Glossa a journal of general linguistics

Blochowiak, Joanna, et al. 2020. What type of subjectivity lies behind French causal connectives? A corpus-based comparative investigation of *car* and *parce que*. *Glossa: a journal of general linguistics* 5(1): 50. 1–36. DOI: https://doi.org/10.5334/gjgl.1077

RESEARCH

What type of subjectivity lies behind French causal connectives? A corpus-based comparative investigation of *car* and *parce que*

Joanna Blochowiak[1], Cristina Grisot[2] and Liesbeth Degand[1]

[1] Université catholique de Louvain, Louvain-la-Neuve, BE

[2] Université de Genève, Genève, CH

Corresponding author: Joanna Blochowiak (joanna.blochowiak@uclouvain.be)

In French, the difference between the causal connectives *parce que* and *car* is traditionally related to the prototypical causal relations they are meant to convey. The main claim is that *car* conveys more subjective relations and is also used in higher register language, whereas *parce que* is equally well-suited to both types of relations. In line with recent studies, this contribution questions the clear-cut distinction between the two connectives on the basis of a comparative corpus investigation with annotation tasks (journalistic and text messaging registers). Our results do not corroborate the traditional hypotheses that *car* is used to express more subjective relations and it is restricted to higher register language. On the contrary, we find that *car* has a strong tendency to be perceived by addressees as providing the information in a more objective way. Our empirical investigation has allowed us to put forth a modified notion of subjectivity which is associated with *car* and *parce que*: we distinguish between the more classic approach – the type of subjectivity related to causal relations, and a novel approach – the evaluative type of subjectivity related to the expressive use of language. We rely on the relevance-theoretic framework to spell out our theoretical proposal.

Keywords: causal connectives; French; descriptive subjectivity; evaluative subjectivity; corpus study; annotation tasks

1 Introduction

This paper addresses the issue of the contemporary use of the French causal connective *car* in opposition to another French causal connective *parce que* ('because'), from both empirical and theoretical perspectives. The question at stake with these two connectives is that native speakers of French would intuitively agree that there are differences between the usage of *car* and *parce que*, and various analyses have sought to pinpoint these differences within different theoretical frameworks (Groupe-λ-1 1975; Moeschler 1987; 2005; Ferrari 1992; Iordanskaja 1993; Debaisieux 2002; 2004; Degand & Pander Maat 2003; Lambrecht et al. 2006; Zufferey 2012). For example, one of the suggestions made is that *car* is used to express subjective causal relations such as in the case of opinions and speech acts whereas *parce que* may be used to express subjective and also objective causal relations between events (Iordanskaja 1993; Debaisieux 2002; 2004; Pander Maat & Degand 2003; Lambrecht et al. 2006; Simon & Degand 2007; Fagard & Degand 2008; Degand & Fagard 2012; Zufferey 2012). Yet the existing empirical and experimental studies have not arrived at clear-cut results confirming the existing theoretical claims (see Zufferey et al. 2018 for a recent attempt). For this reason, in this paper we decided to investigate two types of corpora opposing formal (the French newspaper *Le Monde*) and informal (an *SMS* corpus)

discourse genres in order to examine naturally occurring sentences containing the two connectives. In addition, we designed a connective annotation methodology that would enable us to compare different approaches of the notion of subjectivity expressed by *car* and *parce que*. In Study 1, we annotated the corpus data with a fine-grained classification using five ontological categories (events, actions, opinions, emotions and speech acts). On the basis of the fine-grained classification, we derived the cases of subjective and objective ontological categories. In Study 2, we annotated subjectivity as language comprehenders approach it in an intuitive way. The analyses of the various annotations made in Study 1 and 2 brought into light the fact that *car* has a tendency to be perceived as more objective than it was initially assumed. We interpret these results within a novel approach of subjectivity developed in the relevance-theoretic framework (Sperber & Wilson 2015; Wilson & Carston 2019), where it is understood as the evaluative property related to the speakers' expressive use of language.

This paper is structured as follows. The next two sections are dedicated to the current state of research in the domains of subjective and objective causal relations (Section 2) and French causal connectives (Section 3). Section 4 presents the empirical study carried out for this paper. A general discussion of the results, accompanied by a theoretical proposal aiming to explain our findings, is presented in Section 5. Finally, Section 6 provides a conclusion with a brief summary of the results.

2 Current state of research: objective and subjective relations and causal connectives

The objective/subjective distinction is one of the most frequently applied distinctions in current investigations of causal connectives in corpus-based studies (Pit 2003; 2006; 2007; Pander Maat & Sanders 2001; Pander Maat & Degand 2003; Zufferey 2012), processing experiments (Millis & Just 1994; Noordman & Vonk 1997; Sanders & Noordman 2000; Zufferey et al. 2015 for French) and eye-tracking testing (Canestrelli et al. 2013 for Dutch; Zufferey et al. 2018 for French). The distinction has its origin in Sweetser's (1990) differentiation between three domains of use for connectives: content (1), epistemic (2) and speech act domain (3).

(1) John fell **because** Mary pushed him.
(2) The neighbours must be at home, **because** the lights are on.
(3) Hurry up! **Because** we are late.

In the content domain, *because* conveys a causal relation between eventualities occurring in the world (the event of Mary's pushing John caused the event of John's falling); in the epistemic domain, *because* serves to provide evidence (the presence of lights) for a speaker's opinion or belief (the presence of neighbours); and in the speech act domain, *because* provides a justification for the use of a speech act such as a request. Content relations are considered to be objective because they refer to external, factual reality (causal relations between eventualities which are often independent of the speaker), while the relations conveyed in epistemic and speech act domains are claimed to be subjective as they refer to speakers' internal reality, providing justification for their opinions, beliefs or actions (cf. Sanders & Spooren 2015).

