



Contents lists available at ScienceDirect

Journal of Pragmatics

journal homepage: [www.elsevier.com/locate/pragma](http://www.elsevier.com/locate/pragma)

# Pragmatic and syntactic constraints on French causal connectives: An evaluation of native and non-native speakers' sensitivity



Mathis Wetzel <sup>a, \*</sup>, Ekaterina Tskhovrebova <sup>a</sup>, Pascal M. Gygax <sup>b</sup>,  
Sandrine Zufferey <sup>a</sup>

<sup>a</sup> Institut de langue et de littératures françaises, University of Bern, Länggassstrasse 49, CH-3012 Bern, Switzerland

<sup>b</sup> Department of Psychology, University of Fribourg, Rue P.-A.-de-Faucigny 2, CH-1700 Fribourg, Switzerland

## ARTICLE INFO

### Article history:

Received 5 September 2022

Received in revised form 8 February 2023

Accepted 2 March 2023

### Keywords:

Discourse connectives

French as a second language

Causality

Syntactic constraints

Pragmatic constraints

Individual differences

## ABSTRACT

When writing a text, the choice between seemingly equivalent connectives, such as *but* and *however*, is not arbitrary: aside from preferences linked to the register or modality of the utterance, there are often pragmatic or syntactic constraints that limit the choice to only one particular connective. For French, theoretical approaches have extensively discussed the syntactic and pragmatic constraints that allow only either *parce que* ('because') or *puisque* ('since'). However, experimental findings regarding the actual sensitivity of French speakers to these constraints remain inconclusive. In the current study, we examine the sensitivity of different speaker groups to the constraints associated with the use of *parce que* and *puisque*. In a controlled offline-task, we systematically evaluate the sensitivity of a crowd-sourced heterogeneous group of native speakers (Experiment 1) and a group of non-native speakers (Experiment 2). Results show that native speakers are more sensitive to syntactic constraints whereas pragmatic constraints are more equivocal. Non-native speakers, however, are not sensitive to either of these constraints. In addition, the pragmatically complex connective *puisque* is extremely problematic for non-native speakers, as even the mapping of a seemingly equivalent connective of their L1 did not warrant a good mastery of its pragmatic functions.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

When producing and structuring discourse, speakers can use connectives (words such as *however* or *therefore*) to link logical ideas between discourse segments (e.g., Halliday and Hasan, 1976). While doing so, they constantly need to make decisions about which particular connective is the most appropriate choice to convey the intended coherence relation. For instance, in (1), the speaker could choose between connectives such as *but*, *yet* and *however* in order to indicate the concessive relation.

- 1) I ran. \_\_\_ I did not catch the train.

\* Corresponding author.

E-mail address: [Mathis.wetzel@unibe.ch](mailto:Mathis.wetzel@unibe.ch) (M. Wetzel).

Yet, connectives conveying the same coherence relation are not always fully interchangeable (e.g., Knott, 1996; Yung et al., 2021) and the choice of a specific connective is, therefore, often not arbitrary. The preference for a specific connective can depend, for instance, on the formality of the communicative situation or whether the utterance is expressed in written or spoken language (Crible and Cuenca, 2017). The choice of an appropriate connective also involves taking into account syntactic and pragmatic factors (Fraser, 1998). For example, in (2), the choice to stay inside is presented as a consequence of rainy weather. To indicate this idea explicitly, a speaker could use the consequence connective *therefore*, whereas the other connective of consequence *so* would be incorrect in clause-medial position, due to syntactic constraints.

2) It rained all day, I stayed, *therefore* / \**so*, inside.

Similarly, the use of *however* in (3) is not possible, due to the pragmatic constraints that call for a concessive connective implying protest (see also Fraser, 1998), a constraint fulfilled only by *but* in English.

3) A: I thought you would come.

B: *But* / \* *However*, I did!

Another example of such fine-grained constraints can be found in the causal domain. In French, previous work has extensively discussed in which situations the three causal connectives *car* ('for'), *puisque* ('since') and *parce que* ('because') can be used, by establishing specific profiles for each of them (e.g., Groupe  $\lambda$ -I, 1975; Zufferey, 2012; Schumann et al., 2020, ). Yet, experimental evidence documenting the sensitivity of actual French speakers to the theoretically established constraints of connectives is somewhat sparse. While it is likely that speakers adhere to these constraints in their everyday language practices, it is uncertain to what extent deviant uses are accepted or if these constraints are completely binding. In this regard, the question also arises as to whether some constraints (syntactic vs. pragmatic) are stronger than others.

The strength of these constraints may also depend on individual differences between speakers. While being sensitive to the specificities of each connective is essential for successful communication, it could be the case that less skilled speakers with a somewhat lower sensitivity to the constraints associated with some connectives run the risk of misunderstandings or misinterpretations. This may especially be the case for non-native speakers, for whom pragmatic structures are difficult to acquire (e.g., Bardovi-Harlig and Dörnyei, 1998).

Generally, the use of discourse markers in a second language is well documented (e.g., Müller, 2005; Buysse, 2012, 2014; Huang et al., 2022). In regard to discourse connectives, there are indications that the choice of connectives conveying the same coherence relation is less clear, less intuitive and less informed than those of native speakers. Many corpus studies have observed, for instance, that non-native writers tend to use correct, but less formal connectives with an inappropriate register (e.g., Lee, 2013; Myung-Jeong, 2016; Altenberg and Tapper, 1998) and only rely on the same small set of connectives that they are comfortable with (e.g., Leedham and Cai, 2013; Chen, 2014; Castele and Collewaert, 2013). These observations could be an indication that non-native speakers may not have acquired all the subtle pragmatic characteristics of certain specific connectives and tend to avoid them - at the cost of using an inappropriate connective to convey the same coherence relation instead.

We would argue that more research is needed to investigate in detail the sensitivity of a wide range of French speakers to the syntactic and pragmatic constraints imposed by connectives. Furthermore, it should also take into account non-native speakers, as a group with possibly even lower pragmatic competence. In this paper, we attempt to fill this gap by providing new data in order to determine the extent to which French speakers are sensitive to the constraints allowing either *puisque* ('since') or *parce que* ('because') to be used, by testing two types of populations: first, a heterogeneous group of native speakers, and second, a group of non-native speakers of French. By doing so, our research not only aims to provide experimental data to inform important theoretical issues on connectives (such as the nature of the procedural constraints they encode), but also to document the existence of individual differences between both native and non-native speakers in the mastery of connectives.

## 2. Discourse connectives

The term *discourse connectives* refers to a category of linguistic elements stemming from different grammatical categories. These elements, at the crossroad of pragmatics and semantics, are used to indicate the underlying coherence relations (i.e., the logical links between different propositions) of a text (e.g., Halliday and Hasan, 1976). As illustrated in (4), the connective *but* conveys amongst other relations a concessive relation, as it cancels a salient inference that could be drawn from S1 (see also Bell, 2007): people hurry when they are late.

4) [Peter was late]<sub>S1</sub> *but* [he did not hurry]<sub>S2</sub>.

Most coherence relations can be conveyed by more than one discourse connective, at least in Indo-European languages that possess a vast repertoire of these items. For instance, in (4) the concessive connective *yet* would also represent an appropriate choice. Yet, connectives vary greatly in their specific characteristics (see for instance Bell, 2010, for English concessive connectives), and research has shown that subtle differences between connectives conveying the same coherence relation do have an impact on the way readers process and retrieve information (Schumann et al., 2020; Wetzel et al., 2022b).

One difference among connectives conveying the same coherence relation lies, for instance, in pragmatic uses (e.g., Sweetser, 1990; Fraser, 1998). An example of such a pragmatic use is the difference between *objective* and *subjective* causality.

Subjective causality is established when the point of view or reasoning of the speaker is involved (Degand and Pander Maat, 2003; Sanders et al., 1992; Moeschler, 2005). Consider (5) in which the fact that Marie is not heard by the speaker, led them to conclude that Marie had left. In contrast, the causality in (6) is more objective, as it states the fact that the speaker cannot hear Marie anymore as a logical consequence of her leaving.

- 5) Marie left, *since* I do not hear her anymore.  
 6) I don't hear Marie anymore, *because* she left.

As already indicated in the examples, the difference between objective and subjective causality has important implications for the choice of connective: a large body of research has shown that certain causal connectives, such as *because* in English, are more likely to be used when the causality is objective whereas others, such as *since*, are used to indicate a subjective type of causality (e.g., Sweetser, 1990). Although perhaps not always consciously, speakers have been found to use this criterion while choosing a causal connective in several languages, such as Portuguese (Lopes, 2009), Dutch (Pit, 2007), German (Stukker and Sanders, 2012) or Mandarin Chinese (Li et al., 2013). While in some cases the difference between objective and subjective causality gives an answer to the question of why a seemingly equivalent connective is not appropriate, there are also other types of constraints that influence the decision between connectives, such as syntactic and pragmatic characteristics of language.

### 3. Syntactic and pragmatic constraints for connectives

Different connectives, although conveying the same coherence relation, cannot always be used interchangeably due to different types of constraints that either allow only one connective or exclude others (e.g., Hitzeman, 1995; Zufferey, 2012; Knott, 1996). Constraints can be of a syntactic nature (e.g., Fraser, 1998). Pitler and Nenkova (2009) were able to show, for instance, that the disambiguation of polyfunctional connectives (such as *since*) was facilitated when considering the syntactic position of the connectives. Yet, perceived syntactic constraints for connectives are not always binding or shared by speakers, as discussed by Bell (2007). While grammarians have long proscribed the use of English connectives *and* and *but* in the sentence initial position, these uses become increasingly frequent in corpus data. Still, commonly shared syntactic constraints of connectives do exist, as demonstrated by Fraser (1998). Consider (7) (taken from Fraser, 1998: 306) in which the sentence initial position excludes the use of *on the other hand*, the medial position the uses of *but*, *rather* and *though*, and the sentence final position the use of *on the other hand* or *on the contrary* (Fraser, 1998: 306).

- 7) It may rain. \_\_\_\_\_ (initial position) it \_\_\_\_\_ (medial position) may not \_\_\_\_\_ (final position).

In addition to syntactic constraints, connectives also convey subtle pragmatic differences. For example, the English connective *but* can be used to indicate a “direct, an implied, a presupposed, or an entailed message” (Fraser, 1998: 312) that depends highly on the context. As such, it cannot be arbitrarily interchanged with *however*, as this would create either controversial (8) or simply wrong uses (9). Both examples are taken from Fraser (1998: 313).

- 8) A: Here is a triangle.  
 B: <sup>?</sup>*However* / *But* this has four sides.  
 9) A: Let's go.  
 B: \**However* / *But*, I can't go.

Taken together, although they convey the same coherence relations, connectives differ in many ways and can often not be used, due to pragmatic and syntactic constraints, interchangeably.

We will now present the types of constraints separating the two French connectives *puisque* ('since') and *parce que* ('because') that both indicate a relation of *consequence-cause* but that cannot always be used interchangeably.

### 4. When to choose *puisque* and when to choose *parce que*? A linguistic description

French causal connectives such as *parce que* ('because'), *car* ('for') and *puisque* ('since') have been extensively studied. Prototypically, *parce que* ('because') is described as being used mainly in the content domain and to convey objective causality (Degand and Pander Maat, 2003; Degand and Fagard, 2018; Fagard and Degand, 2008; Lambrecht et al., 2006). This means that it is used to indicate causality between segments that can be easily and logically established, such as in (10), and which is based neither on epistemic interpretation (11) nor on speech acts (12). All examples taken from Sweetser (1990: 77).

