1998; Gordon et al., 2001; Dyke & Lewis, 2003) as the subject NP and the embedded verb are in a local relation in both the RC and PR parse.

Summarizing, there is no reason to expect particularly strong effects locally at the ROI, as no revision of expected syntactic category or semantic revision is needed and because the disambiguation does not involve similarity based interference. Our manipulation (which targets tense inflection) is also not locally problematic, but might encourage regressions to ensure the main clause tense was properly parsed. Hence, we would not expect to find as strong effects in all of the typical measures that show an effect in other studies as the variables underlying the effect across these different studies vary significantly.

Regression path durations, first pass reading times and total reading times were log-transformed and analyzed with Bayesian Gaussian models (linear mixed-effects models that assume a normal distribution of the dependent variable). The proportion of regressions-out was analyzed with a Bayesian mixed-effects model that assumes a Bernoulli distribution of the dependent variable, using a logit link function. All models were fitted in the *brms* package in R.

In all the models, the levels of the factor Tense were coded as 1 (match) and -1 (mismatch), the levels of the factor Verb Type were coded as 1 (perceptual) and as -1 (stative), and the levels of the factor Language were coded as 1 (French) and -1 (English). Data interpretation will follow

DP as argument under the RC analysis, but the whole PR (perception of an event/situation) under the PR analysis. Notice however, that this is a relatively minor revision, as perception of an event also involves perception of the subject of that event (see also Rizzi (1992) for a claim that PRs in fact require direct perception of their subject and that the thematic role of stimulus is shared by the whole PR and the subject DP in this case).

the same principle as in the previous experiments.

Item order effects

In a first attempt to analyze the data, we fit four models, one for each of the four dependent measures (regression path duration, proportion of regressions-out, first pass duration and total reading times), on the data from the embedded verb region. Each model included, besides the experimental factors, two additional continuous covariates. The length of the region, as expressed by the number of characters it is composed of, was included in order to control for the varying length of the region across different trials. Moreover, we included interactions between the experimental factors and the order of presentation of the items in the experiment. This was done in order to control for any potential learning effects during the experiment and their influence on the experimental manipulations (see below for discussion).

The fixed effects part of the models included the main effects and all possible interactions of the three experimental factors: Tense, Verb Type and Language. The covariate Word Length was centered and only its main effect was estimated (without interactions). Finally, all interactions between the three experimental factors and the centered covariate Item Order were estimated, but not its main effect. The random effects part included adjustments for subjects and items of an intercept, of slopes for the main effects and interaction of Tense and Verb Type, and the correlations between intercepts and slopes.

All model parameters were assigned weakly informative priors. The models were run with 4 chains and 3000 iterations each, whereby the first 1500 iterations were discarded. Model convergence was verified by visually checking the convergence of the chains and by making sure that the $\hat{R}$ statistics for all parameter coefficients was equal to 1 and that the number of effective samples per iteration was reasonably high.

The results showed strong evidence for several model terms involving an interaction of Item Order with one or more of the experimental factors. For instance, in the proportion of regressions-out model, the probability of the four-way interaction Tense by Verb Type by Language by Item Order being greater than zero was 0.96 (90% CrI = [0.001, 0.03]). This shows that the order of presentation of the items influenced processing throughout the experiment differently in the various conditions, and that these effects were dissimilar in the two languages. Figure 6 shows these effects in French for regression path duration and Figure 7 for the proportion of regressions-out (in English, since all the conditions are relative clauses, strong item order effects did not emerge;

this is the reason for the four-way interaction including LANGUAGE).

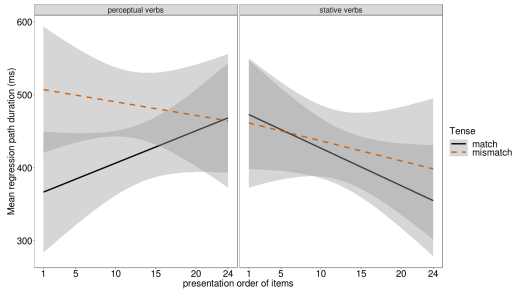


Figure 6: Regression path duration in the embedded verb region in the French data in relation with the order of presentation of the items (x-axis), as a function of Tense (solid line = tense match; dashed line = tense mismatch) and Verb Type (left panel = perceptual verbs; right panel = stative verbs).

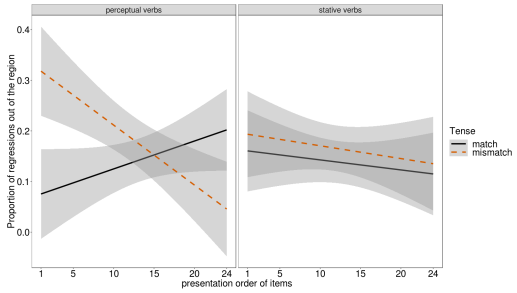


Figure 7: Proportion of regressions out of the embedded verb region in the French data in relation with the order of presentation of the items (x-axis), as a function of Tense (solid line = tense match; dashed line = tense mismatch) and Verb Type (left panel = perceptual verbs; right panel = stative verbs).

In both measures we observe that, as the experiment proceeded and more items were presented, regression path duration and proportion of regressions-out increase in tense-match and decreased in tense-mismatch sentences with perceptual verbs, whereas no such effect emerged under stative verbs. In other words, the difference between tense match and tense mismatch under perceptual verbs in French is reduced while reading more of the experimental items (cf. left panels in Figures 6 and 7).

While we did not expect this effect, following Fernandes et al. (2018), it does have a plausible explanation. Fernandes et al. argue that two aspects of the present study can lead to adaptation. First of all they point out that in our design a PR reading is possible in only one condition (tense match under perceptual verbs), whereas the remaining three conditions can only be parsed as RCs. Moreover, sentences in the PR-compatible condition also allow a RC parse. Fernandes et al. argue that this lack of balance can lead to structural priming effects, generating a stronger preference for the RC reading over the course of the experiment. Fernandes et al. also argue that the design also contains a highly reliable cue which could potentially lead to adaptation: the prediction of a PR is voided when perceptual verbs are in the present tense.

Learning effects of this sort are well-known, as language processing strongly relies on predictive mechanisms which have been shown to adapt to reliable cues (Clayards et al., 2008; Wells et al., 2009; Kamide, 2012; Fine et al., 2010, 2013; Kurumada et al., 2014, a.o.). Repeated exposure to unexpected syntactic structures leads to a reduction of their processing disadvantage over alternative parses (Fine et al. 2013; but see Stack et al. 2018 for contrary evidence). Adaptation effects are thus known to occur and are not necessarily specific to our design and material. Of direct relevance for the present study, Fernandes et al. (2018), using an Italian version of the stimuli from the current experiment, demonstrated that participants indeed adapted to the complex, but highly reliable cue provided by the perceptual verb+present tense, which is always followed by a tense mismatched embedded clause. Moreover, they showed that reducing the reliability of this cue (by adding a small number of unambiguous PRs following perceptual verbs in the present tense) significantly lowered the effect of order of presentation.

Given the strong influence of these effects, in the remainder of this section we will present an analysis of the first half of the trials only, excluding the remaining trials in the data sets of both languages. Our interpretation and discussion will therefore be based on this portion of the data. An analysis of the second half of the experiment, as well as the analysis of the full data set with all items included together, are available under the following link: https://osf.io/v26rx/. This repository contains also the data sets and analysis R scripts for all the studies presented in this article.

Results

Table 3 shows the mean regression path duration and the mean proportion of regressions-out in each sentence region in the French and English experiments. In French, at the embedded

verb region, regression path duration was longer and the proportion of regressions-out higher in sentences with tense mismatch than in sentences with tense match under perceptual verbs. This reflects greater processing difficulty in the former condition. By contrast, when the verb was stative there was no difference between tense match and mismatch. In English, the tense manipulation did not influence processing at the verb region differently under perceptual and stative verbs.

The Bayesian model on the regression path duration in the embedded verb region revealed evidence for the crucial three-way interaction of TENSE by VERB TYPE by LANGUAGE, as indicated by the fact that zero is excluded from the 85% credible intervals of this posterior ($\hat{\beta}$ = -0.03, 85% CrI = [-0.06, -0.002], 95% CrI = [-0.07, 0.009], P($\hat{\beta}$) < 0 = 0.94). The model on the proportion of regressions-out showed even stronger evidence for the same interaction, as suggested by the greater probability that zero is excluded from the posterior distribution ($\hat{\beta}$ = -0.19, 85% CrI = [-0.35, -0.05], 95% CrI = [-0.40, 0.009], P($\hat{\beta}$) < 0 = 0.97). Figure 8 shows the posteriors of the fixed-effects in the model on regression path duration and Figure 9 shows the fixed-effects posteriors in the model on the proportion of regressions-out. Table 4 summarizes the information on the posteriors. [19]

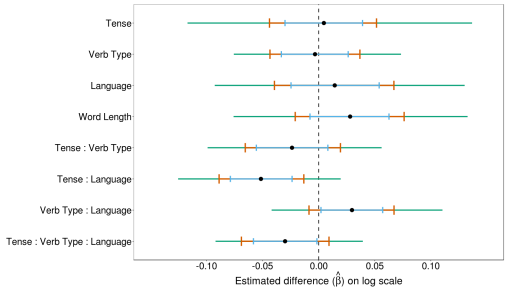


Figure 8: **Regression path duration** in the embedded verb region – posterior distributions of the fixed-effects parameters in the model that includes both languages together. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

[19] In the appendix we provide a table summarizing the first-pass reading times and total reading times data, as well as a summary of the information on the fixed-effects posteriors for the models of these dependent measures. In addition, we provide the results of the analysis in the pre-critical region comprising the second noun and the complementizer, and in the end-of-sentence region in which we check for spill-over effects. The analyses reported in the appendix are not discussed, since they revealed no evidence for the relevant effects.