The mapping between the type of relation and a connective is not one-to-one, and is largely language-specific (cf. Sanders & Sweetser 2009). At one end of the spectrum are languages whose causal connectives follow the subjective/objective distinction quite closely. Dutch has long been considered an example of such languages, with the causal connective *omdat* typically conveying objective relations where *want* is dedicated to

subjective relations (Degand 1996; Sanders & Spooren 2009; Verhagen 2005). Numerous studies have confirmed these tendencies empirically with corpus-based analyses (cf. Stukker & Sanders 2012 for a summary) as well as reading time experiments and eye-tracking testing (Canestrelli et al. 2013). Crucially, it has been observed that the correct match between the connective and the relation it typically expresses speeds up the reading time of the sentences in the critical region – i.e. the region just after the connective (Canestrelli et al. 2013; Zufferey 2014).

At the other end of the spectrum, there are languages which do not have connectives specialized to one type of relation, because a single connective is equally able to express both types of relation, as it is true of English *because*. Nevertheless, this does not mean that there are no differences observed between sentences with *because* expressing the two types of relation. In general, speakers have more difficulty dealing with subjective relations than objective relations, as shown by various experiments involving reading times and reasoning tasks (Noordman & de Blijzer 2000; Sanders & Spooren 2009), as well as eye-tracking techniques (Traxler, Bybee & Pickering 1997). This phenomenon is usually attributed to the inherent complexity of subjective (or epistemic) relations; these are claimed to be harder to process, as they involve complex embedded propositions containing covert or overt epistemic attitude markers such as *I believe*, *I guess*, etc. It has been observed that the insertion of an epistemic marker facilitates the processing of subjective relations, as the presence of such markers makes it explicit that the first segment describes the speaker's internal mental state rather than an external real world event (Traxler, Sanford, Aked & Moxey 1997).[1] Importantly, the presence of the connective *want* in Dutch gives the addressee a cue for the type of relation she should expect; as a result, the difficulties processing subjective relations are cancelled out thanks to the presence of the appropriate connective (Canestrelli et al. 2013).

Finally, there are languages somewhere in the middle of the spectrum, in the sense that there is no clear mapping between a given causal connective and the type of relation it expresses (objective or subjective). This is, for instance, true of the two French causal connectives *parce que* and *car*, which we will discuss in the next section.

3 Current state of research: connectives *car* and *parce que*

The foundational work on French causal connectives was carried out by Groupe-λ-l (1975). This initial research has been fleshed out in various respects, not only for French (Anscombre & Ducrot 1983; Roulet et al. 1985; Bentolila 1986; Iordanskaja 1993; Nølke 1995; Bertin 1997; Moeschler 2003) but also for other languages (Pasch 1983 for German; Van Belle 1989 for Dutch). The analysis by Groupe-λ-l put forth two main dimensions for the contrast between *parce que* and *car*: the status of the utterances related by the connective (proposition or speech act), and their informational status (new vs. old information). Given these elements, *parce que* has been characterized as a propositional operator, as it links two propositions into one complex proposition, asserting the presence of a causal relation between the situations described therein. Regarding informational status, with *parce que* the two linked situations are considered to be known; hence, there is no new information added, except for the causal relation itself. *Car*, on the other hand, is claimed to be a speech act marker, as its role consists in linking two speech acts: the first speech act is asserted, and the second serves as a justification for uttering the first. The speech act followed by *car* constitutes new information, as it must stay open to negotiation and

[1] However, some studies suggest that the presence of an epistemic marker does not always clearly improve the processing of epistemic relations (Canestrelli, Sanders & Mak 2010).

thus possible contestation – or even revocation by means of a justificatory process, as illustrated in (4).

(4) A: The neighbours must be at home, **because** the lights are on.
B: No, they are on vacation. But they have an electronic system which switches the lights on in the evening to make it look like they are at home.

In more recent studies, Groupe-λ-1's proposal is often translated into the cognitive approach proposed by Sweetser, and the differences between French causal connectives are described within her above-mentioned domains of use.

As for other languages, a further parallel between domains of use and types of causality is often drawn, in the sense that the content domain is claimed to pair with the category of objective causality, and the epistemic and speech act domains conversely pair with subjective causality (Pander Maat & Degand 2001; Pit 2003; Stukker & Sanders 2012). Taking these two categories into account, there is a trend in the literature to consider *parce que* as the most polyvalent French causal connective which can be used with any type of relation (objective and subjective), where *car* is considered a prototypical marker of subjective relations (Iordanskaja 1993; Debaisieux 2002; 2004; Degand & Pander Maat 2003; Lambrecht et al. 2006; Simon & Degand 2007; Fagard & Degand 2008; Degand & Fagard 2012; Zufferey 2012). However, it should be stressed that the results with respect to speakers' usage and processing of the two connectives are inconsistent. For instance, Zufferey (2012) reports that these two connectives are interchangeable in many objective and subjective contexts when written, especially among the younger population of French native speakers. Moreover, some recent studies have questioned the subjective character of *car* (Zufferey et al. 2018), and other researchers report that *car* can also be used to express objective relations (Nazarenko 2000).

Given the relative difficulty of straightforwardly relating the use of *car* with the expression of subjective types of relation, an alternative hypothesis on the contemporary use of *car* has been proposed, according to which its use tends to be restricted to high register language, especially in (formal) writing such as newspaper articles (Zufferey 2012; Zufferey et al. 2018). This observation is in line with the fact that *car* is much less attested in oral speech, is even considered extinct in contemporary spoken French (Fagard & Degand 2012), with its use in written texts fluctuating (Frei 1982/1929; Bentolila 1986; Fagard & Degand 2008; Zufferey 2012).

In sum, two main hypotheses can be formulated from prior literature: (i) *car* is mainly used to convey subjective relations (Iordanskaja 1993; Debaisieux 2002, 2004; Degand & Pander Maat 2003; Lambrecht et al. 2006; Simon & Degand 2007; Fagard & Degand 2008; Simon & Degand 2012; Zufferey 2012; but see Nazarenko 2000 and Zufferey et al. 2018 for dissenting claims); and (ii) *car* is restricted to high register language (Zufferey 2012; Zufferey et al. 2018). These hypotheses generate different predictions for an empirical corpus-based study, as detailed in the next section.