- 10) John came back because he loved her.  
 11) John loved her because he came back.  
 12) What are you doing tonight because there's a good movie on.

As such, *parce que* is often opposed to the more subjective *car* ('for', e.g., Bentolila, 1986; Groupe λ-1, 1975, Degand and Fagard, 2018; Simon and Degand, 2007). Zufferey et al. (2018) showed, however, that this distinction between these connectives is blurry in contemporary French. Nowadays, the difference between the connectives is rather perceived in terms of registers: while *car* is considered to be more bound to the written mode, *parce que* is used to express all kinds of causal

relations and even replaces *car* to a great extent in spoken data (see also: Fagard and Degand, 2008; Simon and Degand, 2007; Zufferey, 2012; Degand and Fagard, 2018).

In contrast to *parce que*, several studies (e.g., Degand, 2004, Zufferey and Cartoni, 2012; Scholman et al., 2020) classify *puisque* ('since') as a connective that, amongst other criteria, has a high degree of subjectivity. As such, it can convey a dissociative attitude (Schumann et al., 2020; see also Sperber and Wilson, 1986) by indicating, for instance, annoyance.

Crucially for this paper, *puisque* differs from other French causal connectives in another dimension: its ability to indicate the givenness of a situation (e.g., Ariel, 1985; Zufferey, 2012, 2014; Zufferey and Cartoni, 2012; Schumann et al., 2020). This means that *puisque* is used to indicate already known or obvious information and thus "takes into account the other" speaker (Leeman, 1994: 121, our translation), as in (13) (taken from Zufferey, 2014: 124).

13) *On ne peut pas observer les étoiles à midi, puisqu'il fait jour.*

'You cannot see the stars at midday, CONNECTIVE there is daylight.'

Here, the knowledge that daylight makes it impossible to see the stars is commonly shared (and indicated as such) by both the locutor and the interlocutor.

In a first, more explorative empirical study, Zufferey (2012) assessed the acceptance by highly competent French native speakers for the uses of *car* ('for'), *puisque* ('since') and *parce que* ('because') in different sentences. Participants were first asked to evaluate sentences containing these connectives and then, in a second task, to fill blanks in the same sentences with the three connectives. Importantly, the sentences differed across the three domains defined by Sweetser (i.e., content, epistemic and speech act domain, 1990), allowing thus a first evaluation of the constraints that limit the choice to a particular connective.

One finding was that participants not only showed 100% agreement that *parce que* is correct in the sentence (14) (taken from Zufferey, 2012), but also exclusively chose *parce que* in the fill-in-the-blank task.

14) *C'est \_\_\_\_ il a trop mangé qu'il est malade.*

~ It is \_\_\_\_ he ate too much, that he is sick.

This indicates that only *parce que* can be used in cleft sentences with *c'est* ('it is') whereas this appears not to be compatible with other causal connectives, such as *puisque*.<sup>1</sup> Another constraint only allowing *parce que* (and excluding *puisque*) is that *parce que* can be used to answer a why question – contrary to *puisque* (Groupe λ-1, 1975). Zufferey (2012) observed for sentence (15) that participants preferred at almost 90% *parce que* over *car*, whereas no one indicated *puisque*.

15) *Jean: Pourquoi est-il parti? Pierre: \_\_\_\_ il était fatigué.*

'Jean: Why did he leave? Pierre: \_\_\_\_ he was tired.'

Yet, there are also cases in which *puisque* was highly preferred and *parce que* less accepted, especially, as already discussed, when introducing given information. An example of this constraint is apparent when the sentence contains a citation or paraphrasing of what the interlocutor had said. As such, information is necessarily already known and commonly shared by both interlocutors. In these cases, it appears that native speakers prefer *puisque*, as participants of Zufferey (2012) chose it in 90% of cases for the sentence (16).

16) *Viens te promener! \_\_\_\_ tu dis que tu aimes la montagne.*

'Come for a walk! \_\_\_\_ you say you like the mountains.'

Another type of givenness is present when the situation is so obvious that it is necessarily known by the interlocutor (Groupe λ-1, 1975; Zufferey, 2014). Consider (17): as the interlocutor is necessarily aware that they are present, the information in the second sentence can be said to contain mutually shared information.

17) *Donne-moi un coup de main, puisque tu es là.*

'Give me a hand, CONNECTIVE you are here.'

Yet, findings for this type of constraint in (Zufferey, 2012) are unclear. Although many have shown that *puisque* is indeed used to indicate this type of givenness (Groupe λ-1, 1975; Zufferey, 2014), participants in Zufferey (2012) did not show a clear preference for *puisque* for a sentence that involved an obvious statement.

While the preliminary work of Zufferey (2012) indicates that there are various types of constraints for the French connectives *puisque* and *parce que* and that highly proficient speakers seem to be aware of them, it also raises numerous questions and concerns.

First of all, it should be noted that the results obtained cannot be generalized, and this is for several reasons. For example, Zufferey (2012) only tested speakers that had a high sensitivity to French - namely university students studying French language and literature. Yet, it is unclear whether the observed preferences expressed by these extremely proficient speakers also apply to all French speakers, such as speakers with a lower proficiency and less clear-cut intuitions about the use of connectives. Indeed, there are indications that competence with connectives is quite variable, even among native speakers, and that some of them are less sensitive to the constraints imposed by certain connectives (Tskhovrebova et al., 2022). Another limitation of this study is that only one sentence was used to test each constraint, which makes the interpretation of

<sup>1</sup> Note that the phrase *c'est parce que* ('that is because') is not classified as a different connective (see LexConn, Roze, Danlos & Muller, 2012).

results somewhat difficult. After all, it is possible that the same constraint is more or less pervasive in one sentence than in another. As generalizability can only be achieved if testing is systematic and controlled, the study thus failed to show whether some of the constraints are indeed more binding than others, respectively whether different types of constraints are identified and adhered to differently.

In addition, new research is needed that not only ensures generalizability of the findings of Zufferey (2012, 2014), but also to examine in more detail if and to what extent speakers adhere to the pragmatic and syntactic constraints of these connectives. For instance, it is well-known that native speakers differ in their mastery of connectives (e.g., Wetzel et al., 2020; Tskhovrebova et al., 2022). Yet, we do not know which individual factors may predict a better grasp of syntactic and pragmatic constraints of connectives.

Another question arises regarding non-native speakers. If an L2 speaker fails to understand the pragmatic undertones of *puisque*, the conversation will be severely disrupted and it may even lead to a complete breakdown in communication. Yet, it appears that, despite *puisque* being quite frequent in spoken language, non-native speakers do not master it well. In a multiple-choice cloze task, non-native speakers in Wetzel et al. (2020) showed an unsatisfactory mastery of this connective. Unfortunately, these results do not allow for any conclusion for the extent to which especially pragmatic constraints (and if so, which ones) hindered mastery, as the authors did not test any of the pragmatic constraints of this connective.

Taken together, the work of Zufferey (2012, 2014) hints at the presence of binding constraints of the French connectives *puisque* and *parce que*. However, the shortfalls of this and other studies mentioned above call for a more controlled approach that ensures generalizability of the findings and investigates in a more detailed manner the sensitivity of a more diversified range of speakers to the connectives' constraints.

In the current study, we address these issues by conducting experiments with sentences containing syntactic and pragmatic constraints rendering *parce que* the appropriate connective, and sentences containing the two types of pragmatic constraints rendering only the use of *puisque* appropriate.

## 5. Experiment 1. crowd-sourced group of native French speakers

In this experiment, we assess to what extent native French speakers are sensitive to the different constraints in causal sentences that limit the choice between the two French connectives *parce que* ('because') and *puisque* ('since'). While it is true that syntactic and pragmatic constraints can sometimes overlap, the constraints tested in our experiment were of only one type and were determined by acceptance judgment data from native speakers.

Firstly, we compare constraints that are syntactic or pragmatic (for *parce que*) and secondly, we compare two types of pragmatic constraints (for *puisque*). For the first, we expect that syntactic constraints might be somewhat more salient for speakers. For sentences in which *puisque* is the only possible choice, pragmatic constraints involving citations might be somewhat more salient than statements referring to the obviousness of the situations, as citations more clearly emphasize the fact that the given information is, in fact, already known.

### 5.1. Participants

In order to obtain a broad and heterogeneous group of participants (i.e., with different socio-educative backgrounds), we opted for a crowd-sourcing approach. Crowd-sourcing has proven to be an advantageous approach for studying discourse interpretations and the role of connectives (see Santana et al., 2021 for a discussion). We therefore recruited sixty-three French native speakers (23 women, 1 non-binary, mean age = 32 years old, SD = 9.2, range = 18 to 61) via the online platform Prolific. Participants were remunerated in this experiment with 3.15 GBP. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the ethics protocol was approved by the Swiss National Science Foundation (100012 184882).

### 5.2. Materials and procedure

The experiment was conducted and designed using Qualtrics software (Qualtrics LLC, Provo, USA). After accepting the consent form, all participants performed the main task of the experiment, a multiple-choice cloze task. In this task, sentences were presented in which a connective was omitted and replaced by a blank. Participants had to choose the correct connective among four options for each of the sentences. The sentences were presented without further context and in a randomized order. The four options were *puisque* ('since'), *parce que* ('because'), *alors* ('so') and *mais* ('but') and were also presented in a randomized order.

We tested 20 sentences targeting *parce que* and 20 sentences targeting *puisque*. Among the 20 sentences targeting *parce que*, 10 included syntactic constraints and 10 included pragmatic constraints, with *parce que* being the only appropriate choice for both. As illustrated in (18), the syntactic constraint was elicited by using a cleft structure, e.g., with *c'est* ('that is'), which is incompatible with the other connectives from the experiment (i.e., *puisque*, *alors* or *mais*, Zufferey, 2012). For the pragmatic constraint (19) we used a question. The rationale here is that when a speaker asks something, they necessarily try to obtain

new information, thus ruling *puisque* out as an appropriate connective.<sup>2</sup> Also, the explanation given referred to the real-world domain, excluding *puisque* (Zufferey, 2014).

- 18) *Si elle a compris ça, c'est \_\_\_\_\_ elle est très intelligente.*  
'If she understood that, it's \_\_\_\_\_ she's very smart.'
- 19) *Où est Jean ? - Il ne peut pas venir \_\_\_\_\_ il est malade.*  
'Where is Jean? - He can't come \_\_\_\_\_ he's sick.'

Among the 20 sentences targeting *puisque*, 10 contained a constraint due to the obviousness of the situation (20) and 10 contained sentences in which *puisque* introduces information that has already been given by the interlocutor, and is cited or rephrased by the speaker (21). As such, these sentences represent prototypical uses of *puisque*, and render other choices inappropriate.

- 20) *Je profite de te le dire maintenant, \_\_\_\_\_ tu es encore là.*  
'I take the opportunity to tell you now, \_\_\_\_\_ you are still here.'
- 21) *Je savais que tu allais réparer la table, \_\_\_\_\_ tu me l'avais annoncé hier.*  
'I knew you were going to fix the table \_\_\_\_\_ you told me yesterday.'