French data

Verb type	Tense	Regression path duration				Proportion of regressions-out		
		First noun + verb	Second noun + complementizer	Embedded verb	End of sentence	Second noun + complementizer	Embedded verb	End of sentence
Perceptual	Match	670 (74.88)	493 (70.77)	396 (77.52)	1310 (167.13)	0.2 (0.07)	0.12 (0.06)	0.97 (0.04)
	Mismatch	657 (83.71)	519 (86.69)	495 (88.27)	1218 (166.42)	0.23 (0.08)	0.25 (0.07)	0.91 (0.05)
Stative	Match	845 (86.96)	526 (85.77)	420 (71.58)	1242 (191.1)	0.38 (0.11)	0.14 (0.06)	0.86 (0.08)
	Mismatch	723 (71.04)	483 (66.44)	431 (88.36)	1169 (149.56)	0.32 (0.09)	0.15 (0.07)	0.86 (0.07)

English data

Verb type	Tense	Regression path duration				Proportion of regressions-out		
		First noun + verb	Second noun + complementizer	Embedded verb	End of sentence	Second noun + complementizer	Embedded verb	End of sentence
Perceptual	Match	600 (78.17)	702 (128.33)	445 (110.28)	1179 (185.71)	0.31 (0.1)	0.32 (0.12)	0.91 (0.07)
	Mismatch	656 (71.96)	728 (126.76)	383 (74.64)	1326 (166.37)	0.3 (0.12)	0.22 (0.09)	0.87 (0.1)
Stative	Match	832 (78.85)	562 (96.77)	474 (73.66)	1131 (162.07)	0.22 (0.08)	0.29 (0.11)	0.82 (0.12)
	Mismatch	859 (94.22)	624 (88.87)	442 (91.94)	1236 (162.07)	0.28 (0.11)	0.23 (0.1)	0.83 (0.09)

Table 3: Mean regression path duration in milliseconds and mean proportion of regressions-out (with 95% confidence intervals) in the French and English data, divided into the various regions of the sentence and broken by the four conditions.

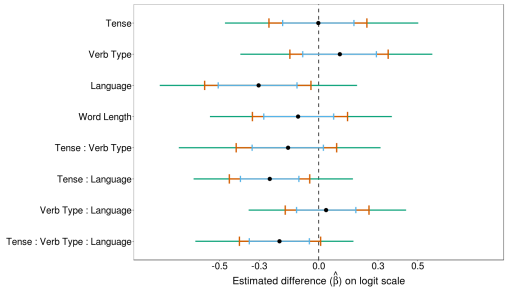


Figure 9: **Proportion of regressions out** of the embedded verb region – posterior distributions of the fixed-effects parameters in the model that includes both languages together. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

Like in the acceptability judgment experiment, in the next step of the analysis we want to break down the three-way interaction and see whether there is evidence for the Tense by Verb Type interaction in French but not in English, as predicted. For this purpose, we fit separate models, one on the French data and one on the English data, for each of the dependent measures in the embedded verb region. Each model included in the fixed-effects part parameters for the main effect of Tense and Verb Type as well as their interaction. In the random-effects part, the models included intercepts for subjects and items.

As can be seen in Figure 10, the analysis of the regression path duration in French showed evidence for the main effect of Tense ($\hat{\beta}$ = -0.05, 95% CrI = [-0.09, -0.003], P($\hat{\beta}$) < 0 = 0.98), meaning that regression path duration was longer for sentences with tense mismatch than with tense match. Importantly, there was also evidence for the interaction Tense by Verb Type ($\hat{\beta}$ = -0.05, 85% CrI = [-0.08, -0.01], 95% CrI = [-0.09, 0.002], P($\hat{\beta}$) < 0 = 0.97). This interaction reflects the greater processing cost of tense-mismatch sentences under perceptual verbs, as compared to tense-match sentences, with no effect of tense under stative verbs. In English, although there was evidence for the main effect of Tense ($\hat{\beta}$ = 0.06, 95% CrI = [0.002, 0.12], P($\hat{\beta}$) < 0 = 0.02), there was no evidence for the main effect of Verb Type and, crucially, not for the interaction of the two factors ($\hat{\beta}$ = 0.003, 85% CrI = [-0.04, 0.04], 95% CrI = [-0.05, 0.06], P($\hat{\beta}$) < 0 = 0.47).

	Regression path duration			Proportion of regressions-out		
Effect	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**
Tense	0.005	[-0.03, 0.04]	0.42	-0.002	[-0.18, 0.18]	0.50
Verb Type	-0.003	[-0.03, 0.03]	0.56	0.12	[-0.08, 0.29]	0.19
Language	0.01	[-0.02, 0.05]	0.29	-0.30	[-0.51, -0.11]	0.99
Word Length	0.03	[-0.008, 0.06]	0.13	-0.10	[-0.28, 0.08]	0.81
Tense : Verb Type	-0.02	[-0.06, 0.008]	0.87	-0.15	[-0.34, 0.02]	0.89
Tense : Language	-0.05	[-0.08, -0.02]	0.99	-0.25	[-0.39, -0.10]	0.99
Verb Type : Language	0.03	[0.002, 0.06]	0.06	0.04	[-0.11, 0.18]	0.37
Tense : Verb Type : Language	-0.03	[-0.06, 0.002]	0.94	-0.19	[-0.35, -0.05]	0.97

Table 4: Regression path duration and proportion of regressions-out in the embedded verb region – summary of the posteriors of the fixed-effects parameters in the models that include both languages together. For each effect we provide the estimated mean of the posterior, the 85% credible intervals and the probability that the posterior is smaller than zero. For the regression path duration, the estimated mean and the 85% credible intervals are on log scale; for the proportion of regressions-out they are on logit scale.

The posteriors of the models for the proportion of regressions-out are shown in Figure 11. These models had similar results to those of the regression path models. In French the model revealed evidence for the main effect of Tense ($\hat{\beta}$ = -0.26, 95% CrI = [-0.51, 0.009], P($\hat{\beta}$) < 0 = 0.98) and for the interaction Tense by Verb Type ($\hat{\beta}$ = -0.26, 95% CrI = [-0.51, 0.007], P($\hat{\beta}$) < 0 = 0.98). In English there was only evidence for the main effect of Tense ($\hat{\beta}$ = 0.23, 85% CrI = [0.04, 0.42], 95% CrI = [-0.02, 0.49], P($\hat{\beta}$) < 0 = 0.96), but not for the interaction ($\hat{\beta}$ = 0.09, 85% CrI = [-0.10, 0.29], 85% CrI = [-0.17, 0.36], P($\hat{\beta}$) < 0 = 0.26) and neither for the main effect of Verb Type. Table 5 summarizes the fixed-effects posteriors from the models of regression path duration and proportion of regressions-out that were fitted separately on the French and the English data.

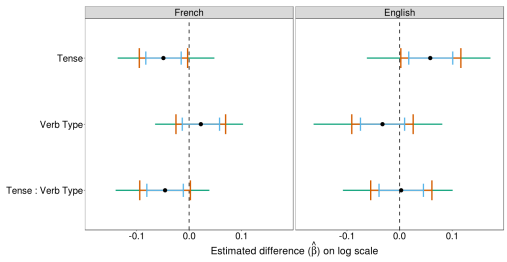


Figure 10: **Regression path duration** at the embedded verb region – posterior distributions of the fixed-effects parameters in the models fitted separately for French (left panel) and English (right panel). The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

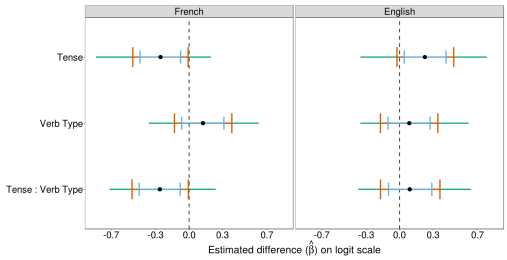


Figure 11: **Proportion of regressions out** of the embedded verb region – posterior distributions of the fixed-effects parameters in the models fitted separately for French (left panel) and English (right panel). The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

In sum, the outcome of the eye-tracking experiments in French and English add to (and support the interpretation of) the results of the previous offline experiments we presented in two ways. First, we replicated the finding of a tense-match advantage under perpceptual verbs in French but not in English. This indicates that the preference for the PR interpretation is active at the earliest stages of processing. Second, it clarifies the results of the acceptability study, as it shows that the source of low rating of unambiguous RCs under stative verbs is different from that of the tense mismatch under perceptual verb. The latter, as shown by the eye-tracking results, is due to reanalysis, while the former is due to independent factors (e.g. inherent complexity of RCs also in the absence of ambiguity).

5. General discussion

In this paper, we set out to investigate the processing of the PR/RC ambiguity, with the double goal of clarifying the timing of this disambiguation and testing the *PR-first Hypothesis*, that is, the claim that the parser displays a structural preference for PRs over RCs.

Previous results, based on RC attachment preferences, indirectly supported this hypothesis. Resolution of this ambiguity in the absence of attachment ambiguities, however, had not been tested directly so far.

We presented three sets of experiments: one sentence completion task assessing effects of PR availability on RC attachment in French, two acceptability judgment tasks and two eye-tracking while reading studies in French and English. Each experiment adds to the results of the previous ones, providing an increasingly clearer picture on the processing of PRs and RCs.