4 Empirical investigation

The empirical work discussed in this section consists of two comparative corpus-based studies with annotation tasks, in which the use of the two connectives was analysed and classified by four independent annotators (students in Language Sciences and Speech Therapy) in two types of corpora – the portion of the *Le Monde* corpus from the year 2012, containing newspaper texts, and the *SMS* corpus from Cougnon (2012), containing text messages. Our choice for these corpora is motivated by their stylistic status: formal written prose for the former and a more informal and conversational style for the lat-

ter.[2] In section 4.1, we explain the method of *pragmatic annotation* and discuss how we calculated inter-annotator agreement rates. In section 4.2, we will formulate a series of predictions yielded by the hypotheses put forward in the literature review with respect to the investigation we performed.

4.1 Methodology

Annotation is the practice of adding interpretative linguistic information to a corpus (Leech 2005), thus enriching the original raw corpus. Annotated corpora often reveal a whole range of uses which would not have been observable and exploitable. For this, annotations ought to be objective and reliable. In this research, we use the method of annotation of raw corpus data, and more precisely pragmatic annotation.

Studies on discursive phenomena (Spooren & Degand 2010; Grisot 2017) pointed out that in the case of pragmatic and discursive annotations, inter-annotator agreement rates measured with chance-corrected coefficients such as the Ķ coefficient (Cohen 1960; Carletta 1996; Artstein & Poesio 2008), are frequently low. One of the methods used to increase the reliability of the data is *double coding*, which consists of a discussion of disagreements (cf. Spooren & Degand 2010; see for example Sanders & Spooren 2009, who use double coding for their analysis of two connectives indicating causality in Dutch). In this case, individual annotation strategies become cooperative strategies since double coding requires making explicit the reasoning on which the interpretation is based and convincing the other annotator of the quality of the reasoning. By means of case studies with annotation tasks, Grisot (2017) shows that data, for which low inter-annotator agreement rates were found, may be considered as reliable (defined as above). She proposes that inter-annotator agreement rates are strongly dependent on the type of linguistic information dealt with in the annotation task: purely pragmatic information, which is highly context-dependent, will result into low inter-annotator agreement rates while semantic (encoded) types of information will result into higher agreement rates. In the current research, inter-annotator agreement was calculated using both percentage and the Ķ coefficient. In both studies, we apply the double coding method: disagreements are discussed in a second phase and we carry out further analysis on the final set of agreements.

Sections 2 and 3 have made it clear that the meaning and use of the connectives *car* and *parce que* are all but a virgin field of linguistic discovery. We therefore designed a connective annotation methodology that would enable us to build on this previous research, yet without being biased in the corpus analysis. To this end, we asked external annotators who were naïve to our research questions and hypotheses to fulfil different types of annotation tasks (for a discussion of the use of naïve coders, see e.g. Bruce & Wiebe 1999; Carletta 1996; Crible & Degand 2019). The coding consists in a targeted annotation of the connectives *car* and *parce que* in context. In our two annotation studies, we used three types of annotation procedures, as given in Figure 1.

We will first briefly present them, and then, describe them in more detail with respect to how they were used in our two studies. In the **guided** annotation procedure, annotators received annotation guidelines consisting of clear definitions allowing them to distinguish among the categories to be annotated and they went through a training phase. In the **non-guided** annotation procedure, annotators did not receive specific annotation guidelines about how to distinguish between the categories of interest. Four different annotators (two pairs) participated in our studies. The first pair of annotators participated in the

[2] We should note however that certain text messages could also be considered as relatively formal. Nevertheless, the analysis of the content of the SMSs considered in our study showed that they were cases of informal and conversational style.

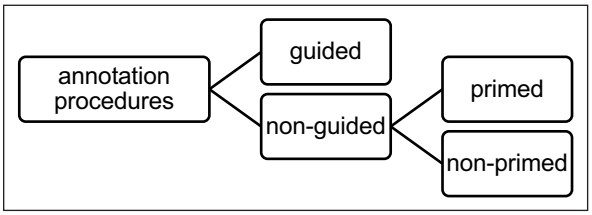

Figure 1: Annotation procedures.

guided annotation task (Study 1) and then they participated in the second annotation task (Study 2), which was a non-guided procedure. As such, their performance in the second annotation task was certainly influenced by the knowledge they acquired through their participation in Study 1. We will call this a **primed**[3] annotation procedure. Finally, two other annotators, who had not participated in Study 1, participated in the second (non-guided) annotation task. As they did not participate in Study 1, their performance was not influenced by previously acquired knowledge. We will call this a **non-primed** annotation procedure.

In sum, Annotation Study 1 focused on the fined-grained ontological classification (five categories: events, actions, opinions, emotions or speech acts). It was a guided annotation process, as each of the five categories was defined and exemplified in the annotation instructions, and the annotators received training before carrying out the annotation. Annotation Study 2 focused on an intuitive classification of subjectivity (two categories: subjective or objective). It was a non-guided annotation process, as annotators were asked to evaluate intuitively whether they considered the sentences to convey information in a more objective or more subjective manner. They did not receive any specific training for this task. Two pairs of annotators participated in Study 2: the first pair was the same as the two annotators from Study 1 (hence, it was a primed annotation task) and the second pair consisted of two new annotators with no previous knowledge (so, it was a non-primed annotation task).

4.2 Hypotheses and predictions

The theoretical analyses, together with previous studies presented in Sections 2 and 3, give rise to two hypotheses, and their subsequent predictions for empirical testing. Table 1 summarizes the predictions generated by these two hypotheses: *the hypothesis of subjective car* and *the hypothesis of high register car*.