In order to ensure that the options *alors* ('so') and *mais* ('but') would be correct from time to time and that not only causes were tested, we also tested 10 correct uses of *alors* in cause-consequence relations, and 10 correct uses of *mais* in concessive relations as filler sentences. We chose *alors* ('so') and *mais* ('but') to that end, as these connectives convey coherence relations that differ in their proximity to the tested consequence-cause relations (for *puisque* and *parce que*). As displayed in (22), *parce que* indicates *non-basic* or *backward* causality (i.e., S2 being the cause of S1; following Sanders et al., 1992), whereas the connective *alors* conveys *basic* or *forward* causality (i.e., S2 being the consequence of S1, Example 23). *Mais* ('but') on the other hand conveys a coherence relation that clearly differs to that of *puisque* and *parce que*, namely a concession, as seen in (24) in which the assumption created in S1 (i.e., that people hurry when they are late) is negated.

- 22) *[Je me dépêche]<sub>S1</sub> parce que [je suis en retard]<sub>S2</sub>.*  
'[I hurry]<sub>S1</sub> because [I am late]<sub>S2</sub>'
- 23) *[Je suis en retard]<sub>S1</sub> alors [je me dépêche]<sub>S2</sub>.*  
'[I am late] so [I hurry]<sub>S2</sub>'
- 24) *[Je suis en retard]<sub>S1</sub> mais [je ne me dépêche pas]<sub>S1</sub>.*  
'[I am late]<sub>S1</sub> but [I do not hurry]<sub>S1</sub>'

Hence, for a sentence that targeted the use of *puisque*, the four options were: 1) a connective that conveyed the correct coherence relation and that was pragmatically appropriate (i.e., *puisque*), 2) a connective indicating the correct coherence relation that was, however, inappropriate due to pragmatic constraints (i.e., *parce que*), 3) a connective that also conveyed causality, however with a wrong order of the segments (i.e., *alors*), and 4) a connective that conveys a different coherence relation (i.e., *mais*).

Research has shown that speakers differ greatly in their ability to read discourse and to master discourse connectives appropriately (e.g., Braze et al., 2007; Wetzel et al., 2020). Therefore, we also conducted two supplementary tasks targeting different dimensions of the language proficiency of our participants.<sup>3</sup>

Firstly, we assessed their sensitivity to ungrammatical uses of language, by using a French Grammar task (Zufferey and Gygax, 2020). In this task, participants rated the correctness of 40 sentences by situating a cursor on a scale reaching from 'I am sure this is incorrect' to 'I am sure this is correct'. Twenty of the sentences contained an error typical of the written mode (such as the incorrect spelling of *différend* ['disagreement'] as its homophone *différent* ['different']). The position of the cursor was converted into a numerical value on a scale from 0 to 100. A total percentage was then calculated for the judgement of all 40 sentences, in which 100% indicates that the person judged all sentences appropriately (i.e., correct sentences as correct, and incorrect sentences as incorrect).

Secondly, we assessed the vocabulary knowledge of our participants by using the French version of the Lextale task (Brysbart, 2013), which has been proven to solidly detect differences in the vocabulary width of both native and non-native speakers (Wetzel et al., 2020). In this task, participants chose the words they recognized as existing ones among 56 existing and 28 non-existing (but pseudo-words) French words. They were advised that choosing a non-existing word would lead to a point deduction. Thus, the maximum score was 56. Once more, a total percentage of words recognized was calculated.

<sup>2</sup> In order to avoid that only sentences targeting pragmatic uses of *parce que* contained a question, we also included five (unrelated) questions in all other conditions. Note that these questions did not alter the sentences nor the choice of the connective.

<sup>3</sup> Importantly, these tasks assess linguistic competences other than pragmatic ones, as their first goal was to control for the language proficiency of our participants. Still, the results enabled us to assess whether a higher language proficiency, such as grammar knowledge, would also predict a greater sensitivity to constraints of connectives.

### 5.3. Analyses and results

All of our data and R - scripts can be retrieved at [https://osf.io/hgfsu/?view\\_only=b3bf160f9854469d8d8f3e602ca04356](https://osf.io/hgfsu/?view_only=b3bf160f9854469d8d8f3e602ca04356).

#### 5.3.1. Descriptive results

In a first step, we assessed whether participants were able to identify the correct coherence relation in sentences targeting *puisque* or *parce que*, thus controlling that they performed well in the task and that the tested sentences unambiguously implied a consequence-cause relation. It appeared that they did, as participants chose in 98% of the cases (95% CI [96.97, 99.03]) either one of the connectives for the sentences of this coherence relation.

In a second step, we assessed whether participants, while choosing between the two connectives to indicate this coherence relation, scored differently for the two connectives, i.e., whether the decision was easier for either sentences targeting *parce que* or *puisque* (independently of the constraints tested). The results are reported in Table 2.

**Table 2**  
Experiment 1. Descriptive results for experimental items.

	mean	95% CI
Sentences targeting <i>parce que</i>	91%	[89.71, 92.83]
Sentences targeting <i>puisque</i>	76%	[73.75, 78.47]

As can be seen, sentences targeting *puisque* elicited considerably lower scores, meaning that participants were more accurate in choosing the appropriate connective for sentences targeting *parce que*.

#### 5.3.2. Analysis

As our dependent variable was binary (i.e., correct or incorrect response for each sentence), we conducted logistic mixed-effects-models that were built successively, meaning that upon adding a fixed effect, the improvement of the resulting model was compared to the model that did not contain this fixed effect (following Baayen, 2008). We used the *glmer* function of the *lme4*-package (Bates et al., 2014) in R (R Core Team, 2020) to build our models and used the *anova()* function of the *stats* - package (R Core Team, 2020) to assess the improvement of the model when fixed effects were added, i.e., to perform the log-likelihood test. We used the *summary()* function of the *base* - package (R Core Team, 2020) to retrieve the estimates and *p* values of our final models. Post hoc tests were conducted using the *emmeans()* function of the *emmeans* - package (Lenth, 2016). All assumptions - such as the absence of multicollinearity (Schreiber-Gregory, 2018) - for logistic mixed-effects models were met. We always opted for a maximal random effect structure, as advocated by Barr et al. (2013).

#### 5.3.3. Measurements of language proficiency

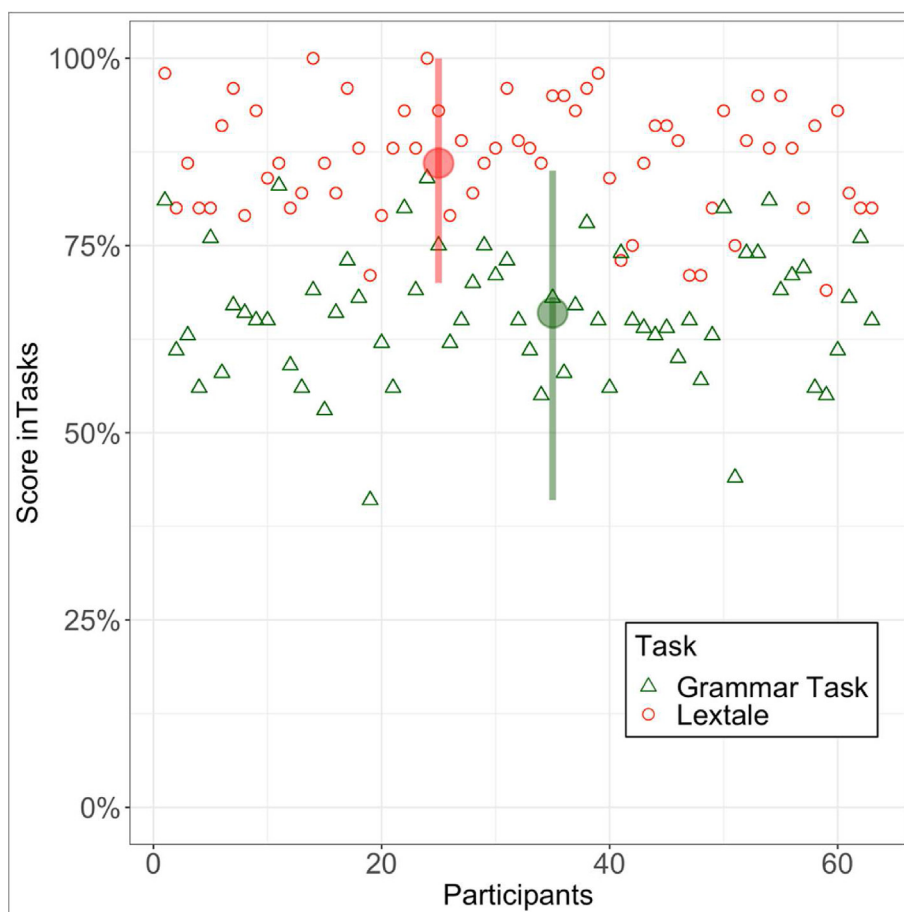
Research has shown that native speakers do not always constitute a homogenous group and may strongly differ in their ability to use connectives (Wetzel et al., 2020). Therefore, we tested whether a greater vocabulary width and a better score in the grammar task would predict a better ability to discriminate between *parce que* and *puisque*. First of all, we indeed observed a high variance in both tasks among the participants, despite the fact that all of them were native speakers. For the Lextale task, scores ranged from 69% to 100% of correctly identified words (mean = 48 [86%]; 95% CI [48.19, 48.47]). For the Grammar task scores, the range was even more pronounced: from 41% to 85% of certainty for a correct response in this task (mean = 66%; 95% CI [65.78, 66.34]). The distribution among the participants of both scores is displayed in Fig. 3.

We contrast coded the predictor variable *Connective* (*puisque* set to 0.5 and *parce que* to -0.5). Adding the contrast of both *Connectives* as a fixed effect to the model (that included the score in the main task as the dependent variable, *Connective* as a random slope per *Participant*, and both *Participant* and *Item* as random intercepts, improved the model's fit ( $\chi(1) = 13.28$ ,  $p < .0001$ ). Yet, adding its interaction with neither the *Score of Lextale* ( $\chi(2) = 2.82$ ,  $p = .24$ ) nor the *Score of the Grammar task* ( $\chi(2) = .47$ ,  $p = .79$ ) improved the model's fit. Hence, although we observed variance among participants in these tasks, they were not predictive of a better ability to choose between *parce que* and *puisque*.

In a next step, we tested whether these scores depended on the different constraints of the connectives, by analyzing their impact for each connective separately.

#### 5.3.4. Comparing syntactic vs. pragmatic constraints of 'parce que'

In order to assess whether syntactic or pragmatic constraints were predictive for the observed scores of *parce que*, we analyzed the data only for this connective. Adding the contrast (pragmatic set to 0.5, syntactic set to -0.5) of both types of *Constraints* to a random model (containing only *Participants* and *Items* as random effects) did improve the model's fit ( $\chi(1) = 8.39$ ,  $p < .005$ ). The output of the final model is reported in Table 3.



**Fig. 1.** Experiment 1. Distribution of the mean scores of the Grammar and Lextale task per Participants. Ranges of means displayed as lines (including the mean for all participants).

**Table 3**

Experiment 1. Output of final, contrast coded model including the type of Constraint as a fixed effect and Participants and Items as random effects.