The first experiment on effects of PR availability on RC attachment in French adds to previous work in this domain and further supports the *PR-first Hypothesis*. Once PR availability is controlled for, by using matrix verbs that only select for entities as complements, a Low Attachment preference is observed also in French. This is despite the fact that previously this language has consistently been shown to display a High Attachment preference for RCs. When PRs are made available, using perceptual matrix verbs, which can also select events, the usual High Attachment preference is observed, as previously shown for Italian, Portuguese and Spanish, but not for non-PR languages like English (Grillo et al., 2015). This first study, aside from adding an important

	Regression path duration			Proportion of regressions-out		
Effect	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**
French						
Tense	-0.05	[-0.08, -0.02]	0.98	-0.26	[-0.44, -0.08]	0.98
Verb Type	0.02	[-0.01, 0.06]	0.18	0.12	[-0.07, 0.31]	0.18
Tense : Verb Type	-0.05	[-0.08, -0.01]	0.97	-0.26	[-0.45, -0.08]	0.98
English						
Tense	0.06	[0.02, 0.10]	0.02	0.23	[0.04, 0.42]	0.04
Verb Type	-0.03	[-0.07, 0.01]	0.86	0.09	[-0.10, 0.27]	0.25
Tense : Verb Type	0.003	[-0.04, 0.04]	0.47	0.09	[-0.10, 0.29]	0.26

Table 5: Regression path duration and proportion of regressions-out in the embedded verb region – summary of the posteriors of the fixed-effects parameters in the models fitted separately for French (top) and English (bottom). For each effect we provide the estimated mean of the posterior, the 85% credible intervals and the probability that the posterior is smaller than zero. For the regression path duration, the estimated mean and the 85% credible intervals are on log scale; for the proportion of regressions-out they are on logit scale.

piece to the RC attachment literature, provides a baseline for the following studies that test the interplay between the verb type and tense manipulations. Having established that PR availability plays a role in the processing of embedded finite clauses in French, we moved on to ask whether this effect is also observable in the absence of attachment ambiguities and, if so, how does it unfold in time.

To this aim, we designed an acceptability judgment task which capitalizes on a well-known asymmetry between PRs and RCs: the constraint on the PR tense to be anaphoric to, or match that of the matrix clause. We compared acceptability ratings of (perfectly grammatical) embedded clauses which either matched or mismatched the matrix clause in tense specification. The (mis)matching clauses were embedded within either perceptual or stative verbs. We reasoned that a PR preference might generate a higher acceptability for PR-compatible (tense matching) embedding, over PR-incompatible (tense mismatching) embedded clauses. As RCs do not require tense matching, we did not expect any effects. We further predicted the effect to be language dependent. A disadvantage for tense mismatch under perceptual verbs should only be observed in PR-languages (e.g. French), but not in non-PR languages (e.g. English). The reason is that it depends on PR availability and it is not tied, for instance, to an interaction between the semantics of the matrix predicate and tense (mis)match.

The results fully support our predictions, showing an interaction between verb type and tense in the desired direction and only for French. A mismatch in tense between the matrix and embedded predicates leads to significantly lower acceptability rate under perceptual verbs. Since all the target sentences used in this experiment were perfectly grammatical, we attribute the lower acceptability to the processing cost of reanalysis, triggered by the tense mismatch, from the originally preferred PR to the more complex RC.

Finally, we conducted two eye-tracking studies, in French and English, using the same design and materials from the acceptability studies. This final set of experiments further strengthens the interpretation of the acceptability judgments studies and contributes a valuable insight into the timing of the PR/RC ambiguity resolution. We were able to replicate the tense match advantage observed uniquely in French under perceptual verbs in eye-fixations at the disambiguating embedded verb region. Shorter regression path duration and a smaller proportion of regressions-out were found for tense match than for tense mismatch sentences at the embedded verb exclusively under perceptual verbs and only in French but not in English.

It is worth noting that the evidence for the various main effects and interactions in the eye-tracking data is based on posterior distributions that exclude zero with probability slightly smaller than 1.0. Importantly, this can be expected since we let participants read sentences that are all grammatical and with a structure that is relatively easy to parse. As discussed above, reanalysis in the present case does not involve factors known to significantly raise complexity in processing in garden-path sentences or complex structures involving word order change (like object relative clauses). Specifically, no major modification of argument structure, like inversion of thematic role assignment or similarity based interference, is involved here. Hence, we cannot expect to observe an effect-size comparable to the one found in typical studies on garden-path effects and other processing difficulties.

These results are consistent with the proposed preference for PRs over RCs and indicate that this is a syntactic preference present at the earliest stages of parsing and not determined by later intepretive components. Alternative interpretations of the results, e.g. in terms of frequency distribution of PR vs. RCs are difficult to test. This is because establishing the relative frequency of these string identical structures (i.e. deciding whether a given occurrence of *see* + *DP* + *that* in a corpus should be counted a PR or a RC) is problematic as it would require making non-trivial decisions about disambiguation, which in general cannot be done automatically and, more often than not, cannot be done conclusively. Assuming this could be done successfully, there remains the problem of *directionality of the effect*. That is, are PRs preferred because they are more frequent or are they more frequent because they are more preferred on independent grounds? While frequency effects on processing complexity are well-known, we find that grounding frequency distribution in independently motivated principles of economy of computation provides a more principled explanation for the regularity of the effects described here. The preference for PRs over RCs, in fact, aligns with a more general preference for more parsimonious structures.

Taken together, our results provide strong direct support for the *PR-first Hypothesis*. A preference for PRs emerges both in acceptability judgments and eye-fixations with perfectly grammatical sentences. Forcing a RC reading of otherwise PR-compatible sentences leads to lower acceptability rating and greater processing cost as expressed by longer regression path duration and a bigger proportion of regressions-out. PR availability also leads to stronger preference for High Attachment of RCs in French and, crucially, its unavailability leads to Low Attachment, supporting a universal preference to attach incoming material to the most local host as one of the central

factors in RC attachment ambiguities.

These results further show that cross-linguistic asymmetries in parsing preferences of RC attachment are epiphenomenal and greatly modulated by PR availability (Grillo, 2012), among other grammatical factors (Gilboy et al., 1995; Frazier & Clifton, 1996; Hemforth et al., 2000, 2015).

Acknowledgements

Early versions of this work were presented at CUNY 2015 and AMLaP 2015 (Grillo, Hemforth, Pozniak & Santi 2015a,b *Pseudo Relatives are easier than Relative Clauses: (eye-tracking) evidence from Tense*). We thank the conferences' organizers, reviewers and participants for their comments and suggestions. We also thank 3 anonymous JML reviewers for insightful comments and criticisms on a previous version of this manuscript. We also want to thank Christoph Scheepers and the Institute of Neuroscience and Psychology at the University of Glasgow for giving us the opportunity to test participants in their eye tracking lab. Thanks to Xinjia Rachel Shen for help with data collection and to Bruno Fernandes for insightful comments and for help with data collection in London.

This project was partly funded by the grant "Universal Parsing and Grammatical Variation" (Investigador FCT IF/00876/2013) awarded by the Fundação para a Ciência e a Tecnologia to Nino Grillo, by the Deutsche Forschungsgemeinschaft DFG *Leibniz Prize* AL 554/8–1 (to Artemis Alexiadou) and by the Excellence Cluster "Empirical Foundations of Linguistics" (ANR/CGI 10-LABX-0083).

References

Agresti, A. (2012). *Categorical Data Analysis*. Wiley.

Aguilar, M., & Grillo, N. (2016). Testing the effect of pseudo relatives on relative clause attachment in Spanish. In *AMLaP (Architectures and Mechanisms for Language Processing)*. Bilbao, Spain.

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–281.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*, 390–412.

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

Brysbaert, M., & Mitchell, D. (1996). Modifier attachment in sentence parsing: Evidence from dutch. *Quarterly Journal of Experimental Psychology*, *49A*, 664–695.

Bürkner, P.-C. (2018). brms: An R package for bayesian multilevel models using stan. *Journal of statistical software*, *80*, 1–28.

Burzio, L. (1986). *Italian Syntax: A Government and Binding Approach*. Dordrecht: D. Reidel.

Casalicchio, J. (2013). *Pseudorelative, gerundi e infiniti nelle varietà romanze: Affinità solo superficiali e corrispondenze strutturali*. Ph.D. thesis Università degli Studi di Padova.

Christensen, R. H. B., & Brockhoff, P. B. (2013). Analysis of sensory ratings data with cumulative link models. *Journal de la Societe Francaise de Statistique & Revue de Statistique Appliquee*, *154*, 58 – 79.

Cinque, G. (1992). The Pseudo-Relative and Acc-*ing* constructions after verbs of perception. In *University of Venice Working Papers in Linguistics*. Università di Venezia.

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, *108*, 804–809.

Clifton, C. J., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain*. Amsterdam, Netherlands: Elsevier.

Colonna, S., & Pynte, J. (2001a). Relative clause attachment in french: The role of fodor's "same size sister" constraint. Poster presented at the Prosody in Processing workshop, Utrecht, July 5-6.

Colonna, S., Pynte, J., & Mitchell, D. (2000). Relative clause attachment in french: The role of constituent length. Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA, March 30 - April 1.

Côté, M.-H. (1999). Quantification existentielle sur des événements et structure des pseudorelatives. In F. Corblin, C. Dobrovie-Sorin, & J.-M. Marandin (Eds.), *Empirical issues in formal syntax and semantics 2* (pp. 169–190). The Hague: Thesus.