As shown in Table 4–1, ***the hypothesis of subjective car*** yields the general prediction that *car* is better suited to subjective relations. We should find more uses of *car* expressing subjective relations than uses expressing objective relations, in both the *Le Monde* corpus and the *SMS* corpus. The second hypothesis – ***the hypothesis of high register car*** – generates the general prediction that we should find *car* used with greater frequency in the contexts of high register language. A significantly higher number of sentences with *car* should be observed in a journalistic corpus like that of *Le Monde* than in the *SMS* corpus, given that text messages are characterized as low register language. Yet here we need to consider a possible adjustment of this prediction. On the one hand, it may be that *SMS* users opt for *car* over *parce que* because *car* is a shorter word and takes less

[3] In other words, we did not specifically prime the annotators as it is done in psycholinguistic or psychological experimental setups. In our study, the term "primed" is used to describe exclusively the fact that by participating in a study, annotators acquire knowledge that they use when performing their annotations in another study.

Table 1: Summary of hypotheses and predictions in terms of corpus analysis.

Hypotheses	*SMS*	*Le Monde*
HYPOTHESIS OF SUBJECTIVE CAR	Significantly more occurrences of *car* with subjective relations than with objective relations	Significantly more occurrences of *car* with subjective relations than with objective relations
HYPOTHESIS OF HIGH REGISTER CAR	Significantly less *car* than *parce que*	Significantly more *car* than *parce que*

time to type; if this is indeed the case, we might expect to observe a much higher use of *car* at the expense of *parce que* in the *SMS* corpus. On the other hand, SMS users very often employ a whole array of abbreviations for frequently used words (*bon aprèm* "good after-noon", *bon ap* "enjoy your meal", *uni* or *univ* "university", *à+* "see you later", *qqch* "some"). One such word is also *parce que*, which could be abbreviated to *pcq, pq* or *pc que*. In other words, the fact that *car* is a shorter word than *parce que* might not, in the end, play a major role in text messaging. We will take all these possibilities into account in our analyses of the *SMS* corpus data.

4.3 Annotation Study 1

4.3.1 Annotators

Two female third-year students in Language Sciences and Speech Therapy participated in Study 1. They were native speakers of French and were paid for their work.

4.3.2 Materials

The corpus study compares two types of texts: journalistic style (with the newspaper *Le Monde*) and SMS style (with the corpus of Belgian SMS). In the *SMS* corpus, *car* is well represented although statistically less frequent than *parce que*: 1076 occurrences of *car* and 1494 occurrences of *parce que* (that is, 1299 of *parce que*, 193 of *parce qu'*, 2 of *pq*, 1 of *pc que* and 0 of *pcq*) in 51,232 SMSs (approx. 981,000 words) ($p > .05$).[4] In the *Le Monde* corpus (year 2012, approx. 7,594,000 words), *car* is significantly more frequent than *parce que*, with 2907 occurrences of *car* and 1970 occurrences of *parce que* ($p > .05$).

From these two corpora, a total of 420 excerpts were randomly selected for the study: 215 (19,442 words) from the *Le Monde* corpus (108 *parce que* and 107 *car*) and 205 (6207 words) from the *SMS* corpus (108 *parce que* and 97 *car*). All excerpts selected for the analysis contained complex sentences in which the connective linked at least two sentences. No occurrence of sentence initial *parce que* was included in our data set.

4.3.3 Annotation procedure

In this annotation task, annotators carried out a **fine-grained classification of ontological categories.** This was a guided annotation procedure as the annotators received guidelines consisting of the definition of each of the categories to be annotated. Before the actual annotation task, annotators went through a training session which took place as follows. First, we presented them the fine-grained ontological classification (as described below), with one straightforward (coined) example for each of the categories,

[4] The p value is the result of a significance test using the z-ratio score.

to be sure that annotators understood them. Second, they received a training set of data made of 17 corpus excerpts (9 occurrences of *car* and 8 occurrences of *parce que* selected from the two corpora), which they annotated in an individual manner. Third, their annotations of the training set were discussed in a group discussion with no intervention of the researcher, during which each annotator made explicit how she performed the categorization. After the training session, annotators received the actual set of data (204 occurrences of *car* and 216 occurrences of *parce que* from the two corpora) and annotated it in an individual manner.

The fine-grained ontological classification was composed of five categories: *events* (englobing both dynamic events and states) as in (5), *actions* as in (6), *opinions* as in (7), *emotions* as in (8) and *speech acts* as in (9). We asked the two annotators to determine which of the five categories was expressed by means of a given *parce que* or *car* sentence.

(5) Le bâtiment s'est écroulé **parce que/car** il y avait un tremblement de terre.
the building collaps-REFL-PC because there be-IMP an earthquake.
'The building collapsed because there was an earthquake.'

(6) Véronique a poussé Max **parce que/car** il l'a provoqué.
Véronique push-PC Max because he her provoke-PC.
'Véronique pushed Max because he provoked her.'

(7) Les voisins sont à la maison **parce que/car** les lumières sont allumées.
the neighbours be-PRES at home because the lights be-PRES on.
'The neighbours are at home because the lights are on.'

(8) Je suis triste **parce que/car** mon chat est mort.
I be-PRES sad because my cat be-PRES dead.
'I am sad because my **parce que** cat is dead.'

(9) Dépêche-toi **parce que/car** on est en retard.
hurry up you because we be-PRES late.
'Hurry up because we are late.'