	Estimate	SE	z	Pr(> z )
(Intercept)	3.28	.36	9.05	<2 <sup>e-16</sup>
Syntactic constraint vs. pragmatic constraint	1.96	.66	.30	<.005

Hence, French participants scored not only better for *parce que* (i.e., 91% of correct responses) but they also scored better when there was a syntactic constraint (mean of 96% of correct answers for the syntactic constraint vs. a mean of 86% of correct answers for the pragmatic constraint). Results are displayed in Fig. 1.

Given the observed difference between the two types of constraints for *parce que* among native speakers, the question arises as to whether the pragmatic constraint elicited lower scores or the syntactic constraint particularly high scores. A comparison of the scores for the filler sentences provides some indication in that matter. As a reminder, these sentences targeted uses of *alors* ('so') and *mais* ('but'). The mean scores of the filler sentences and the two types of constraints for *parce que* are reported in Table 4.

It appears that, numerically, pragmatic constraints elicited particularly low scores, as sentences targeting *alors* ('so'), *mais* ('but') and syntactic constraints for *parce que* ('because') elicited ceiling level scores.

### 5.3.5. Comparing the two types of pragmatic constraints of 'puisque'

We now focus on the other connective, *puisque* ('since'), and assess whether the scores observed for this connective can be predicted by the type of constraint used. Adding the type of Constraint, the initial model (containing only Participants and Items as random effects) did not improve ( $\chi(1) = .09, p = .77$ ). Hence, as can be seen in Fig. 2, participants did not score better for one of the types of pragmatic constraints (a mean of 76% of corrects answers for both types of constraints).



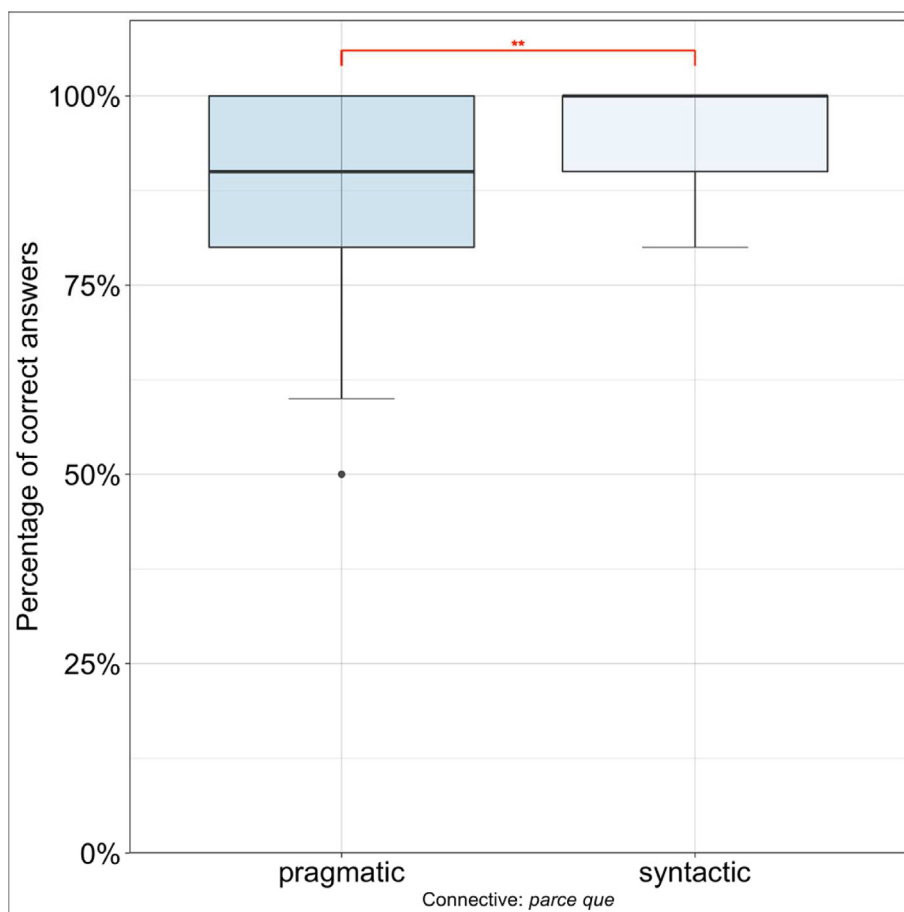


Fig. 2. Experiment 1. Box and whisker plot for the percentage of correct answers (means per participants) for sentences targeting *parce que*.

Table 4

Experiment 1. Mean scores for sentences targeting *alors* ('so'), *mais* ('but') and *parce que* ('because').

	Percentage of correct answers	95% CI
<i>Alors</i>	97%	[95.45, 98.20]
<i>mais</i>	96%	[94.69, 97.69]
<i>parce que</i> (syntactic constraints)	96%	[94.88, 97.82]
<i>parce que</i> (pragmatic constraints)	86%	[83.49, 88.89]

#### 5.4. Discussion

In this study, we tested the sensitivity of a heterogeneous group of French speakers to the pragmatic and syntactic constraints conditioning the use of two causal connectives, namely *puisque* ('since') and *parce que* ('because').

First of all, the finding that participants chose in nearly all the cases either *puisque* or *parce que* for sentences implying a consequence-cause relation shows three things: Firstly, participants performed well on the task (i.e., they did not choose the connectives randomly). Secondly, the tested sentences conveyed an unambiguous coherence relation. Third and finally, participants were able to identify it correctly and, thus, to discriminate between connectives conveying a consequence-cause relation from connectives conveying other relations such as *mais* and *alors*.

We also observed a notable difference in the way participants discriminated between the use of these two connectives, as sentences targeting the pragmatically complex connective *puisque* elicited lower performance scores than the ones targeting *parce que*. One explanation for this difference may stem from the fact that *puisque* is less frequent in corpus data than *parce que* (e.g., Zufferey and Cartoni, 2012). This difference may also be explained by the fact that sentences targeting *puisque* involved pragmatic constraints whereas sentences targeting *parce que* contained syntactic constraints, which are most probably more salient. Indeed, when comparing scores for sentences targeting *parce que*, our results show that participants scored less well for the sentences that contained pragmatic constraints, compared to the sentences containing syntactic

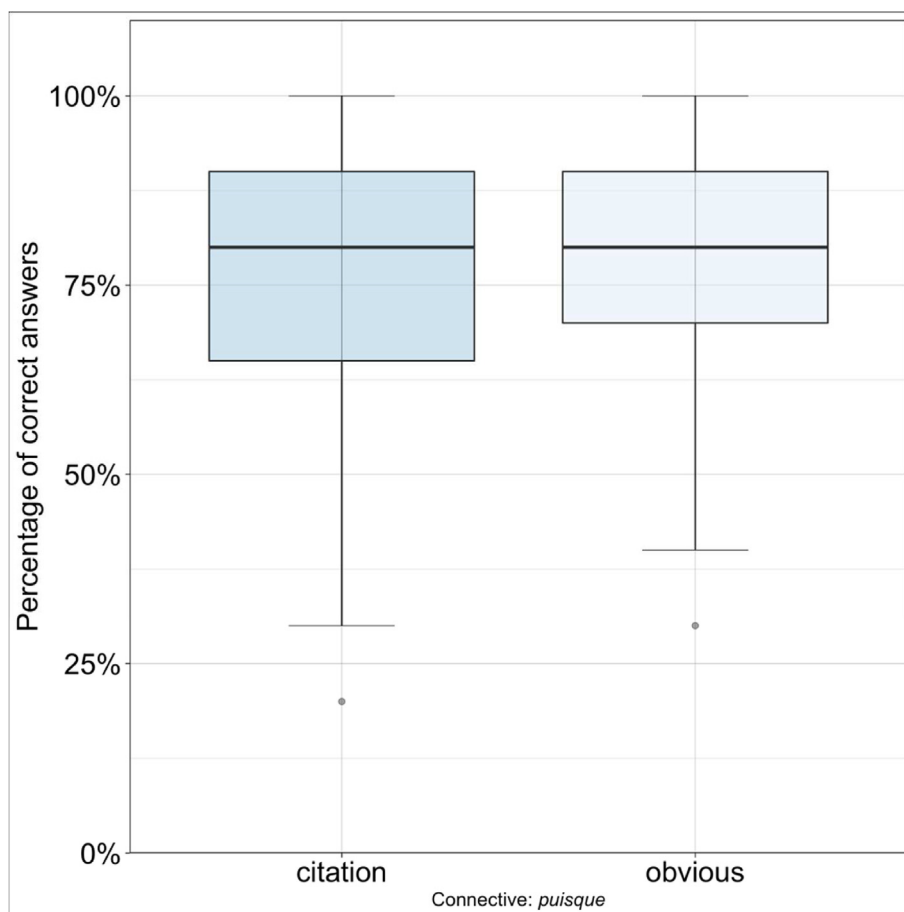


Fig. 3. Experiment 1. Box and whisker plot for the percentage of correct answers (means per participants) for sentences targeting *puisque*.

constraints. When comparing these scores with the scores for the filler sentences targeting *alors* ('so') and *mais* ('but'), we observed that syntactic constraints elicited similar ceiling levels, whereas scores for pragmatic constraints were considerably lower across the board. We conclude, thus, that native speakers have a high sensitivity for syntactic cues, and this is independent of language proficiency. Regarding the two pragmatic constraints of *puisque*, we did not observe clear variations, as most speakers identified the pragmatic constraint just as well under its two forms. Contrary to findings by (Zufferey (2012), *obvious* given statements did not elicit a lower preference for *puisque* than for given statements that contained a *citation*. As already mentioned, this difference most probably stems from the fact that only one sentence was tested in Zufferey (2012), which apparently did not meet the criteria of obviousness sufficiently.

Finally, our study also supports the finding that native speakers are a heterogeneous group regarding their language proficiency (e.g., Kidd et al., 2018), as we found some variance among participants in the *Lextale* and *Grammar task*. Interestingly, however, neither *Lextale* nor *Grammar task* scores were predictive of a better sensitivity of the tested constraints. We believe that it might be that the participants of this study were actually sufficiently proficient with both connectives and picked up generally well on their constraints (however, to a lesser extent for pragmatic uses). In other words, although individual differences among native speakers were apparent, in terms of language proficiency, these differences did not impact the results on the connective task.

This raises the question of how less proficient speakers of French would cope with the task, and whether a greater variance in proficiency would allow us to make predictions regarding the sensitivity to these constraints. In a second experiment, we thus replicated the first experiment with a group of participants that should have less clear-cut intuitions about the syntactic and pragmatic constraints allowing certain connectives, namely non-native speakers.

## 6. Experiment 2: non-native speakers of French

In our second experiment, we replicated the first experiment with non-native speakers of French. We expected non-native speakers to also have a better grasp of the syntactic constraints of connectives as these are more salient, whereas pragmatic

constraints are more subtle and, therefore, more difficult to acquire (see for instance Bardovi-Harlig and Hartford, 1990, 1993; Bardovi-Harlig and Dörnyei, 1998). Regarding the sentences targeting *puisque*, we did not have evidence-based expectations about a potential preference for constraints including a citation over obvious statements (or *vice-versa*).

### 6.1. Participants

Via the Prolific platform we recruited seventy-five native monolingually raised English speakers that were fluent in French (60 women, 2 non-binary, 40.8 years old, SD = 15.8, range = 22 to 84). Participants were remunerated in this experiment with 3.15 GBP. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the ethics protocol was approved by the Swiss National Science Foundation (100012 184882).