Crain, S., & Steedman, M. (1985). On not being led up to garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language processing: Psychological, computational, and theoretical perspectives*. Cambridge University Press.

Cuetos, F., Mitchell, D., & Corley, M. (1996). Parsing in different languages. In M. Carreiras, J. García-Albea, & N. S. Gallés (Eds.), *Language processing in Spanish* (pp. 145–187). Mahwah, N.J.: Lawrence Erlbaum.

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing. *Cognition*, *30*, 73–105.

De Vincenzi, M., & Job, R. (1993). Some observations on the universality of the late closure strategy. *Journal of Psycholinguistic Research*, *22*, 189, 206.

De Vincenzi, M., & Job, R. (1995). An investigation of late closure: The role of syntax, thematic structure, and pragmatics in initial interpretation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1303–1321.

Desmet, T., Brysbaert, M., & Baecke, C. D. (2002a). The correspondence between sentence production and corpus frequencies in modifier attachment. *Quarterly Journal of Experimental Psychology*, *55A*, 879–896.

Desmet, T., De Baecke, C., & Brysbaert, M. (2002b). The influence of referential discourse context on modifier attachment in Dutch. *Memory and Cognition*, *30*, 150–157.

Desmet, T., & Gibson, E. (2003). Disambiguation preferences and corpus frequencies in noun phrase conjunction. *Journal of Memory and Language*, *49*, 353–374.

Dobrovie-Sorin, C., Ellsiepen, E., & Hemforth, B. (2016). Why are distributive readings dispreferred? In E. Carrilho, A. Fiis, M. Lobo, & S. Pereira (Eds.), *Romance Languages and Linguistic Theory 10: Selected papers from 'Going Romance' 28, Lisbon*. John Benjamins Publishing Company.

Drummond, A. (2013). Ibex farm. *Online server: http://spellout. net/ibexfarm*, .

Dyke, J. A. V., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285 – 316.

Ehrlich, K., Fernández, E., Fodor, J., Stenshoel, E., & Vinereanu, M. (1999). Low attachment of relative clauses: New data from Swedish, Norwegian and Romanian. Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing. New York, NY, March 18-20.

Fernandes, B. (2012). *Attachment Preferences in Prepositional Infinitive Constructions*. Master's thesis Centro de Linguística da Universidade Nova de Lisboa.

Fernandes, B., Alexiadou, A., Chow, W.-Y., Santi, A., & Grillo, N. (2018). Adaptation to complex cues in sentence comprehension. In *31st CUNY Conference on Human Sentence Processing* (pp. 441–442). UC Davis.

Fine, A., Qian, T., Jaeger, T. F., & Jacobs, R. (2010). Syntactic adaptation in language comprehension. In *Proceedings of the 2010 workshop on cognitive modeling and computational linguistics* (pp. 18–26).

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PLoS ONE*, *8*.

Fodor, J. D. (1998a). Learning to Parse? *Journal of Psycholinguistic Research*, *27*, 285–319.

Fodor, J. D. (1998b). Parsing to Learn? *Journal of Psycholinguistic Research*, *27*, 339–374.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.), *Proceedings of the North East Linguistic Society*. GSLA, University of Massachusetts, Amherst.

Frazier, L. (1990). Parsing modifiers. special purpose routines in the human sentence processing mechanism? In D. Balota, & G. F. d'Arcais (Eds.), *Comprehension Processes in Reading* (pp. 303–330). Hillsdale, N.J.: Lawrence Erlbaum.

Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

Frenck-Mestre, C., & Pynte, J. (2000). Romancing syntactic ambiguity: Why the French and the Italians don't see eye to eye. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp. 549–564). Oxford, UK: Elsevier.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, *59*, 23–59.

Gibson, E., & Schütze, C. (1999). Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. *Journal of Memory and Language*, *40*, 263–279.

Gilboy, E., Sopena, J., Clifton, C., & Frazier, L. (1995). Argument structure and association preferences in spanish and english complex NPs. *Cognition*, *54*, 131–167.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1411 – 1423.

Graffi, G. (1980). Su alcune costruzioni "pseudorelative". *Rivista di Grammatica Generativa*, *5*, 115–139.

Grillo, N. (2012). Local and universal. In V. Bianchi, & C. Chesi (Eds.), *Enjoy Linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday* (pp. 234–245). Siena, Italy: CISCL Press.

Grillo, N., & Costa, J. (2014). A novel argument for the universality of parsing principles. *Cognition*, *133*, 156–187.

Grillo, N., Costa, J., Fernandes, B., & Santi, A. (2015). Highs and lows in english attachment. *Cognition*, *144*, 116 – 122.

Grillo, N., Fernandes, B., & Costa, J. (2012a). Attachment preferences in Prepositional Infinitive Constructions in European Portuguese. In *AMLaP (Architectures and Mechanisms for Language Processing)*. Riva del Garda, Italy.

Grillo, N., Fernandes, B., & Costa, J. (2012b). Universally local attachment: New evidence from PIC. In *ERP Experimental Psycholinguistics Conference*. Madrid, Spain.

Grillo, N., & Moulton, K. (2016). Event kinds and the pseudo relative. In B. Prickett, & C. Hammerly (Eds.), *NELS 46: Proceedings of the Forty-Sixth Annual Meeting of the North East Linguistics Society* (pp. 11–20). Amherst MA: GLSA volume 2.

Grillo, N., & Spathas, G. (2014). Tense and Aspect modulate RC attachment: Testing the PR hypothesis in Greek. In *DGfS 36 Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft*. Oldenburg, Germany.

Grillo, N., Tomaz, M., Lourenço Gomes, M., & Santi, A. (2013). Pseudo relatives vs. Relative clauses: Greater preference, Lower costs. In *AMLaP (Architectures and Mechanisms for Language Processing)*. Marseille, France.

Grillo, N., & Turco, G. (2016). Prosodic disambiguation and attachment height. In *Proceedings of Proceedings of the 8th International Speech Prosody Conference. Boston, May 31–June 3 2016*.

Guasti, M. (1988). La pseudorelative et les phénomenènes d'accord. *Rivista di Grammatica Generativa*, *13*, 35–57.

Guasti, M. (1992). Progressive in the complements of perception verbs. In Bonomi, Casalegno, & Zwarts (Eds.), *Proceedings of the Gargano Conference on "Perceptual Reports"*.

Gutierrez-Ziardegi, E., Carreiras, M., & Laka, I. (2004). Bilingual sentence processing: Relative clause attachment in basque and spanish. In *CUNY Conference on Human Sentence Processing*. University of Maryland.

Hemforth, B., Fernandez, S., Clifton, C., Frazier, L., Konieczny, L., & Walter, M. (2015). Relative clause attachment in german, english, spanish and french: Effects of position and length. *Lingua*, *166*, 43–64.

Hemforth, B., Konieczny, L., & Scheepers, C. (2000). Syntactic attachment and anaphor resolution: Two sides of relative clause attachment. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and Mechanisms for Language Processing* (pp. 259–281). Cambridge, UK: Cambridge University Press.

Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E.-J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic bulletin & review*, *21*, 1157–1164.

Kamide, Y. (2012). Learning individual talkers structural preferences. *Cognition*, *124*, 66–71.

Kayne, R. (1975). *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Koenig, J.-P., & Lambrecht, K. (1999). French relative clauses as secondary predicates. In F. Corblin, C. Dobrovie-Sorin, & J.-M. Marandin (Eds.), *Empirical issues in Formal Syntax and Semantics 2* (pp. 191–214). Thesus.

Konieczny, L., Hemforth, B., Scheepers, C., & Strube, G. (1997). The role of lexical heads in parsing: Evidence from german. *Language and Cognitive Processes*, *12*, 307 – 348.

Kruschke, J. (2015). *Doing bayesian data analysis: a tutorial introduction with R, JAGS and STAN*. London: Academic Press.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*, 573.

Kruschke, J. K., & Liddell, T. M. (2017). Bayesian data analysis for newcomers. *Psychonomic bulletin & review*, (pp. 1–23).

Kurumada, C., Brown, M., Bibyk, S., Pontillo, D., & Tanenhaus, M. (2014). Rapid adaptation in online pragmatic interpretation of contrastive prosody. In *Proceedings of the Annual Meeting of the Cognitive Science Society*. volume 36.

Lambrecht, K. (2000). Prédication seconde et structure informationnelle : la relative de perception comme construction présentative. *Langue française*, *127*, 49–66.

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception*. Oxford, UK: Elsevier Science Ltd.

Lovrić, N. (2003). *Implicit prosody in silent reading: Relative clause attachment in Croatian*. Ph.D. thesis City University of New York.

Mackintosh, J. (2012). Run. Creative Commons 4.0, https://www.flickr.com/photos/59207552@N08/7870735914.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type i error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan (Vol. 122)*. Boca Raton: CRC Press.

Miles, W. R. (1930). Ocular dominance in human adults. *The journal of general psychology*, *3*, 412–430.

Mitchell, D., Cuetos, F., & Zagar, D. (1990). Reading in different languages: is there a universal mechanism for parsing sentences? In D. Balota, & G. F. d'Arcais (Eds.), *Comprehension processes in reading* (pp. 285–302). Hillsdale, N.J.: Lawrence Erbaum.

Mitchell, D. C., & Brysbaert, M. (1998). Challenges to recent theories of crosslinguistic variation in parsing: Evidence from Dutch. In D. Hillert (Ed.), *Sentence Processing: A Crosslinguistic Perspective* (pp. 313–335). San Diego, CA: Academic Press.