Even though this way of classifying sentences with connectives is not exactly the same as Sweetser's domains of use, certain parallels can be drawn. The category of *events* is closely related to Sweetser's content domain (in turn corresponding to the objective type of causal relations), the category of *opinions* typically indicates the epistemic domain, and the category of *speech acts* indicates the domain of speech acts, both of which are directly transposable to the subjective type of causal relations. In addition, our ontological classification explicitly separates the categories of *actions* and *emotions* as a potential source of additional information. In general, human actions can be classified as sources of causal relations (cf. Davidson 1967), and thus can belong to the content domain. However, the presence of human individuals as sources of causality can create room for subjectivity, a feature intended to be captured by the scale of speaker involvement (Pander Maat & Degand 2001; Degand & Pander Maat 2003). In other words, in distinguishing the category of *actions* from that of *events*, we aimed to control the noise they might produce in the data with respect to the subjective/objective distinction. The same goes for the category of *emotions*, which may also be somewhat nuanced. *Emotive* may be seen as subjective *per se*, since their source indisputably lies in human beings. Yet if we acknowledged

Table 2: Subjectivity classification derived from ontological categories defined for fined-grained classification.

	OBJECTIVE CATEGORIES	SUBJECTIVE CATEGORIES
SUBJECTIVITY1	*events + actions*	*opinions + emotions + speech acts*
SUBJECTIVITY2	*events*	*actions + opinions + emotions + speech acts*

that *emotions* take part in ordinary causal relations, we would be forced to accept that *emotions* also belong to the content domain (i.e. objective causal relations), which is in conflict with our intuitions about their eminently subjective character. In sum, the rationale of the ontological classification constructed for our corpus study is to avoid possible interference between the ontological categories described in sentences with connectives and their eventual subjective/objective distinction, on which we will explicitly focus in the remaining classifications.

On the basis of the fined-grained ontological classification, we then established two **'derived' classifications of subjectivity**. The paradigmatic category associated with objective uses of connectives is the category of *events*. Another category which we might also consider to be linked to the notion of objectivity (albeit to a lesser degree) is *action*. So, the first derived subjectivity classification (**Subjectivity1**) comprises *events* and *actions* under the tag of objectivity; the remaining three categories (i.e. *opinions*, *emotions* and *speech acts*) are labelled as subjective. The second derived subjectivity classification (**Subjectivity2**) kept only *events* as objective, with all other members from the ontological classification annotated as subjective. Table 2 summarizes the two derived subjectivity categorizations.

4.3.4 Analysis and results

4.3.4.1 Analysis 1: fine-grained ontological classification

The inter-annotator agreement rate was 58%; for a five-category classification, this corresponds to a Ķ value of .46. In a second phase we applied the double coding method: the two annotators discussed the disagreements during a free discussion, without the presence of the experimenter, and had to refer to the annotation guidelines. After this discussion, the inter-annotator agreement rate increased to 98%. The remaining 2% (ten corpus excerpts) were judged "ambiguous", and discarded from the analysis. Hereafter, we take into account the agreements issued from phase 2 of the annotation.

Figure 2 shows the distribution of the five categories for each connective in the *Le Monde* corpus and the *SMS* corpus.

We observe that the most frequent categories in the *Le Monde* corpus, regardless of the connective used, are *opinions* and *events*, followed by *actions*. The two least frequent categories are *emotions* and *speech acts*, and this is true for both connectives. A Chi-Square[5] test showed that the two distributions are similar, as each category was expressed equally frequently with *parce que* and *car* ($\chi^2(4) = .327$, $p = .988$). The situation changes for the *SMS* corpus. We observe significant differences between *events* and *opinions*, such that *events* are expressed at a significantly higher frequency with the connective *car* than with *parce que*, whereas *opinions* are more frequent with *parce que* than *car* ($\chi^2(4) = 10.559$, $p = .339$).

[5] All chi-square tests (from Study 1 and Study 2) are Pearson's Chi-Square. They were performed using the SPSS statistical software, with the function Crosstabs, and the graphs are interpreted on the basis of the SPSS output.

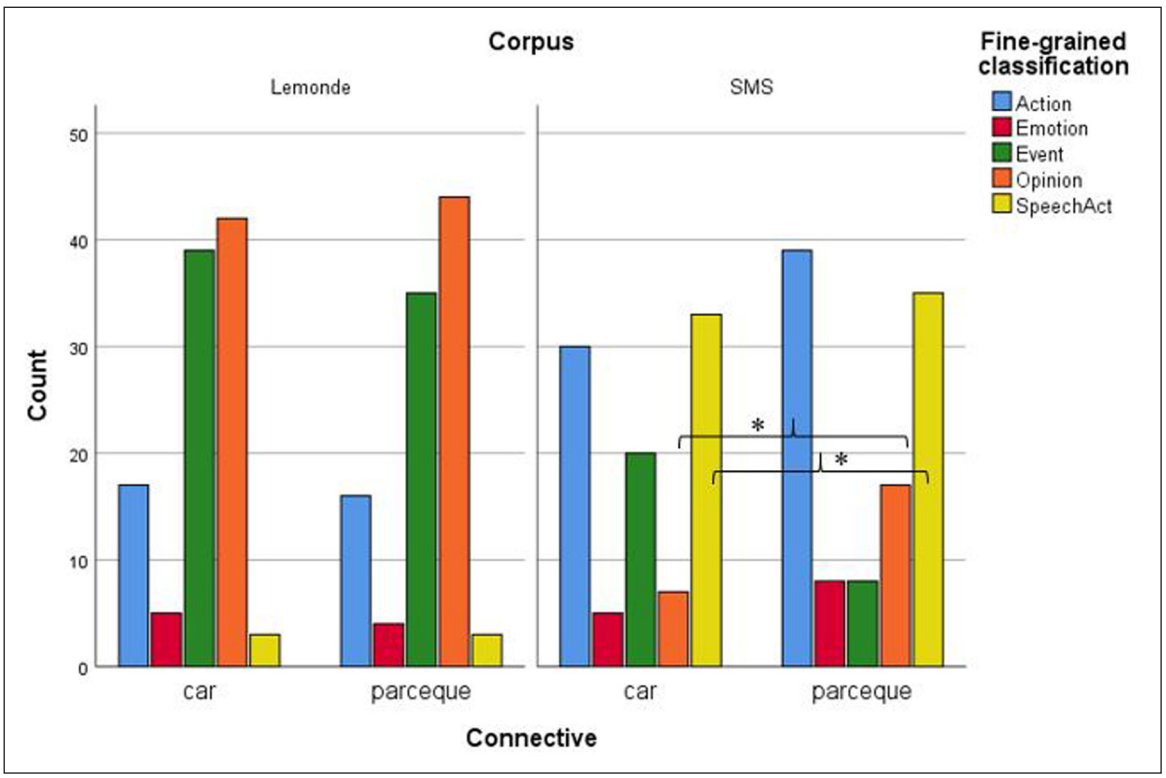

Figure 2: Connective and fine-grained ontological classification in each corpus.