### 6.2. Design and procedure

The design and procedure were the same as in Experiment 1 (in Section 5.2).

In addition, we wanted to ensure that the non-native speakers generally understood the meanings of *parce que* ('because') and *puisque* ('since') and were aware of their functions. We did so by asking an open metalinguistic question at the end of the experiment, targeting the difference between *puisque* and *parce que*. (i.e., "What is the difference between *puisque* and *parce que*?"). Participants could either give an unlimited written response (in any language they wished) or indicate *Je ne sais pas* ("I don't know"). When giving an answer, they also indicated how certain they were of their response by indicating a certainty percentage.

### 6.3. Analyses & results

All analyses were conducted as described in Section 5.3.

#### 6.3.1. Descriptive results

As in the first experiment, we assessed whether participants were able to identify the intended coherence relation in sentences targeting *puisque* or *parce que*. Once more, it appeared that participants were able to do so, as they chose for 92% of the sentences either one of the connectives (95% CI [90.18, 94.56]).

Mean scores for both sentences targeting the two connectives (independently of their constraints) are reported in Table 5. As can be seen, sentences targeting *puisque* ('since') elicited extremely low scores.

**Table 5**

Experiment 2. Descriptive results for experimental items.

	Percentage of correct answers	95% CI
Sentences targeting <i>parce que</i>	81%	[79.43, 83.37]
Sentences targeting <i>puisque</i>	46%	[43.61, 48.66]

#### 6.3.2. Measurements of language proficiency

In order to avoid multicollinearity of the two measurements, we conducted separate models for the Lextale and the Grammar task.

In the Lextale task, non-native speakers obtained a mean score of 23.4 (42%, 95% CI [23.06, 23.74]). In order to assess whether this score was predictive for the mastery of both connectives for the non-native speakers, we conducted logistic mixed effects models on the data of the non-native speakers. As in the first experiment, we contrast coded the predictor variable *Connective* (*parce que* set to 0.5, *puisque* set to -0.5). Adding this contrast to the null model (containing only *Connective* as a random slope per *Participant*, and both *Participant* and *Item* as random intercepts) as a fixed effect improved the model,  $\chi(1) = 41.81, p = 1.01 \times 10^{-10}$ , as did further adding *Lextale*,  $\chi(2) = 8.65, p < .05$ . The output of the final model is reported in Table 6.

**Table 6**

Experiment 2. Output of final model including the contrast of Connectives and Lextale as interacting fixed effects and *Connective* as a random slope per *Participants* and *Items* as a random effect.

	Estimate	SE	z	Pr(> z )
(Intercept)	.98	.15	6.52	$7.03 \times 10^{-11}$
<i>puisque</i> vs. <i>parce que</i>	-2.14	.31	-6.93	$4.30 \times 10^{-12}$
Lextale	.02	.01	2.61	<.01
Connective ( <i>puisque</i> ): Lextale	-0.02	.01	-1.16	.25

Importantly, for non-native speakers, *Lextale* scores predicted a better mastery for both connectives. Supplementary analyses (which can be accessed here: [https://osf.io/hgfsu/?view\\_only=b3bf160f9854469d8d8f3e602ca04356](https://osf.io/hgfsu/?view_only=b3bf160f9854469d8d8f3e602ca04356)) showed a similar predicting effect of the *Lextale* scores for the filler sentences targeting *alors* ('so') and *mais* ('but'). (see Fig. 4)

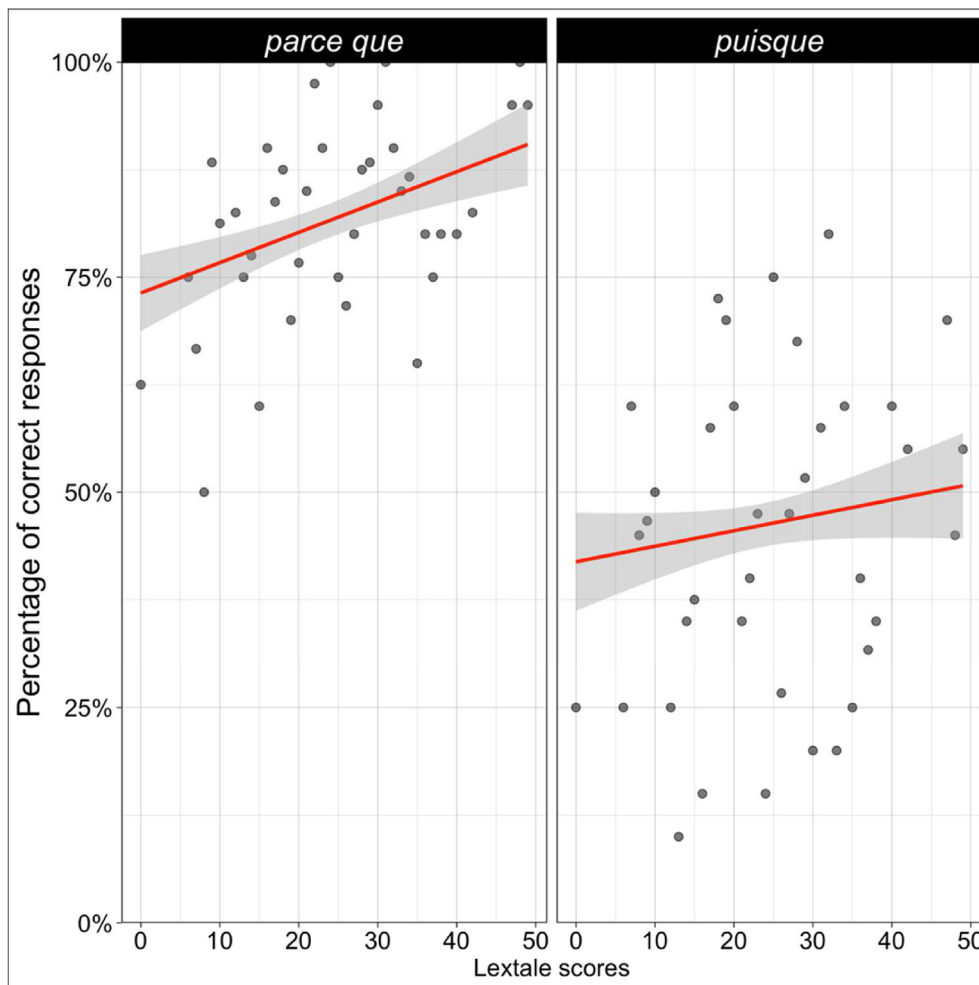


Fig. 4. Experiment 2. Correlations of *Lextale* scores with the mean scores for sentences targeting *parce que* ('because') and *puisque* ('since').

In the Grammar task, non-native speakers had a mean score of 56% (95% CI [55.82, 56.27]).

Adding the contrast of *Connective* to the null model improved the model's fit,  $\chi(1) = 41.81, p = 1.01e^{-10}$ , as did further adding *Grammar Task Score*,  $\chi(2) = 16.89, p < .001$ . The output of the final model<sup>4</sup> is reported in Table 7.

Table 7

Experiment 2. Output of final model including the contrast of *Connectives* and *Grammar task* as interacting fixed effects and *Connective* as a random slope per *Participants* and *Items* as a random effect.

	Estimate	SE	z	Pr(> z )
(Intercept)	.97	.13	7.39	1.48e <sup>-13</sup>
<i>puisque</i> vs. <i>parce que</i>	-1.94	.28	-6.96	3.48e <sup>-12</sup>
<i>Grammar Task</i>	.04	.01	4.32	1.59e <sup>-05</sup>
<i>Connective (puisque): Grammar Task</i>	<.001	.02	-0.01	.99

<sup>4</sup> We did not conduct a model including both *Lextale* and *Grammar task* scores in order to meet the assumption of the absence of multi-collinearity for logistic mixed-effects models (Schreiber-Gregory, 2018). Indeed, a positive, moderate correlation was found between the two scores ( $r(73) = .40, p = 2.2e^{-16}$ ).

As illustrated in Fig. 5 and by the main effect of *Grammar Task* in Table 7, a higher score in the grammar task predicted a better score for both connectives (as well as for the sentences targeting *alors* and *mais*, see supplementary analyses).

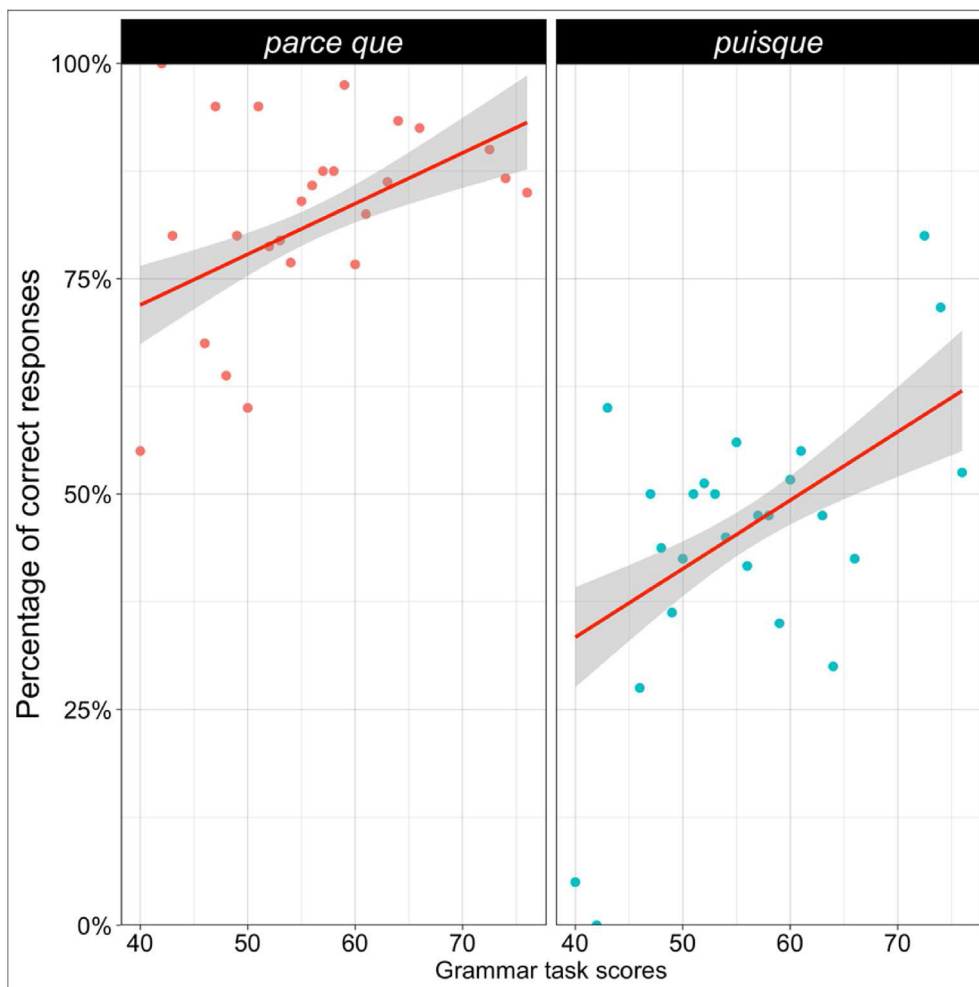


Fig. 5. Experiment 2. Correlations of the Grammar task scores with the mean scores for sentences targeting *parce que* ('because') and *puisque* ('since').