Mitchell, D. C., Brysbaert, M., & Swanepoel, S. (2000). Modifier attachment in Dutch: Testing aspects of Construal Theory. In D. H. A. Kennedy, R. Radach, & J. Pynte (Eds.), *Reading as a Perceptual Process* (pp. 493–516). Oxford, UK: Elsevier.

Mitchell, D. C., & Cuetos, F. (1991). The origin of parsing strategies. In C. Smith (Ed.), *Current issues in Natural Language Processing* (pp. 1–12). Center for Cognitive Science, U. of Austin, TX.

Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, *24*, 469–488.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic bulletin & review*, *23*, 103–123.

Moulton, K., & Grillo, N. (2015). Pseudo relatives: Big but Transparent. In T. Bui, & D. Ozyildiz (Eds.), *NELS 45: Proceedings of the Forty-fifth Annual Meeting of the North East Linguistic Society,* (pp. 193–202). MIT Cambridge, MA.

Nicenboim, B., Roettger, T. B., & Vasishth, S. (2018). Using meta-analysis for evidence synthesis: The case of incomplete neutralization in german. *Journal of Phonetics*, *70*, 39 – 55.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational ideas–part ii. language and linguistics compass, 10 (11), 591–613.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in l1 and l2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *25*, 501–528.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria.

Radford, A. (1975). Pseudo-relatives and the unity of subject-raising. *Archivum Linguisticum*, *6*, 32–64.

Rafel, J. (1999). *Complex Small Clauses*. Ph.D. thesis Universitàt Autonoma de Barcelona.

Rizzi, L. (1992). Direct Perception, Government and Thematic Sharing. *Geneva Generative Papers*, *1*, 39–52.

Shen, X. (2006). *Late assignment of syntax theory: evidence from Chinese and English*. Ph.D. thesis University of Exeter.

Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using stan: A tutorial for psychologists, linguists, and cognitive scientists. quantitative methods for psychology.

Stack, C. M. H., James, A. N., & Watson, D. G. (2018). A failure to replicate rapid syntactic adaptation in comprehension. *Memory and Cognition*, *46*, 864 – 877.

Tomaz, M., Lourenço Gomes, M., Santi, A., & Grillo, N. (2014). A concordância de número em construções relativas e pseudorelativas em Português europeu. In *Textos Selecionados do XXIX Encontro Nacional da Associação Portuguesa de Linguística*.

Vasishth, S., & Nicenboim, B. (2016). Statistical methods for linguistic research: Foundational ideas–parti i. *Language and Linguistics Compass*, *10*, 349–369.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic bulletin & review*, *14*, 779–804.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive psychology*, *58*, 250–271.

Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early-closure attachment on first-pass reading times in French. *Quarterly Journal of Experimental Psychology*, *50A*, 421–438.

Zwerink, S. (2011). My hero: Usain bolt. Creative Commons, 4.0. https://www.flickr.com/photos/stevenzwerink/6237256592. Accessed: 2018–03–05.

Appendix A. Items from Experiment 1: French completion task

Mean % of High Attachment is indicated for each item.

1. a. Jean voit le fils du médecin qui bricole. 53.84
   J. sees the son of the doctor that tinkers.
   b. Pierre partage la maison avec le fils du médecin qui bricole. 68.75
   P. shares the house with the son of the doctor that tinkers.
2. a. Kelly entend la grand-mère de la fille qui fait le ménage. 38.88
   K. hears the grandmother of the girl that does the housework.
   b. Kelly travaille avec la grand-mère de la fille qui fait le ménage. 0
   K. works with the grandmother of the girl that does the housework.
3. a. Jean entend le professeur du garçon qui chante. 50
   J. hears the professor of the boy that sings.

b. Jean court avec le professeur du garçon qui chante. 0

J. runs with the professor of the boy that runs.

4. a. L'écrivain regarde la tante de la fille qui jongle. 15

The writer looks at the aunt of the girl that juggles.

b. L'écrivain est marié à la tante de la fille qui jongle. 5.26

The writer is married to the aunt of the girl that juggles.

5. a. Marie écoute le fils du policier qui murmure. 92.3

M. listens to the son of the policeman that whispers.

b. Marie est employée par le fils du policier qui murmure. 53.33

M. is employed by the son of the policeman that whispers.

6. a. Marie observe l'ami du député qui cuisine. 70.58

M. observes the friend of the congressman that cooks.

b. Marie est fiancé à l'ami du député qui cuisine. 20

M. is engaged with the friend of the congressman that cooks.

7. a. Jeanne surprend la domestique de l'actrice qui vole. 75

J. surprises the maid of the actress that steals.

b. Jeanne s'entraîne avec la domestique de l'actrice qui vole. 23.07

J. trains with the maid of the actress that steals.

8. a. L'avocat surprend le chauffeur du voisin qui nage. 42.10

The lawyer surprises the chauffeur of the neighbour that swims.

b. L'avocat s'entraîne avec le chauffeur du voisin qui nage. 31.57

The lawyer trains with the chaffeur of the neighbour that swims.

9. a. David observe la fille de la domestique qui s'entraîne. 92.30

D. observes the daughter of the maid that trains.

b. Marc est divorcé de la fille de la domestique qui s'entraîne. 25

M. is divorced from the daughter of the maid that trains.

10. a. Alain observe la nièce de l'infirmière qui patine. 84.21

A. observes the niece of the nurse that skates.

b. Alain est lié à la nièce de l'infirmière qui patine. 44.44

A. is linked to the niece of the nurse that skates.

11. a. Jeanne photographie le collègue du boucher qui court. 75

J. photographs the colleague of the butcher who runs.

	b.Jeanne danse avec le collègue du boucher qui court.	38.46
	J. dances with the colleague of the butched that runs.	
12.	a. Cathy regarde l'ami du juge qui peint.	80
	C. looks at the friend of the judge that paints.	
	b. Cathy est fiancée à l'ami du juge qui peint.	15.78
	C. is engaged with the friend of the judge that paints.	
13.	a. Lily imagine l'amie de la fleuriste qui travaille.	42.85
	L. imagines the friend of the florist that works.	
	b. Lily fait la fête avec l'amie de la fleuriste qui travaille.	68.75
	L. parties with the friend of the florist that works.	
14.	a. Rachel rêve de l'ami du frère qui boit.	21.05
	R. dreams of the friend of the brother that drinks.	
	b. Rachel est marié à l'ami du frère qui boit.	15
	R. is married to the friend of the brother that drinks.	
15.	a. David dessine le petit-fils de l'homme qui fume.	66.66
	D. draws the granddaughter of the man that smokes.	
	b. David est employé par le petit-fils de l'homme qui fume.	30.76
	D. is employed by the granddaughter of the man that smokes.	
16.	a. Philippe filme l'agent de l'acteur qui ronfle.	60
	P. films the agent of the actor that snores.	
	b. Philippe passe du temps avec l'agent de l'acteur qui ronfle.	15.78
	P. spends time with the agent of the actor that snores.	
17.	a. Le pompier enregistre le cousin de l'avocat qui siffle.	64.28
	The firefighter records the cousin of the lawyer that whistles.	
	b.Le pompier est employé par le cousin de l'avocat qui siffle.	31.25
	The firefighter is employed by the cousin of the lawyer that whistles.	
18.	a. Léa aperçoit l'ami du cordonnier qui danse.	73.68
	L. perceives the friend of the shoemaker that dances.	
	b. Léa est fiancée à l'ami du cordonnier qui danse.	35
	L. is engaged to the friend of the shoemaker that dances.	
19.	a. Sally photographie la belle-fille de l'infirmière qui étudie.	56.25
	S. photographs the daughter-in-law of the nurse that studies.	

b.Sally collabore avec la belle-fille de l'infirmière qui étudie. 57.14

S. collaborates with the daughter-in-law of the nurse that studies.

20. a. Le chanteur regarde le frère du PDG qui saigne. 90

The singer looks at the brother of the CEO that bleeds.

b. Le chanteur étudie avec le frère du PDG qui saigne. 27.77

The singer studies with the brother of the CEO that bleeds.

21. a. Le policier filme l'amie de la sœur qui tricote. 69.23

The policeman films the friend of the sister that knits.

b.Le policier est marié à l'amie de la sœur qui tricote. 50

The policeman is married to the friend of the sister that knits.

22. a. L'architecte imagine la sœur de la collègue qui danse. 36.84

The architect imagines the sister of the colleague that dances.

b. L'architecte est divorcé de la sœur de la collègue qui danse. 5

The architect is divorced from the sister of the colleague that dances.

23. a. David voit le professeur de l'ami qui pilote. 46.66

D. sees the professor of the friend that flies.

b. David fait la fête avec le professeur de l'ami qui pilote. 14.28

D. parties with the professor of the friend that flies.

24. a. Le voisin écoute le fils du concierge qui chante. 89.47

The neighbour listens to the son of the porter that sings.

b. Le voisin va à l'université avec le fils du concierge qui chante. 31.57

The neighbour goes to the university with the son of the porter that sings.

Appendix B. Items for French *acceptability* and *eye-tracking* studies

Mean acceptability rate (scale 1–10) is indicated for each item.