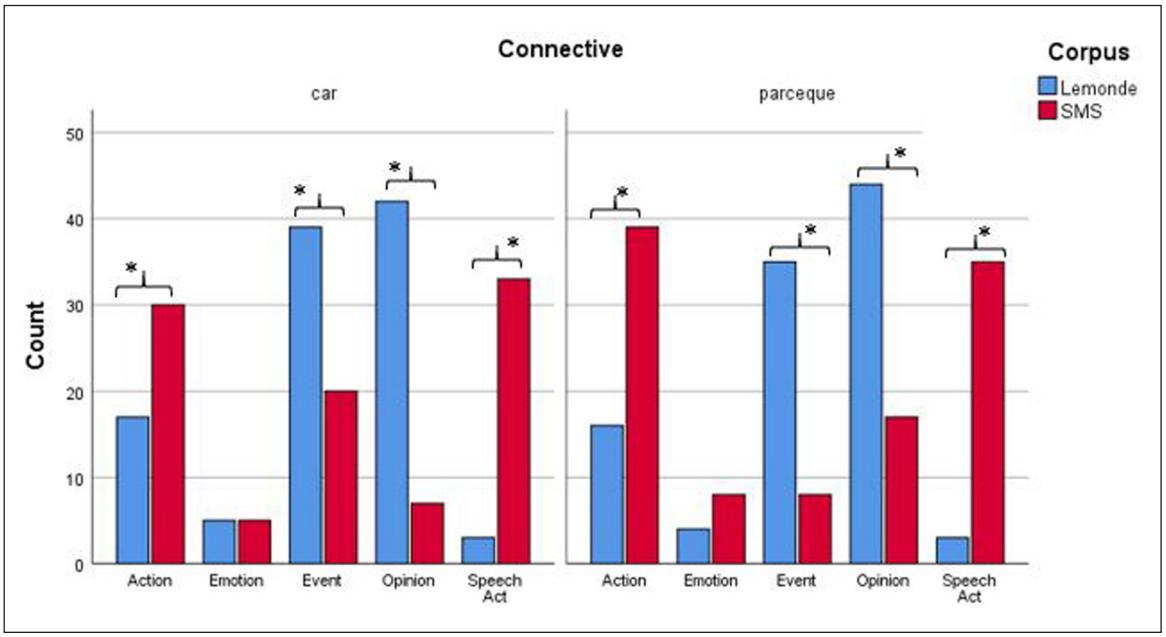

Figure 3: Connective and fine-grained ontological classification by corpus.

From a different perspective, Figure 3 presents for each category the differences between connectives across the two types of corpus.

The comparison between the two types of data shows that *car* is used more often with *actions* and *speech acts* in the *SMS* corpus than the *Le Monde* corpus, and less often to express *events* and *opinions* in the SMS corpus than the *Le Monde* corpus ($\chi^2(4) = .327$, $p = .988$). *Parce que* follows a similar trend: it is employed more often to talk about *actions* and *speech acts* in *SMS* data than in the *Le Monde* data, and less often about *events* and *opinions* in *SMS* than in the *Le Monde* corpus ($\chi^2(4) = 10.559$, $p = .032$).

4.3.4.2 Analysis 2: two subjectivity classifications derived from the fine-grained ontological classification

As shown in Figure 4, relating to the Subjectivity1 classification (*events* and *actions*: objective, *opinions*, *emotions* and *speech acts*: subjective), our data – taking both types of corpus together – showed no significant differences, whether between the two connectives or between the two relations ($\chi^2(1) = 1.401$, $p = .237$).

As shown in Figure 5, when analyzing the data in each corpus separately, no differences were found between the two connectives with respect to Subjectivity1, whether for the *Le Monde* corpus ($\chi^2(1) = .167$, $p = .683$) or for the *SMS* corpus ($\chi^2(1) = 1.528$, $p = .237$).