### 6.3.3. Comparing syntactic vs. pragmatic constraints of 'parce que'

Adding the contrast of *Constraint* (i.e., *syntactic*, *pragmatic*) as a fixed effect to the model did not improve its fit ( $\chi(1) = .11$ ,  $p = .74$ ). We conclude that non-native speakers did not benefit more from one of the constraints tested than from the other (a mean of 81% of correct answers for pragmatic constraints vs. a mean of 82% of correct answers for syntactic constraints) (see Fig. 6).

### 6.3.4. Comparing the two types of pragmatic constraints of 'puisque'

For the data of the sentences targeting *puisque*, the observation was similar: *Constraints* (i.e., *citation*, *obvious*) as fixed effect to the model did not improve its fit,  $\chi(1) = .12$ ,  $p = .73$  (as can be seen in Fig. 7, a mean of 45% of correct answers for *citation* vs. a mean of 47% of correct answers for *obvious*).

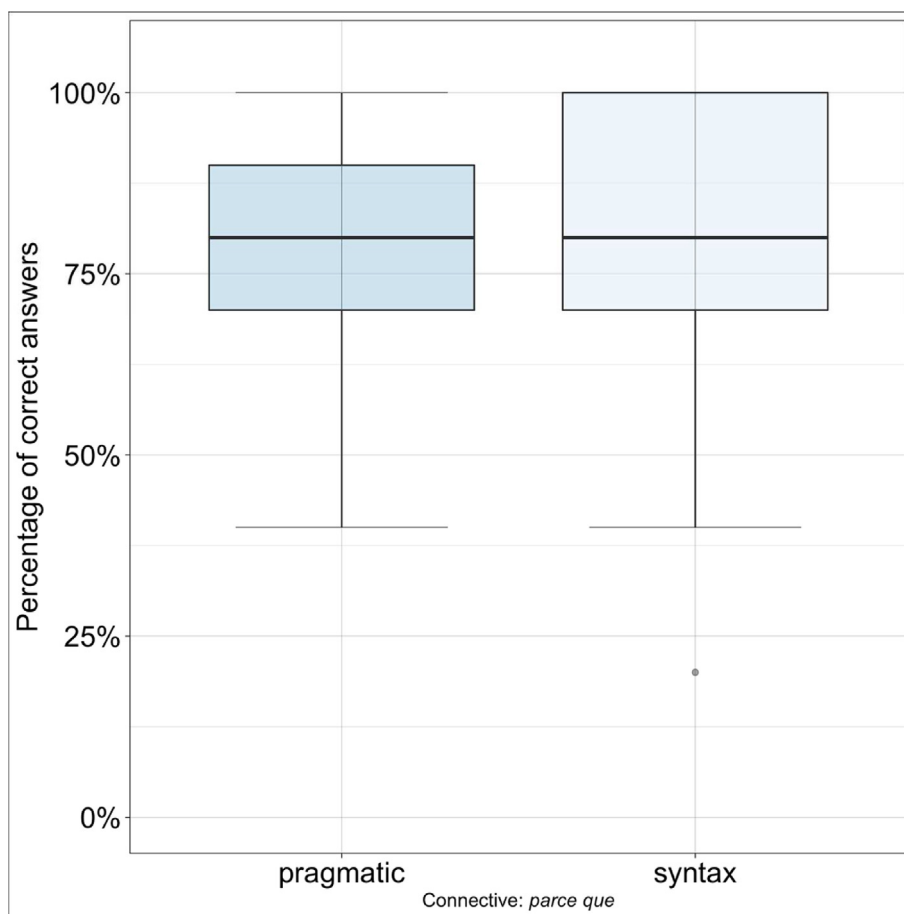


Fig. 6. Experiment 2. Percentage of correct answers for *parce que*.

### 6.3.5. Results of participants based on their answers to the open meta linguistic question

Given the low scores observed for *puisque*, we wanted to assess in a supplementary and more exploratory analysis whether some non-native participants were actually able to master *puisque*, and if so, which factors would predict a better mastery. In order to do so, we analyzed the explanations given by the participants to the metalinguistic question “What is the difference between *parce que* and *puisque*?”. We expected that participants who had a correct and explicit metalinguistic understanding about the different pragmatic functions of the two connectives were more likely to score higher for *puisque* than participants who only regarded *puisque* as the translation of *since* (without a conscious understanding of their difference). As such, the analysis of the answers may provide important insights into how a better mastery of the pragmatically rich embedded connective *puisque* can be attained.

All answers were classified into the following categories: *incorrect*, *partially correct* and *correct*. Responses from participants who did not give an answer and indicated “I don’t know” were classified as *no answer* ( $n = 18$ ). The answers of participants who included incorrect assumptions about the difference between these connectives (or the lack thereof) formed the group of *incorrect* answers ( $n = 18$ ). Answers were rated as *partially correct* when they stated that *puisque* is the equivalent or translation of *since* (and *parce que* of *because*), however without further indications ( $n = 27$ ). *Correct* were answers that involved a correct and more detailed description of both connectives, for example, the fact that *puisque* is used, unlike *parce que*, to introduce already known information ( $n = 12$ ).

Three of the authors independently rated the 75 answers. In cases when there was a disagreement between two raters (3 cases, 4% of all ratings), we decided based on the decision of the third rater. There were no cases of disagreement of all three raters. Fleiss’ Kappa of inter-rater agreement indicated an *almost perfect agreement* (following Landis and Koch, 1977; see Table 8) for all groups. All answers and our evaluation of them can be found in the supplementary analyses, accessible here: [https://osf.io/hgfsu/?view\\_only=b3bf160f9854469d8d8f3e602ca04356](https://osf.io/hgfsu/?view_only=b3bf160f9854469d8d8f3e602ca04356).

The mean scores for each participant’s group are reported in Table 9.

Hence, the post hoc classification of participants into four groups based on their answers indicates that participants who gave a correct answer scored higher for *puisque* than those who gave no answer, an incorrect answer or only a partially correct answer.

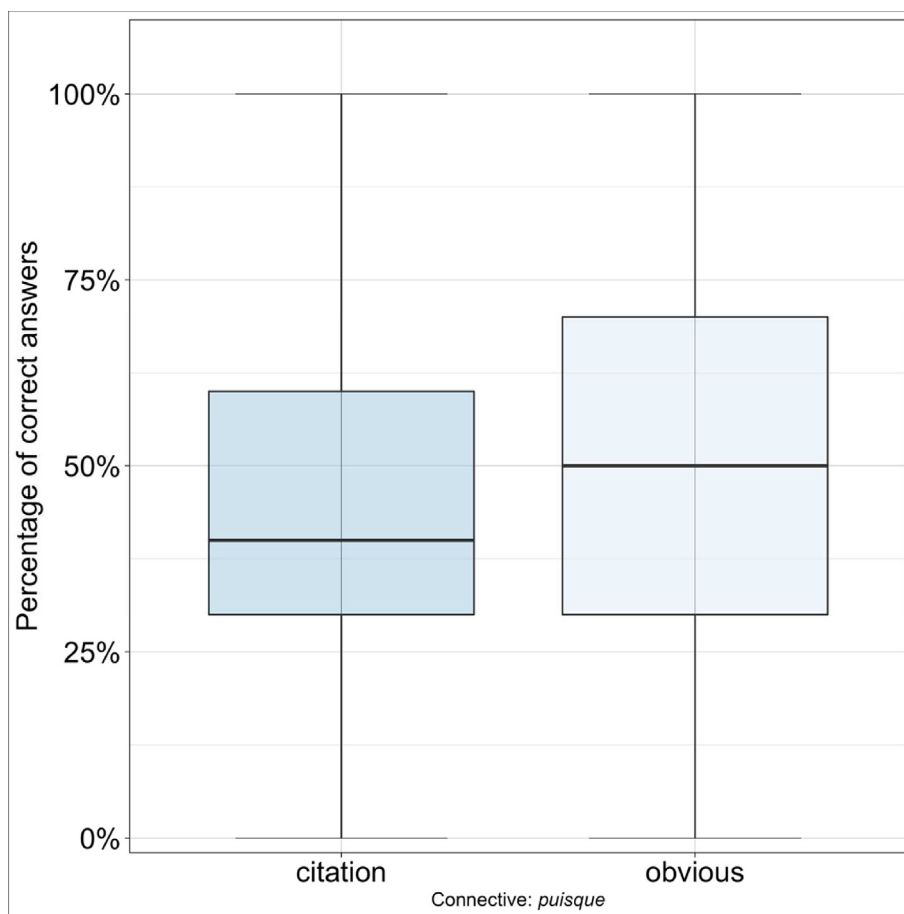


Fig. 7. Experiment 2. Percentage of correct answers for *puisque*.

Table 8

Fleiss' Kappa of inter-rater agreement for the classification of participants in four groups.

category	K	z	P
no text	1	15	<.0001
incorrect	.95	14.29	<.0001
partially correct	.96	14.13	<.0001
correct	.97	13.49	<.0001

#### 6.4. Discussion

In our second experiment, we replicated the first experiment with non-native participants, who hypothetically have fuzzier intuitions about when to use which connective. Regarding the decision to use either *puisque* ('since') or *parce que* ('because'), our results demonstrate that non-native speakers showed important differences in the mastery of these two connectives, as they mastered *parce que* considerably better than *puisque*. This difference could in fact be explained by a poor mastery of the latter. Since participants chose mainly between *parce que* and *puisque* for the causal relations included in the experiment, the 46% rate of correct responses for non-native speakers might even be considered to represent chance level (although they had four choices). As such, it can be assumed that the lack of effect for the pragmatic constraints for *puisque* was due to the fact that non-native speakers were simply unaware of both constraints. Also, our results indicate that non-

Table 9

Mean scores of correct answers for *puisque* in the main task (in percentage) for the four groups of participants.

Group (according to answers to meta-linguistic question)	Achieved mean score for sentences targeting <i>puisque</i>	CI 95%
no answer	37%	[31.67, 41.67]
incorrect	39%	[33.83, 43.95]
partially correct	49%	[44.29, 52.75]
correct	66%	[59.80, 71.88]

native speakers did not benefit more from the syntactic constraint of *parce que* than from the pragmatic constraint. We believe that this might be due to the lower sensitivity of non-native speakers to pick up on syntactic cues, whereas this does not necessarily imply a general lack of understanding of the coherence relations under scrutiny. Indeed, participants were able to identify the correct coherence relation, i.e., in 92% of the cases, as they selected either *puisque* or *parce que* to indicate a consequence-cause relation (filler sentences were also judged correctly in 87% of the sentences).

## 7. General discussion

Across two experiments, we evaluated the sensitivity of French speakers to correctly use connectives in consequence-cause relations, as well as the constraints that allow the use of either *parce que* ('because') or *puisque* ('since'). The constraints that imposed only *parce que* as the appropriate choice were syntactic (i.e., use of a cleft sentence) and pragmatic (i.e., replying to a question). The constraints allowing only *puisque* were both of a pragmatic nature.