Legend:

a. Perception–Match
b. Perception–Mismatch
c. Stative–Match
d. Stative–Mismatch

1. a. Pierre a vu le garçon qui arrosait la fille avec le tuyau. 7.5

b. Pierre voit le garçon qui arrosait la fille avec le tuyau. 7.5

c. Pierre a été ami avec le garçon qui arrosait la fille avec le tuyau. 6.2

d. Pierre est ami avec le garçon qui arrosait la fille avec le tuyau. 7.9

2. a. Léa a entendu le clown qui imitait le magicien. 7.6
   b. Léa entend le clown qui imitait le magicien. 7.4
   c. Léa a été fiancée au clown qui imitait le magicien. 8.0
   d. Léa est fiancée au clown qui imitait le magicien. 8.0

3. a. Le policier a surpris le juge qui discutait avec le ministre. 8.5
   b. Le policier surprend le juge qui discutait avec le ministre. 7.3
   c. Le policier a couru avec le juge qui discutait avec le ministre. 6.9
   d. Le policier court avec le juge qui discutait avec le ministre. 6.6

4. a. L'écrivain a regardé le journaliste qui menaait le sénateur. 8.1
   b. L'écrivain regarde le journaliste qui menaait le sénateur. 7.2
   c. L'écrivain s'est entraîné avec le journaliste qui menaait le sénateur. 7.6
   d. L'écrivain s'entraîne avec le journaliste qui menaait le sénateur. 7.2

5. a. Marie a écouté le ministre qui critiquait le président. 8.5
   b. Marie écoute le ministre qui critiquait le président. 7.7
   c. Marie a été mariée au ministre qui critiquait le président. 7.9
   d. Marie est mariée au ministre qui critiquait le président. 8.0

6. a. Sarah a aperu le policier qui frappait le chauffeur. 9.4
   b. Sarah aperoit le policier qui frappait le chauffeur. 6.3
   c. Sarah a divorcé du policier qui frappait le chauffeur. 6.9
   d. Sarah divorce du policier qui frappait le chauffeur. 6.6

7. a. Jeanne a vu le professeur qui cherchait l'étudiant. 8.5
   b. Jeanne voit le professeur qui cherchait l'étudiant. 6.2
   c. Jeanne a été fiancée au professeur qui cherchait l'étudiant. 5.3
   d. Jeanne est fiancée au professeur qui cherchait l'étudiant. 7.1

9. a. Léa a observé le bijoutier qui irritait le client. 7.5
   b. Léa observe le bijoutier qui irritait le client. 7.3
   c. Léa a collaboré avec le bijoutier qui irritait le client. 7.1
   d. Léa collabore avec le bijoutier qui irritait le client. 7.2

10. a. Le détective a filmé le commerant qui trompait le fournisseur. 8.2
    b. Le détective filme le commerant qui trompait le fournisseur. 7.3
    c. Le détective a été employé par le commerant qui trompait le fournisseur. 6.9
    d. Le détective est employé par le commerant qui trompait le fournisseur. 5.4

12. a. Léa a espionné le professeur qui accueillait le doyen. 8.1
b. Léa espionne le professeur qui accueillait le doyen. 7.8
c. Léa a vécu avec le professeur qui accueillait le doyen. 7.8
d. Léa vit avec le professeur qui accueillait le doyen. 7.5

13. a. Léo s'est représenté la serveuse qui agaçait la dame. 5.7
b. Léo se représente la serveuse qui agaçait la dame. 6.2
c. Léo se représente la serveuse qui agaçait la dame. 7.6
d. Léo est marié à la serveuse qui agaçait la dame. 7.4

14. a. Thomas a regardé la vendeuse qui aidait la cliente. 8.7
b. Thomas regarde la vendeuse qui aidait la cliente. 6.9
c. Thomas a été fiancé à la vendeuse qui aidait la cliente. 7.3
d. Thomas est fiancé à la vendeuse qui aidait la cliente. 7.4

15. a. David a rencontré la danseuse qui courait avec la chanteuse. 7.2
b. David rencontre la danseuse qui courait avec la chanteuse. 6.8
c. David a été ami avec la danseuse qui courait avec la chanteuse. 7.0
d. David est ami avec la danseuse qui courait avec la chanteuse. 7.4

18. a.(match-perception) Léa a enregistré la conductrice qui insultait la victime. 9.4
b. Léa enregistre la conductrice qui insultait la victime. 6.7
c. Léa a logé chez la conductrice qui insultait la victime. 6.9
d. Léa loge chez la conductrice qui insultait la victime. 6.5

19. a. Sally a entendu la soprano qui impressionnait la ballerine. 7.2
b. Sally entend la soprano qui impressionnait la ballerine. 6.4
c. Sally a été amie avec la soprano qui impressionnait la ballerine. 7.4
d. Sally est amie avec la soprano qui impressionnait la ballerine. 7.4

20. a. Le chanteur a écouté la présidente qui critiquait la journaliste. 8.1
b. Le chanteur écoute la présidente qui critiquait la journaliste. 6.6
c. Le chanteur a collaboré avec la présidente qui critiquait la journaliste. 6.7
d. Le chanteur collabore avec la présidente qui critiquait la journaliste. 7.0

21. a. Le caméraman a observé la chirurgienne qui aidait la sage-femme. 8.6
b. Le caméraman observe la chirurgienne qui aidait la sage-femme. 7.3
c. Le caméraman est sorti avec la chirurgienne qui aidait la sage-femme. 6.9
d. Le caméraman sort avec la chirurgienne qui aidait la sage-femme. 7.7

24. a. Le chorégraphe a épié la scénariste qui encourageait la pianiste. 7.4
    b. Le chorégraphe épie la scénariste qui encourageait la pianiste. 7.0
    c. Le chorégraphe a été hébergé par la scénariste qui encourageait la pianiste. 6.5
    d. Le chorégraphe est hébergé par la scénariste qui encourageait la pianiste. 7.2

Appendix C. Items English *acceptability* and *eye-tracking* studies

Mean acceptability rate (scale 1–10) is indicated for each item.

Legend:
- a. Perception–Match
- b. Perception–Mismatch
- c. Stative–Match
- d. Stative–Mismatch

1. a. Peter saw the boy that sprayed water over the girl. 7.6
   b. Peter sees the boy that sprayed water over the girl. 7.5
   c. Peter was friends with the boy that sprayed water over the girl. 8.2
   d. Peter is friends with the boy that sprayed water over the girl. 8.3
2. a. (match-perception) Leah heard the clown that imitated the magician. 7.7
   b. Leah hears the clown that imitated the magician. 7.2
   c. Leah was engaged to the clown that imitated the magician. 7.7
   d. Leah is engaged to the clown that imitated the magician. 7.8
3. a. The policeman was pointing at the judge that argued with the minister. 8.1
   b. The policeman is pointing at the judge that argued with the minister. 7.6
   c. The policeman was jogging with the judge that argued with the minister. 7.2
   d. The policeman is jogging with the judge that argued with the minister. 7.2
4. a. The writer was watching the journalist that threatened the senator. 7.3
   b. The writer is watching the journalist that threatened the senator. 7.4
   c. The writer was training with the journalist that threatened the senator. 7.8
   d. The writer is training with the journalist that threatened the senator. 6.9
5. a. Mary listened to the minister that criticized the president. 7.4
   b. Mary listens to the minister that criticized the president. 7.5
   c. Mary was married to to the minister that criticized the president. 7.9
   d. Mary is married to to the minister that criticized the president. 7.7

6. a. Sarah caught sight of the policeman that hit the driver. 8.7
b. Sarah catches sight of the policeman that hit the driver. 7.8
c. Sarah was divorced from the policeman that hit the driver. 7.1
d. Sarah is divorced from the policeman that hit the driver. 8.3

7. a. Jean saw the professor that looked for the student. 7.6
b. Jean sees the professor that looked for the student. 5.8
c. Jean was engaged to the professor that looked for the student. 5.4
d. Jean is engaged to the professor that looked for the student. 5.9

8. a. Jack observed the postman that attacked the neighbour. 7.8
b. Jack observes the postman that attacked the neighbour. 7.6
c. Jack worked with the postman that attacked the neighbour. 7.9
d. Jack works with the postman that attacked the neighbour. 7.7

9. a. Leah was watching the jeweller that irritated the customer. 6.9
b. Leah is watching the jeweller that irritated the customer. 7.4
c. Leah was working for the jeweller that irritated the customer. 7.8
d. LLeah is working for the jeweller that irritated the customer. 7.6

10. a. The detective filmed the shopkeeper that cheated the supplier. 7.8
b. The detective films the shopkeeper that cheated the supplier. 7.1
c. The detective was employed by the shopkeeper that cheated the supplier. 7.0
d. The detective is employed by the shopkeeper that cheated the supplier. 6.8

11. a. Peter was photographing the butler that attacked the gardener. 7.5
b. Peter is photographing the butler that attacked the gardener. 8.3
c. Peter was living with the butler that attacked the gardener. 7.0
d. Peter is living with the butler that attacked the gardener. 7.8

12. a. Leah spied on the professor that met the dean. 7.0
b. Leah spies on the professor that met the dean. 6.9
c. Leah lived with the professor that met the dean. 7.4
d. Leah lives with the professor that met the dean. 6.2

14. a. Tom was watching the shop assistant that helped the customer. 8.2
b. Tom is watching the shop assistant that helped the customer. 7.5
c. Tom was engaged to the shop assistant that helped the customer. 7.5
d. Tom is engaged to the shop assistant that helped the customer 7.8