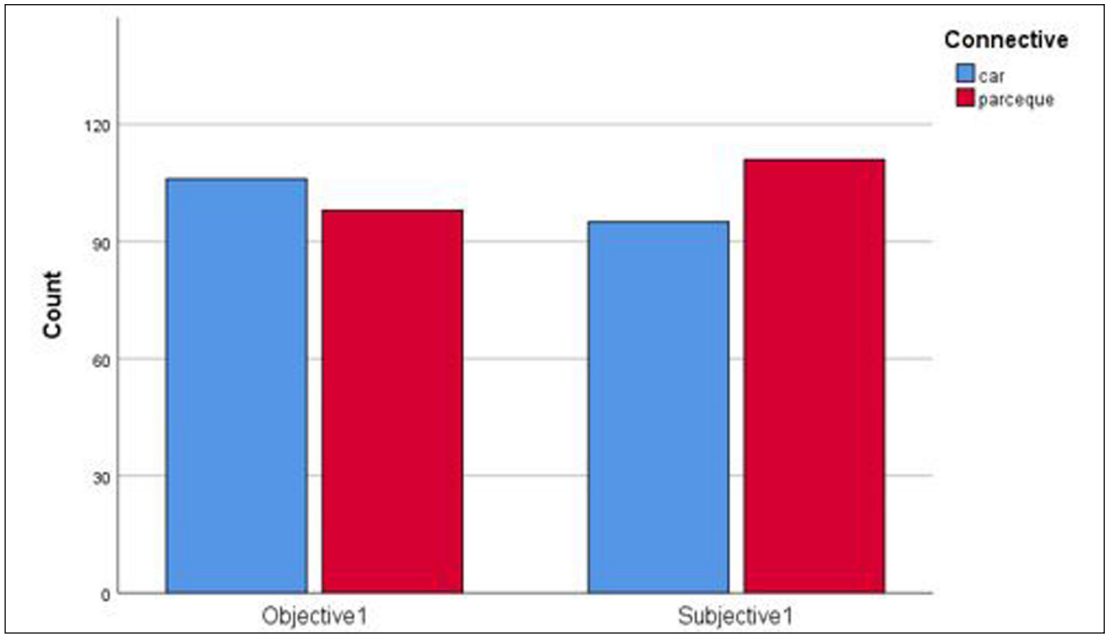

Figure 4: Connective and Subjectivity1 (derived classification).

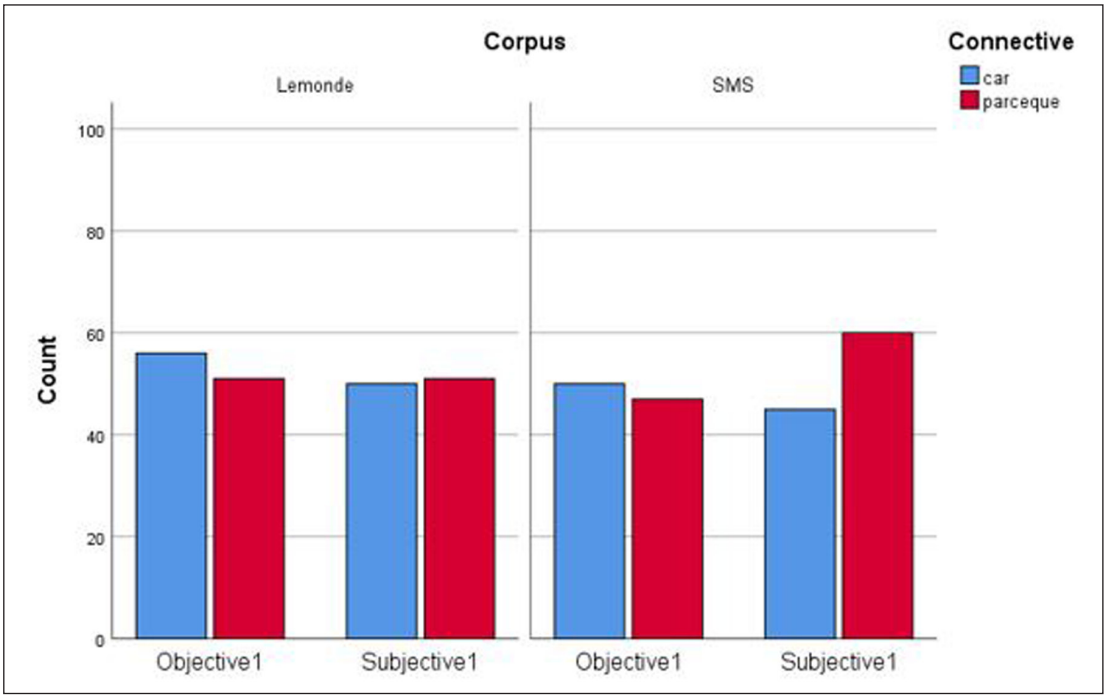

Figure 5: Connective and Subjectivity1 (derived classification) by corpus.

As shown in Figure 6 regarding the Subjectivity2 classification (*events*: objective, *actions*, *opinions*, *emotions* and *speech acts*: subjective), we observe significant overall differences between the two connectives with respect to the subjective/objective opposition. In particular, when *actions* are categorized as subjective (as is the case with Subjectivity2), both connectives are used with significantly greater frequency with subjective categories ($\chi^2(1) = 4.225$, $p = .040$). Also, the connective *car* has a significantly higher number of objective uses than *parce que* ($\chi^2(1) = 4.225$, $p = .040$). So it seems that when the ontological category of *action* is categorized as subjective, it has an impact on the categorization of *parce que* as more subjective than *car*.

As demonstrated in Figure 7, the analysis per corpus shows that the overall difference between subjective and objective categories with both connectives is significant for

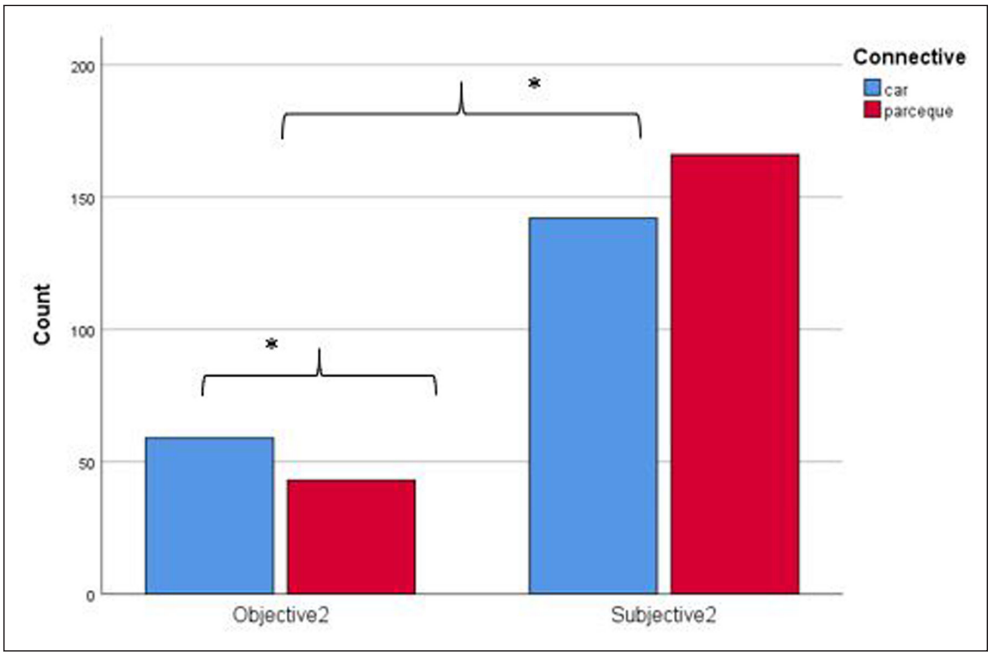

Figure 6: Connective and Subjectivity2 (derived classification).