In the first experiment, when specifically addressing different uses of the connectives under scrutiny, we observed a high consensus among speakers of the heterogeneous group of native speakers on the syntactic constraint imposing *parce que* ('because') as the appropriate connective. As we did not observe predicting effects of our language proficiency measures in Experiment 1 (Lextale and Grammar tasks) we conclude that the knowledge of the constraint is pervasive amongst native speakers. In Experiment 2, non-native speakers did lack sensitivity to this constraint, yet show some global proficiency effect. Regarding the constraints for *parce que*, we observed in Experiment 1 lower performance scores when the constraint was a pragmatic one. We conclude that pragmatic cues indicating the necessity to choose a certain connective are not as clear-cut as syntactic ones.

The type of pragmatic constraint (as tested by comparing two types of constraints for *puisque*), however, does not seem to be decisive, as native speakers in Experiment 1 picked up both pragmatic constraints for *puisque* similarly well. At first glance, this finding seems to contradict the one reported by Zufferey (2012), who observed that native participants accepted less univocally *puisque* in a sentence containing a citation, while showing a high acceptance for *puisque* for sentences containing an obvious statement. However, this difference is in fact not all that surprising: it most likely stems from the fact that we used multiple sentences for this condition, whereas Zufferey (2012) only used one sentence, which, as discussed in the study, "might not have been salient enough" (Zufferey, 2012:152) to indicate obviousness. It should be noted, however, that some of the pragmatic constraints we presented may have been more salient than others, although they were absorbed by a general (lack of) effect. Future studies may want to identify more precisely the effect of saliency on the sensitivity to pragmatic constraints.

We would now like to discuss in more detail the results of the non-native participants for sentences targeting *puisque*, as these were low and could, as discussed above, even be considered at chance-level.

### 7.1. Factors predicting the mastery of *puisque* for non-native speakers

Firstly, we observed that non-native speakers with a higher language proficiency, as measured by the Lextale task, showed a better mastery of *puisque* than those with a lower language proficiency. This was the case for all connectives, including the connectives of the filler sentences. Similarly, a higher grammar proficiency was predictive for a better mastery of *puisque*. An interesting finding in this respect is that declarative knowledge also predicted a better mastery for *puisque*: Participants who gave a correct answer to the question about the difference between *puisque* and *parce que* obtained a higher score for sentences targeting *puisque*. This finding indicates that a better mastery of this connective can be obtained, provided language learners are consciously aware of its pragmatic functions. Yet, this finding may be somewhat limited by the relatively low number of participants in the four groups. Also, it should be noted that participants with a good meta-understanding of the pragmatic dimension of *puisque* had, on average, only 66% of correct responses for sentences targeting this connective - which could be considered low in light of the scores obtained for *parce que* (90%) or the filler connectives *alors* (98%) and *mais* (99%).

The analysis of the explanations given by the participants leads to a second interesting observation. As indicated by several participants, the French connective *puisque* can be translated by the English *since* (see also Zufferey and Cartoni, 2012). Yet, this mapping alone did not help the participants to better master the French connective. In fact, these participants' scores can even be considered at chance level for sentences with *puisque* (49% of correct responses), as participants generally chose between *parce que* and *puisque* for the causal sentences. Hence, although *puisque* is commonly translated by *since*, it appears that a deeper understanding is needed for this connective. This observation is reinforced by the fact that we found several cases in which the mapping of the English *since* onto French *puisque* even led to erroneous assumptions about the French connective. As can be seen in the example of response (25) (classified as *incorrect* in the analysis), participants mistakenly mapped functions of English *since* onto French *puisque*.

25) *Puisque* is "since", so there's a time element. For example, if I was saying that I hadn't spoken to someone since they got back from a holiday, I would use "puisque". *Parce que* means *because* so there's no time involved in that, it's just an explanation

*Since* serves two functions of temporality or causality (Zamel, 1983), whereas *puisque* is not used to indicate temporal links (Zufferey and Cartoni, 2012). By mapping all functions of *since* to *puisque*, non-native speakers erroneously fall prey to negative transfer from their L1. It has indeed been shown that negative transfer can create confusion about the different functions of an L2-connective when these do not correspond fully to the functions of a seemingly equivalent L1-connective (Wetzel et al., 2022a).



The mapping of the function of English *since* onto French *puisque* - although a common translation pair - is not enough to ensure a solid mastery of *puisque* and, on the contrary, even led to erroneous assumptions about its functions. However, participants who had an increased language proficiency and participants who had a correct meta-understanding of the functions of *puisque* were able to score significantly higher for this connective. Future research could assess how this meta-understanding can be taught and retained, preventing language learners from using connectives as devices “that may be sprinkled over a text in order to give it an ‘educated’ or ‘academic’ look” (Crewe, 1990: 316).

## 8. Conclusion and future research avenues

Our study provides several insights into French native and non-native speakers' sensitivity to pragmatic and syntactic constraints. First, our findings show that pragmatic constraints are less salient than syntactic constraints and that this difference applies to different groups of speakers. As such, the results obtained validate the description of French connectives (Degand and Pander Maat, 2003; Degand and Fagard, 2018; Fagard and Degand, 2008; Lambrecht et al., 2006). Also, our systematic approach to pragmatic constraints evokes a more general, theoretical issue: in Experiment 1 we observed that native speakers picked up on pragmatic cues less well than on syntactic ones, which questions whether pragmatic cues are really constraining in a strict sense (i.e., making one connective the only possible choice) or whether they elicited only a very high preference for a certain connective (while not totally excluding the other). Indeed, while constructing the experimental sentences, we estimated that in all tested sentences targeting *puisque* the only appropriate connective would be *puisque*, but, as documented in Experiment 1, not all participants agreed on this. Future research is therefore needed to examine the extent to which pragmatic cues can represent strict constraints that limit the choice for a specific connective in an exclusive manner.

Finally, as the scores from our language proficiency measures showed wide ranges, our study is in line with studies showing that native speakers are, in fact, a heterogeneous group in terms of language proficiency (e.g., Kamalski et al., 2008; Van Silfhout, Evers-Vermeul, Mak & Sanders, 2014; Scholman et al., 2020). Yet, the proficiency measurements in our experiment were not predictive of a higher sensitivity to the pragmatic constraints among native speakers. While we were able to observe a link between a higher pragmatic sensitivity and a higher language proficiency for non-native speakers, we conclude that more research is needed that identifies factors predicting a higher sensitivity to pragmatic cues in native speakers. In this regard, future research could address factors pertaining to language competence other than lexical and grammar competences. For example, metalinguistic knowledge or awareness may be good candidates, as they have been shown to correlate with a better mastery of pragmatic competence in L2 (e.g., Rose, 2005). In terms of language development, future research could also address the grounding factors of the acquisition of the sensitivity to pragmatic and syntactic cues, both on a psychosocial level (e.g., socio-educative context, etc.) and a developmental level (e.g., age of acquisition of other linguistic properties, etc.).

## Funding

This work was supported by the Swiss National Science Foundation (SNSF) under Grant number (100012 184882) granted to SZ.

## Declaration of competing interest

The authors declare none.

## Ethics and consent

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the ethics protocol was approved by the Swiss National Science Foundation (100012 184882).

## Data availability

All of our data and R - scripts can be retrieved at [https://osf.io/hgfsu/?view\\_only=b3bf160f9854469d8d8f3e602ca04356](https://osf.io/hgfsu/?view_only=b3bf160f9854469d8d8f3e602ca04356).

## References

- Altenberg, Bengt, Tapper, Marie, 1998. The use of adverbial connectors in advanced Swedish learners' written English. In: Granger, S. (Ed.), *Learner English on Computer*. Routledge, London, pp. 80–93.
- Ariel, M., 1985. The discourse functions of given information. *Theor. Ling.* 12, 99–113.
- Baayen, Harald, 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press, Cambridge.
- Bardovi-Harlig, Kathleen, Dörnyei, Zoltán, 1998. Do Language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *Tesol Q.* 32 (2), 233–259. <https://doi.org/10.2307/3587583>.
- Bardovi-Harlig, Kathleen, Hartford, Beverly, 1990. Congruence in native and nonnative conversations: status balance in the academic advising session. *Lang. Learn.* 40 (4), 467–501. <https://doi.org/10.1111/j.1467-1770.1990.tb00603.x>.