15.	a. David was meeting the dancer that jogged with the singer.	7.3
	b. David is meeting the dancer that jogged with the singer.	7.8
	c. David was friends with the dancer that jogged with the singer.	6.6
	d. David is friends with the dancer that jogged with the singer.	7.4
18.	a. Leah was recording the driver that insulted the victim.	8.4
	b. (mismatch-perception)Leah is recording the driver that insulted the victim.	7.4
	c. Leah lived with the driver that insulted the victim.	7.4
	d. Leah lives with the driver that insulted the victim.	7.7
19.	a. Sally heard the soprano that impressed the ballerina.	7.2
	b. Sally hears the soprano that impressed the ballerina.	7.6
	c. Sally was friends with the soprano that impressed the ballerina.	7.1
	d. Sally is friends with the soprano that impressed the ballerina.	7.6
20.	a. The singer heard the manager that criticized the journalist.	6.8
	b. The singer hears the manager that criticized the journalist	6.8
	c. (match-stative)The singer worked with the manager that criticized the journalist.	8.0
	d. The singer works with the manager that criticized the journalist.	7.3
21.	a. The cameraman was watching the surgeon that helped the midwife.	6.9
	b. The cameraman is watching the surgeon that helped the midwife.	7.4
	c. The cameraman went out with the surgeon that helped the midwife.	8.0
	d. The cameraman goes out with the surgeon that helped the midwife.	8.0
22.	a. The architect saw the girl that pushed the lady.	8.0
	b. The architect sees the girl that pushed the lady.	7.4
	c. The architect was friends with the girl that pushed the lady.	7.4
	d. The architect is friends with the girl that pushed the lady.	7.3
23.	a. David caught sight of the manager that bothered the clerk.	8.0
	b. David catches sight of the manager that bothered the clerk.	7.0
	c. David was trained by the manager that bothered the clerk.	6.4
	d. David is trained by the manager that bothered the clerk.	6.4
24.	a. The choreographer spied on the scriptwriter that encouraged the piano player.	6.3
	b. The choreographer spies on the scriptwriter that encouraged the piano player.	6.9
	c. The choreographer housed the scriptwriter that encouraged the piano player.	7.0
	d. The choreographer houses the scriptwriter that encouraged the piano player.	4.7

Appendix D. Justification of the choice of a Bayesian analysis

The Bayesian framework for data analysis allows the incorporation of previous information (prior) into the experimental data that have been collected, yielding a new probability distribution (posterior) that indicates how the prior information should be updated in the light of the observed data (Kruschke 2015; Kruschke & Liddell 2017). The advantages of Bayesian analysis over the traditional frequentist methods are discussed in detail in numerous publications (Wagenmakers 2007; Kruschke 2013, 2015; McElreath 2016; Nicenboim & Vasishth 2016; Sorensen et al. 2016). Here we will only briefly mention some of the motivations for opting for this method.

The most important characteristic of Bayesian analysis is a straightforward interpretation of the results. The posterior distribution of a main effect or an interaction provides information on how reliable the evidence for the effect or the interaction is. This contrasts with *Frequentist* null hypothesis significance testing methods which provide information on the null hypothesis (i.e., the possibility that there is no effect) and not on the hypothesis actually being tested by the experimenter (Vasishth & Nicenboim 2016). Moreover, uncertainty around effects is expressed in a Bayesian framework by means of credible intervals, for instance 95% credible intervals, defined as the portion within which we can be certain with probability 0.95 that the true parameter lies. Again, the credible intervals thus provide direct information on the results. By contrast, traditional confidence intervals merely provide information on sampling techniques (Wagenmakers 2007; Hoekstra et al. 2014; Morey et al. 2016; Nicenboim & Vasishth 2016; Vasishth & Nicenboim 2016).

Another benefit, specific to the use of (generalized) linear mixed-effects models, is flexibility of model fitting. Linear mixed-effects models are known to be most reliable with large amounts of data (Matuschek et al. 2017). Small data sets can be a problem, especially when fitting maximal models, namely models with the maximal structure of random effects allowed by the design (Barr et al., 2013), because of convergence problems. Bayesian model fitting, by contrast, does not fail because of model complexity.

Appendix E. Analysis results with other eye-tracking measures

French data

Verb type	Tense	First pass reading times				Total reading times			
		First noun + verb	Second noun + complementizer	Embedded verb	End of sentence	First noun + verb	Second noun + complementizer	Embedded verb	End of sentence
Perceptual	Match	665 (65.93)	365 (39.81)	312 (32.02)	358 (64.78)	1191 (121.94)	819 (96.59)	650 (79.79)	466 (95.04)
	Mismatch	657 (68)	375 (45.94)	341 (40.45)	380 (45.87)	1108 (89.97)	830 (104.05)	639 (66.29)	531 (71.41)
Stative	Match	838 (75.77)	304 (32.64)	304 (29.15)	382 (70.02)	1514 (136.81)	730 (71.84)	621 (68.64)	498 (89.64)
	Mismatch	723 (49)	319 (30.71)	330 (39.14)	360 (44.43)	1311 (92.41)	750 (122.17)	603 (65.7)	444 (80.25)

English data

Verb type	Tense	First pass reading times				Total reading times			
		First noun + verb	Second noun + complementizer	Embedded verb	End of sentence	First noun + verb	Second noun + complementizer	Embedded verb	End of sentence
Perceptual	Match	600 (62.88)	373 (58.44)	282 (46.58)	481 (121.99)	1103 (148.9)	904 (128.61)	569 (80.88)	607 (130.65)
	Mismatch	656 (79.74)	399 (63.2)	306 (47.24)	407 (75.58)	1146 (134.89)	902 (120.51)	566 (89.13)	624 (87.41)
Stative	Match	812 (78.46)	390 (64.36)	349 (52.86)	414 (80.17)	1268 (123.6)	832 (121.72)	622 (92.73)	535 (92.5)
	Mismatch	850 (85.51)	422 (53.05)	316 (46.32)	405 (65.84)	1262 (125.99)	825 (117.16)	630 (84.46)	620 (88.31)

Table E.6: First pass reading times and total reading times in milliseconds (with 95% confidence intervals) in the French and English data, divided into the various regions of the sentence and broken by the four conditions.

	First pass reading times			Total reading times		
Effect	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**
Tense	-0.004	[-0.03, 0.02]	0.58	-0.01	[-0.04, 0.02]	0.71
Verb Type	-0.02	[-0.04, 0.006]	0.86	-0.01	[-0.04, 0.02]	0.69
Language	0.02	[-0.01, 0.05]	0.17	0.04	[0.002, 0.08]	0.07
Word Length	0.04	[0.02, 0.07]	0.01	0.07	[0.03, 0.11]	0.01
Tense : Verb Type	-0.01	[-0.04, 0.01]	0.77	0.004	[-0.02, 0.03]	0.41
Tense : Language	-0.01	[-0.03, 0.01]	0.74	-0.001	[-0.03, 0.02]	0.54
Verb Type : Language	0.03	[0.01, 0.06]	0.01	0.04	[0.01, 0.06]	0.02
Tense : Verb Type : Language	0.01	[-0.01, 0.03]	0.24	-0.02	[-0.04, 0.009]	0.82

Table E.7: **First pass reading times** and **total reading times** in the **embedded verb region** – summary of the posteriors of the fixed-effects model parameters. For each effect we provide the estimated mean of the posterior, the 85% credible intervals and the probability that the posterior is smaller than zero. For both dependent variables, the estimated mean and the 85% credible intervals are on log scale.

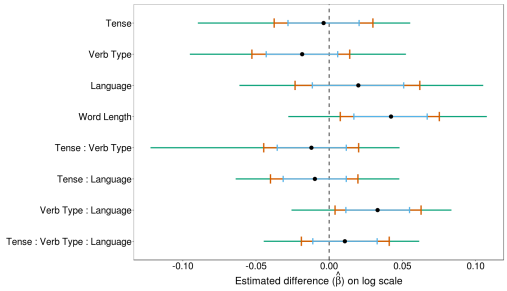


Figure E.12: **First pass reading times** in the **embedded verb region** – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

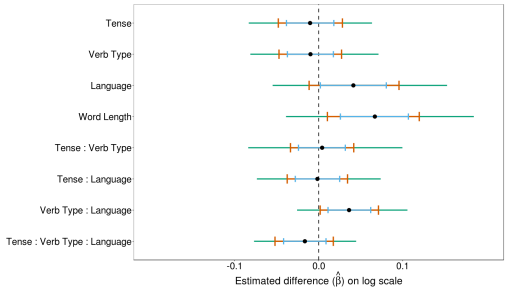


Figure E.13: **Total reading times** in the **embedded verb region** – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

	First pass reading times			Total reading times		
Effect	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**
Tense	-0.02	[-0.05, 0.01]	0.83	-0.009	[-0.03, 0.02]	0.69
Verb Type	0.002	[-0.03, 0.04]	0.46	0.03	[-0.001, 0.05]	0.08
Language	0.006	[-0.03, 0.04]	0.41	-0.03	[-0.07, 0.009]	0.86
Word Length	0.14	[0.11, 0.18]	0	0.16	[0.13, 0.19]	0
Tense : Verb Type	0.03	[0.003, 0.06]	0.06	0.0007	[-0.03, 0.03]	0.48
Tense : Language	0.03	[-0.001, 0.05]	0.08	0.002	[-0.02, 0.03]	0.44
Verb Type : Language	0.04	[0.01, 0.07]	0.03	0.004	[-0.02, 0.03]	0.41
Tense : Verb Type : Language	-0.01	[-0.04, 0.01]	0.77	-0.009	[-0.03, 0.02]	0.71

Table E.8: **First pass reading times** and **total reading times** in the **pre-critical region** comprising the second noun and the complementizer – summary of the posteriors of the fixed-effects model parameters. For each effect we provide the estimated mean of the posterior, the 85% credible intervals and the probability that the posterior is smaller than zero. For both dependent variables, the estimated mean and the 85% credible intervals are on log scale.

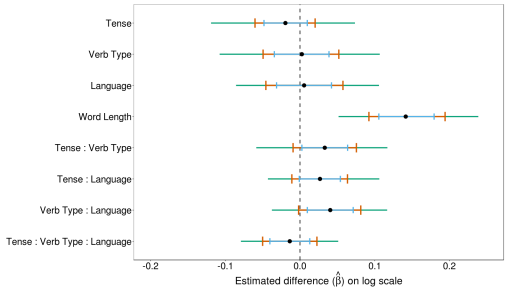


Figure E.14: **First pass reading times** in the **pre-critical region** (second noun + complementizer) – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

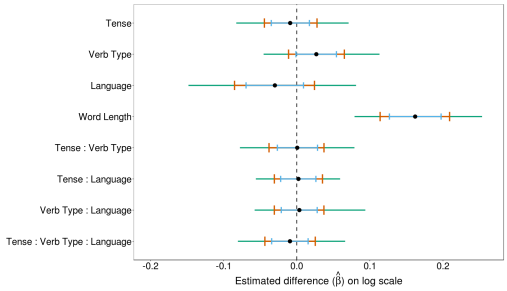


Figure E.15: **Total reading times** in the **pre-critical region** (second noun + complementizer) – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

	First pass reading times			Total reading times		
Effect	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**
Tense	0.009	[-0.02, 0.04]	0.33	-0.03	[-0.05, 0.004]	0.89
Verb Type	0.02	[-0.02, 0.05]	0.23	0.02	[-0.009, 0.05]	0.16
Language	-0.05	[-0.1, -0.02]	0.98	-0.06	[-0.1, -0.02]	0.99
Word Length	0.18	[0.14, 0.21]	0	0.15	[0.11, 0.18]	0
Tense : Verb Type	0.01	[-0.02, 0.05]	0.29	0.02	[-0.01, 0.05]	0.18
Tense : Language	-0.01	[-0.04, 0.02]	0.72	0.04	[0.01, 0.06]	0.02
Verb Type : Language	-0.02	[-0.05, 0.007]	0.86	-0.02	[-0.05, 0.005]	0.88
Tense : Verb Type : Language	-0.003	[-0.03, 0.02]	0.56	-0.008	[-0.04, 0.02]	0.68

Table E.9: **First pass reading times** and **total reading times** in the **post-critical end-of-sentence region** – summary of the posteriors of the fixed-effects model parameters. For each effect we provide the estimated mean of the posterior, the 85% credible intervals and the probability that the posterior is smaller than zero. For both dependent variables, the estimated mean and the 85% credible intervals are on log scale.

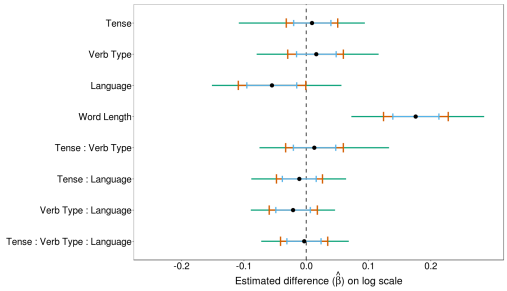


Figure E.16: **First pass reading times** in the **post-critical end-of-sentence region** – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

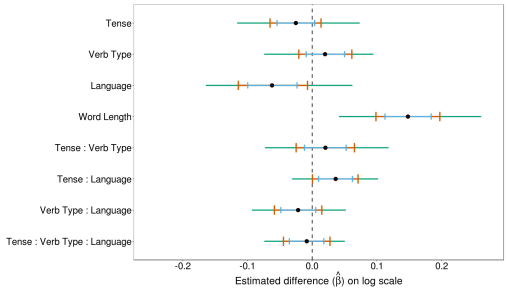


Figure E.17: **Total reading times** in the **post-critical end-of-sentence region** – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

	Regression path duration			Proportion of regressions-out		
Effect	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**
Tense	-0.03	[-0.07, 0.01]	0.87	-0.03	[-0.21, 0.15]	0.61
Verb Type	0.01	[-0.02, 0.05]	0.25	-0.19	[-0.36, -0.03]	0.95
Language	-0.1	[-0.14, -0.06]	1	-0.03	[-0.17, 0.11]	0.60
Word Length	0.07	[0.03, 0.11]	0.009	-0.006	[-0.19, 0.17]	0.51
Tense : Verb Type	0.02	[-0.01, 0.05]	0.22	-0.02	[-0.18, 0.14]	0.58
Tense : Language	0.04	[0.01, 0.06]	0.03	0.06	[-0.07, 0.18]	0.27
Verb Type : Language	-0.03	[-0.05, 0.001]	0.91	-0.22	[-0.35, -0.09]	0.99
Tense : Verb Type : Language	-0.02	[-0.04, 0.01]	0.79	-0.04	[-0.16, 0.09]	0.66

Table E.10: **Regression path duration** and **proportion of regressions-out** in the **pre-critical region** comprising the second noun and the complementizer – summary of the posteriors of the fixed-effects model parameters. For each effect we provide the estimated mean of the posterior, the 85% credible intervals and the probability that the posterior is smaller than zero. For the regression path duration, the estimated mean and the 85% credible intervals are on log scale; for the proportion of regressions-out they are on logit scale.

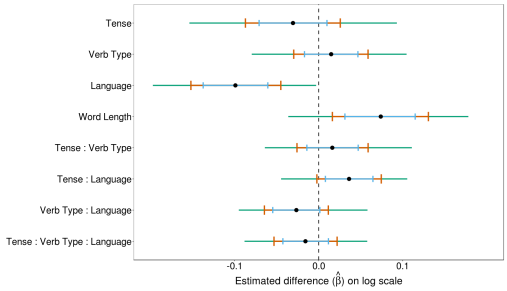


Figure E.18: **Regression path duration** in the **pre-critical region** (second noun + complementizer) – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

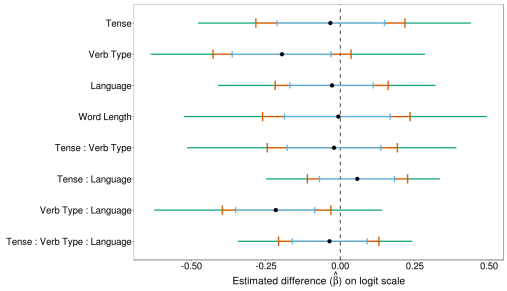


Figure E.19: **Proportion of regressions-out** of the **pre-critical region** (second noun + complementizer) – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

	Regression path duration			Proportion of regressions-out		
Effect	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**	**Estimated mean ($\hat{\beta}$)**	**85% credible intervals**	**P($\hat{\beta}$) < 0**
Tense	-0.005	[-0.04, 0.03]	0.59	0.25	[-0.02, 0.54]	0.09
Verb Type	0.02	[-0.006, 0.05]	0.12	0.40	[0.12, 0.69]	0.03
Language	-0.01	[-0.05, 0.03]	0.67	0.25	[0.002, 0.50]	0.07
Word Length	0.06	[0.03, 0.09]	0.005	-0.17	[-0.43, 0.08]	0.84
Tense : Verb Type	0.01	[-0.01, 0.04]	0.24	0.17	[-0.15, 0.49]	0.23
Tense : Language	0.04	[0.02, 0.07]	0.01	0.13	[-0.08, 0.35]	0.18
Verb Type : Language	-0.01	[-0.04, 0.02]	0.71	0.07	[-0.14, 0.3]	0.32
Tense : Verb Type : Language	0.01	[-0.01, 0.04]	0.22	-0.01	[-0.23, 0.20]	0.55

Table E.11: **Regression path duration** and **proportion of regressions-out** in the **post-critical end-of-sentence region** – summary of the posteriors of the fixed-effects model parameters. For each effect we provide the estimated mean of the posterior, the 85% credible intervals and the probability that the posterior is smaller than zero. For the regression path duration, the estimated mean and the 85% credible intervals are on log scale; for the proportion of regressions-out they are on logit scale.

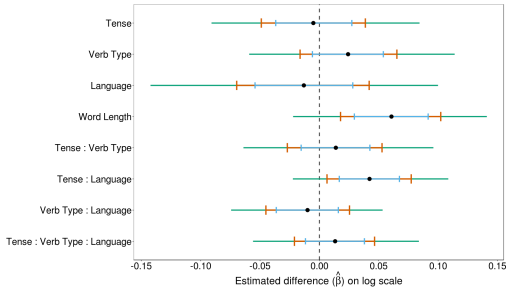


Figure E.20: **Regression path duration** in the **post-critical end-of-sentence region** – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

1274

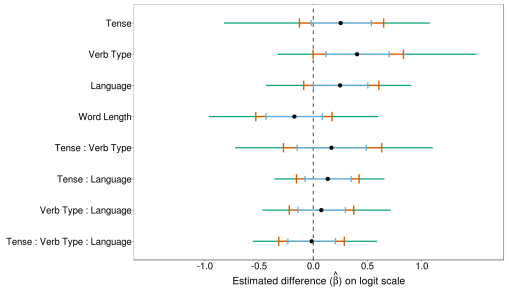


Figure E.21: **Proportion of regressions-out** of the **post-critical end-of-sentence region** – posterior distributions of the fixed-effects model parameters. The black dot marks the posterior's mean. The outer bars show the 95% credible intervals; the inner bars show the 85% credible intervals. Zero is marked with a vertical dashed line.

1275