- Bardovi-Harlig, Kathleen, Hartford, Beverly, 1993. Learning the rules of academic talk: a longitudinal study of pragmatic change. *Stud. Sec. Lang. Acquis.* 15 (3), 279–304. <https://doi.org/10.1017/S0272263100012122>.
- Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68 (3), 255–278.
- Bates, Douglas, Mächler, Martin, Bolker, Ben, Walker, Steve, 2014. Fitting linear mixed-effects models using lme4. *ArXiv:1406.5823 [Stat]*. <http://arxiv.org/abs/1406.5823>.
- Bell, David, 2007. Sentence-initial and and but in academic writing. *Pragmatics* 17 (2), 183–201. <https://doi.org/10.1075/prag.17.2.01bel>.
- Bell, David, 2010. Nevertheless, still and yet: concessive cancellative discourse markers. *J. Pragmat.* 42 (7), 1912–1927. <https://doi.org/10.1016/j.pragma.2009.12.010>.
- Bentolila, Fernand, 1986. Car en français écrit. *Langues Linguistique* 22 (2), 95–115.
- Braze, David, Tabor, Whitney, Shankweiler, Donald, Mencl, Einar, 2007. Speaking up for vocabulary: reading skill differences in young adults. *J. Learn. Disabil.* 40 (3), 226–243. <https://doi.org/10.1177/00222194070400030401>.
- Brysaert, Marc, 2013. Lextale\_FR a fast, free, and efficient test to measure language proficiency in French. *Psychol. Belg.* 53 (1), 23–37.
- Buysse, L., 2012. So as a multifunctional discourse marker in native and learner speech. *J. Pragmat.* 44 (13), 1764–1782.
- Buysse, L., 2014. “So what’s a year in a lifetime so.” Non-prefatory use of so in native and learner English. *Text Talk* 34 (1), 23–47.
- Castele, An Vande, Collewaert, Kim, 2013. The use of discourse markers in Spanish language learners’ written compositions. *Procedia - Social and Behavioral Sciences* 95, 550–556. <https://doi.org/10.1016/j.sbspro.2013.10.681>.
- Chen, Ping Ju, 2014. The comparison of intermediate and advanced Chinese learners’ use of English adverbial connectors in academic writing. *International Journal on Studies in English Language and Literature* 2 (8), 85–92.
- Crewe, William, 1990. The illogic of logical connectives. *ELT J.* 44 (4), 316–325. <https://doi.org/10.1093/elt/44.4.316>.
- Crible, Ludvine, Cuenca, Maria-Josep, 2017. Discourse markers in speech: distinctive features and corpus annotation. *Dialogue & Discourse* 8 (2), 149–166.
- Degand, Liesbeth, Pander Maat, Henk, 2003. A contrastive study of Dutch and French causal connectives on the speaker involvement scale. In: *LOT Occasional Series 1. LOT, Netherlands Graduate School of Linguistics, Amsterdam*, pp. 175–199. <https://dspace.library.uu.nl/handle/1874/295389>.
- Degand, Liesbeth, 2004. Contrastive analyses, translation, and speaker involvement: the case of Puisque and Aangezien. *Language, Culture and Mind* 1–20.
- Degand, Liesbeth, Fagard, Benjamin, 2018. Inter-subjectification des connecteurs: le cas de car et parce que. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto* 3, 119–136, 0.
- Fagard, Benjamin, Degand, Liesbeth, 2008. La fortune des mots: grandeur et décadence de “car”. *Congrès Mondial de Linguistique Française 2008*. [https://doi.org/10.1051/cmlf08213\\_036](https://doi.org/10.1051/cmlf08213_036).
- Fraser, Bruce, 1998. Contrastive discourse markers in English. In: Jucker, Andreas, Ziv, Yael (Eds.), *Discourse Markers. Description and Theory*. John Benjamins Publishing, Amsterdam, pp. 301–326.
- Groupe λ-1, 1975. Car, parce que, puisque. *Rev. Romane* 10, 248–280.
- Halliday, Michael, Hasan, Ruqaiya, 1976. *Cohesion in English*. Longman, London.
- Hitzeman, Janet, 1995. A Constraint-Based Grammar of English Temporal Connectives. *ESPRIT Basic Research Project*, 6665, R2.
- Huang, L.F., Lin, Y.L., Gráf, T., 2022. Development of the use of discourse markers across different fluency levels of CEFR: a learner corpus analysis. *Pragmatics* 33 (1), 49–77.
- Kamalski, Judith, Lentz, Leo, Sanders, Ted, Zwaan, Rolf, 2008. The forewarning effect of coherence markers in persuasive discourse: evidence from persuasion and processing. *Discourse Process* 45 (6), 545–579. <https://doi.org/10.1080/01638530802069983>.
- Kidd, Evan, Donnelly, Seamus, Christiansen, Morten, 2018. Individual differences in language acquisition and processing. *Trends Cognit. Sci.* 22 (2), 154–169. <https://doi.org/10.1016/j.tics.2017.11.006>.
- Knott, Alistair, 1996. A Data-Driven Methodology for Motivating a Set of Coherence Relations. University of Edinburgh. PhD Thesis. <https://era.ed.ac.uk/handle/1842/583>.
- Lambrech, Knud, Bordeaux, Julia, Reichle, Robert, 2006. Cognitive constraints on assertion scope: the case of spoken French parce que. In: Nishida, Chiyo, Montreuil, Jean-Pierre (Eds.), *New Perspectives on Romance Linguistics: Volume 1: Morphology, Syntax, Semantics, and Pragmatics. Selected Papers From the 35th Linguistic Symposium on Romance Languages*. John Benjamin, Amsterdam, pp. 143–154.
- Landis, Richard, Koch, Gary, 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1), 159–174. <https://doi.org/10.2307/2529310>.
- Lee, Kent, 2013. Korean ESL learners’ use of connectors in English academic writing. *Engl. Lang. Teach.* 25, 81–103. <https://doi.org/10.17936/pkelt.2013.25.2.005>.
- Leedham, Maria, Cai, Guozhi, 2013. Using a corpus approach to explore the influence of teaching materials on Chinese students’ use of linking adverbials. *J. Sec. Lang. Writ.* 22 (4), 374–389. <https://doi.org/10.1016/j.jslw.2013.07.002>. Besides ... on the other hand.
- Leeman, Danielle, 1994. Remarques sur puisque et sur car. In: Flament-Boistrancourt, Danielle (Ed.), *Théories, données et pratiques en français langue étrangère*. Presses Universitaires de Lille, Lille, pp. 113–128.
- Lenth, Russell, 2016. Least-squares means: the R package lsmeans. *J. Stat. Software* 69 (1), 1–33. <https://doi.org/10.18637/jss.v069.i01>.
- Li, Fang, Evers-Vermeul, Jacqueline, Sanders, Ted, 2013. Subjectivity and result marking in Mandarin. *Chinese Language and Discourse* 4 (1), 74–119. <https://doi.org/10.1075/cld.4.1.03li>.
- Macário, Lopes, Ana Christina, 2009. Justification: a coherence relation. *Pragmatics* 19 (2), 241–252.
- Moeschler, Jacques, 2005. Connecteurs pragmatiques, inférences directionnelles et représentations mentales. *Cah. Chronos* 12, 35–50.
- Müller, S., 2005. Discourse Markers in Native and Non-native English Discourse. *Discourse Markers in Native and Non-native English Discourse*, vols. 1–310. John Benjamins, Amsterdam.
- Myung-Jeong, Ha, 2016. Linking adverbials in first-year Korean university EFL learners’ writing: a corpus-informed analysis. *Comput. Assist. Lang. Learn.* 29 (6), 1090–1101. <https://doi.org/10.1080/09588221.2015.1068814>.
- Pit, Mirna, 2007. Cross-linguistic analyses of backward causal connectives in Dutch, German and French. *Lang. Contrast* 7 (1), 53–82. <https://doi.org/10.1075/lic.7.1.04pit>.
- Pitler, Emily, Nenkova, Ani, 2009. Using syntax to disambiguate explicit discourse connectives in text. *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers on - ACL-IJCNLP '09* 13. <https://doi.org/10.3115/1667583.1667589>.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rose, K.R., 2005. On the effects of instruction in second language pragmatics. *System* 33 (3), 385–399. <https://doi.org/10.1016/j.system.2005.06.003>.
- Roze, C., Danlos, L., Muller, P., 2012. LEXCONN: a French lexicon of discourse connectives. *Discours. Revue de Linguistique, Psycholinguistique et Informatique. A Journal of Linguistics, Psycholinguistics and Computational Linguistics* 10. <https://doi.org/10.4000/discours.8645>. Article 10.
- Sanders, Ted, Spooren, Wilbert, Noordman, Leo, 1992. Toward a taxonomy of coherence relations. *Discourse Process* 15 (1), 1–35. <https://doi.org/10.1080/01638539209544800>.
- Santana, Andrea, Spooren, Wilbert, Nieuwenhuisen, Dorien, Sanders, Ted, 2021. Do Spanish Causal Connectives Vary in Subjectivity? what Crowdsourcing Data Reveal about Native Speakers’ Preferences. *Text & Talk*. <https://doi.org/10.1515/text-2019-0102>.
- Scholman, Merel, Demberg, Vera, Sanders, Ted, 2020. Individual differences in expecting coherence relations: exploring the variability in sensitivity to contextual signals in discourse. *Discourse Process* 1–18. <https://doi.org/10.1080/0163853X.2020.1813492>.
- Schreiber-Gregory, Deanna, 2018. Ridge Regression and multicollinearity: an in-depth review. *Model Assisted Statistics Appl.* 13 (4), 359–365. <https://doi.org/10.3233/MAS-180446>.

- Schumann, J., Zufferey, S., Oswald, S., 2020. The linguistic formulation of fallacies matters: the case of causal connectives. *Argumentation* 35 (3), 361–388. <https://doi.org/10.1007/s10503-020-09540-0>.
- Simon, Anne Catherine, Degand, Liesbeth, 2007. Connecteurs de causalité, implication du locuteur et profils prosodiques: le cas de car et de parce que. *J. Fr. Lang. Stud.* 17 (3), 323–341. <https://doi.org/10.1017/S095926950700302X>.
- Sperber, Dan, Wilson, D., 1986. *Relevance: Communication and Cognition* (2). Blackwell Publishers, Hoboken.
- Stukker, Ninke, Sanders, Ted, 2012. Subjectivity and prototype structure in causal connectives: a cross-linguistic perspective. *J. Pragmat.* 44 (2), 169–190. <https://doi.org/10.1016/j.pragma.2011.06.011>.
- Sweetser, Eve, 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure* (40). Cambridge University Press, Cambridge.
- Tskhovrebova, E., Zufferey, S., Gygax, P., 2022. Individual variations in the mastery of discourse connectives from teenage years to adulthood. *Lang. Learn.* 72 (2), 412–455. <https://doi.org/10.1111/lang.12481>.
- Van Silfhout, Gerdineke, Evers-Vermeul, Jacqueline, Mak, Willem, M., Sanders, Ted, 2014. Connectives and layout as processing signals: how textual features affect students' processing and text representation. *J. Educ. Psychol.* 106 (4), 1036–1048. <https://doi.org/10.1037/a0036293>.
- Wetzel, M., Zufferey, S., Gygax, P., 2020. Second language acquisition and the mastery of discourse connectives: assessing the factors that hinder L2-learners from mastering French connectives. *Languages* 5 (3), 35. <https://doi.org/10.3390/languages5030035>.
- Wetzel, M., Crible, L., Zufferey, S., 2022a. Processing clause-internal discourse relations in a second language: a case study of specifications in German and French. *J. Second Lang. Stud.* 5 (2), 206–234. <https://doi.org/10.1075/jsls.21032.wet>.
- Wetzel, M., Zufferey, S., Gygax, P., 2022b. How robust is discourse processing for native readers? The role of connectives and the coherence relations they convey. *Front. Psychol.* 13, 1–13. <https://doi.org/10.3389/fpsyg.2022.822151>.
- Yung, Frances, Jungbluth, Jana, Demberg, Vera, 2021. Limits to the rational production of discourse connectives. *Front. Psychol.* 12. <https://doi.org/10.3389/fpsyg.2021.660730>.
- Zamel, Vivian, 1983. Teaching those missing links in writing. *ELT J.* 37 (1), 22–29. <https://doi.org/10.1093/elt/37.1.22>.
- Zufferey, S., 2012. "Car, parce que, puisque" revisited: three empirical studies on French causal connectives. *J. Pragmat.* 44 (2), 138–153. <https://doi.org/10.1016/j.pragma.2011.09.018>.
- Zufferey, S., 2014. Givenness, procedural meaning and connectives. The case of French *puisque*. *J. Pragmat.* 62, 121–135. <https://doi.org/10.1016/j.pragma.2013.09.022>.
- Zufferey, S., Cartoni, B., 2012. English and French causal connectives in contrast. *Languages in Contrast. Int. J. Contrast. Linguist.* 12 (2), 232–250. <https://doi.org/10.1075/lic.12.2.06zuf>.
- Zufferey, S., Gygax, P., 2020. "Roger Broke His Tooth. 'However', He Went to the Dentist": why some readers struggle to evaluate wrong (and right) uses of connectives. *Discourse Process.* 57 (2), 184–200. <https://doi.org/10.1080/0163853X.2019.1607446>.
- Zufferey, S., Mak, W., Verbrugge, S., Sanders, T., 2018. Usage and processing of the French causal connectives 'car' and 'parce que'. *J. French Lang. Stud.* 28 (1), 85–112. <https://doi.org/10.1017/S0959269517000084>.

**Mathis Wetzel** is a PhD student at the University of Bern, Switzerland. His dissertation deals with the acquisition of discourse connectives in French as a second language within the SNF-funded project "Discovering discourse: the acquisition of connectives in L1 and L2" by Prof. Dr. Zufferey.

**Ekaterina Tskhovrebova** is a PhD student at the University of Bern, Switzerland. Her dissertation deals with the acquisition of discourse connectives in French as a first language within the SNF-funded project "Discovering discourse: the acquisition of connectives in L1 and L2" by Prof. Dr. Zufferey.

**Pascal M. Gygax** is senior lecturer at the Department of Psychology, University of Fribourg, Switzerland, where he leads the Psycholinguistics and Applied Social Psychology Group. His research interests include issues related to gender, social cognition and the effect of language on thought.

**Sandrine Zufferey** is full Professor of French linguistics at the University of Bern in Switzerland. She works on discourse processing, and more specifically on the first and second language acquisition of discourse connectives. She is more generally interested in cross-linguistic comparisons, and the way differences of linguistic encoding between languages affect cognitive processes.