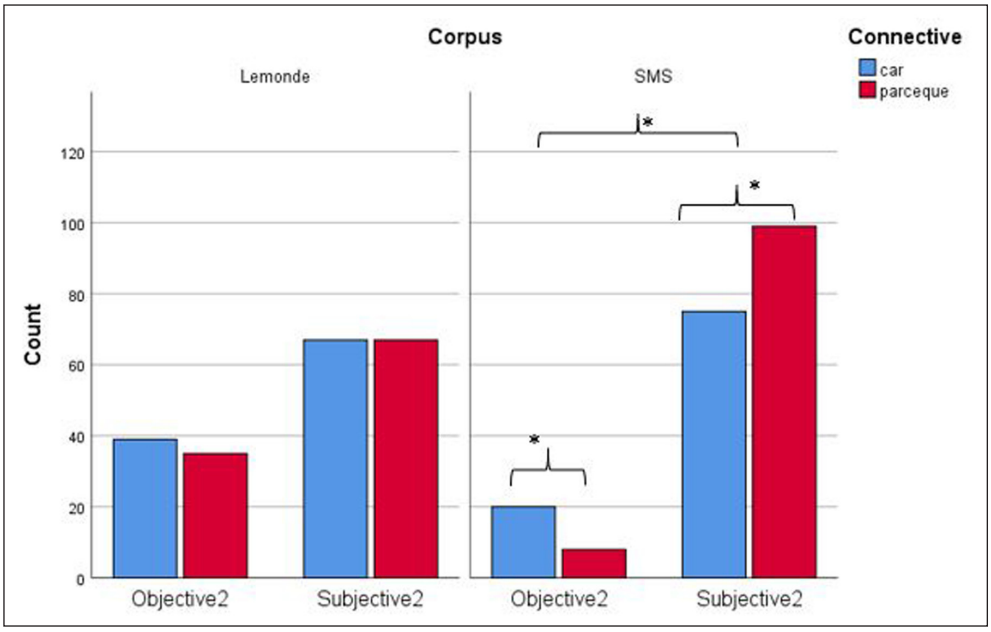

Figure 7: Connective and Subjectivity2 (derived classification) by corpus.

the *SMS* corpus ($\chi^2(1) = 7.768$, $p = .007$) but not the *Le Monde* corpus ($\chi^2(1) = .139$, $p = .709$). Besides, *car* is significantly more often objective than subjective, whereas *parce que* is significantly more often subjective than objective ($\chi^2(1) = 7.768$, $p = .007$).

4.4 Annotation Study 2

4.4.1 Annotators

In this study, we performed two types of annotation tasks: one primed and one non-primed. As mentioned in section 4.1, in the primed annotation task, the annotators were the same who participated in Study 1. As such, we assume that their performance in Study 2 was primed by their participation in Study 1. In contrast, in the non-primed annotation task, we had two new annotators, who had not made previous annotations that could have influenced their performance. All four annotators studied Language Sciences and Speech Therapy. They were all native speakers of French and were paid for their work.

4.4.2 Materials

The same material was used as in Study 1. More precisely, 420 excerpts were randomly selected for the study: 215 (19,442 words) from the *Le Monde* corpus (108 *parce que* and 107 *car*) and 205 (6207 words) from the *SMS* corpus (108 *parce que* and 97 *car*).

4.4.3 Procedure

In this study, we use a non-guided annotation procedure and we propose a novel means of classifying corpus excerpts as subjective or objective based on the annotators' intuitions of subjectivity, rather than on precise contextual features. To annotate this coarse-grained **classification of intuitive subjectivity**, the annotators were asked to judge whether they considered the sentences with *parce que* and *car* to convey information in a 'more objective' or 'more subjective' way. Importantly for the objective/subjective opposition, we did not provide the annotators with a pre-established set of features defining the notions of subjectivity and objectivity, asking them instead to evaluate sentences in a holistic and intuitive way. This procedure is based on the assumption that hearers automatically (and to some extent unconsciously) evaluate the presentation of information as more or less objective/subjective while reading or listening to a text.

4.4.4 Results of the primed annotation task

The inter-annotator agreement rate was 53%, which is close to the chance level and the Ķ was not computed. This type of inter-annotator agreement is in line with Grisot's (2017) suggestion that for highly context-dependent types of interpretation, low inter-annotator rates are expected (cf. also Spooren & Degand 2010). In a second phase, we applied the double coding method: the two annotators discussed the disagreements during a free discussion, without the presence of the experimenter, and could develop a cooperative annotation strategy. During this phase, they had to make explicit the reasoning on which their interpretation is based, convincing the other annotator of the quality of the reasoning. After this discussion, the two annotators were asked to explain how they resolved the disagreements. They reported that they developed a cooperative strategy and that the task was easier to perform in duo. Due to the double coding method, the inter-annotator agreement rate increased to 100%. Hereafter, we take into account the agreements issued from the second phase. Figure 8 shows the total number of cases classified as subjective vs. objective by connective, regardless of corpus. In total, a significantly higher number of sentences were classified as objective with *car* than with *parce que*, and more sentences were classified as subjective with *parce que* than with *car* ($\chi^2(1) = 7.935$, $p = .005$).

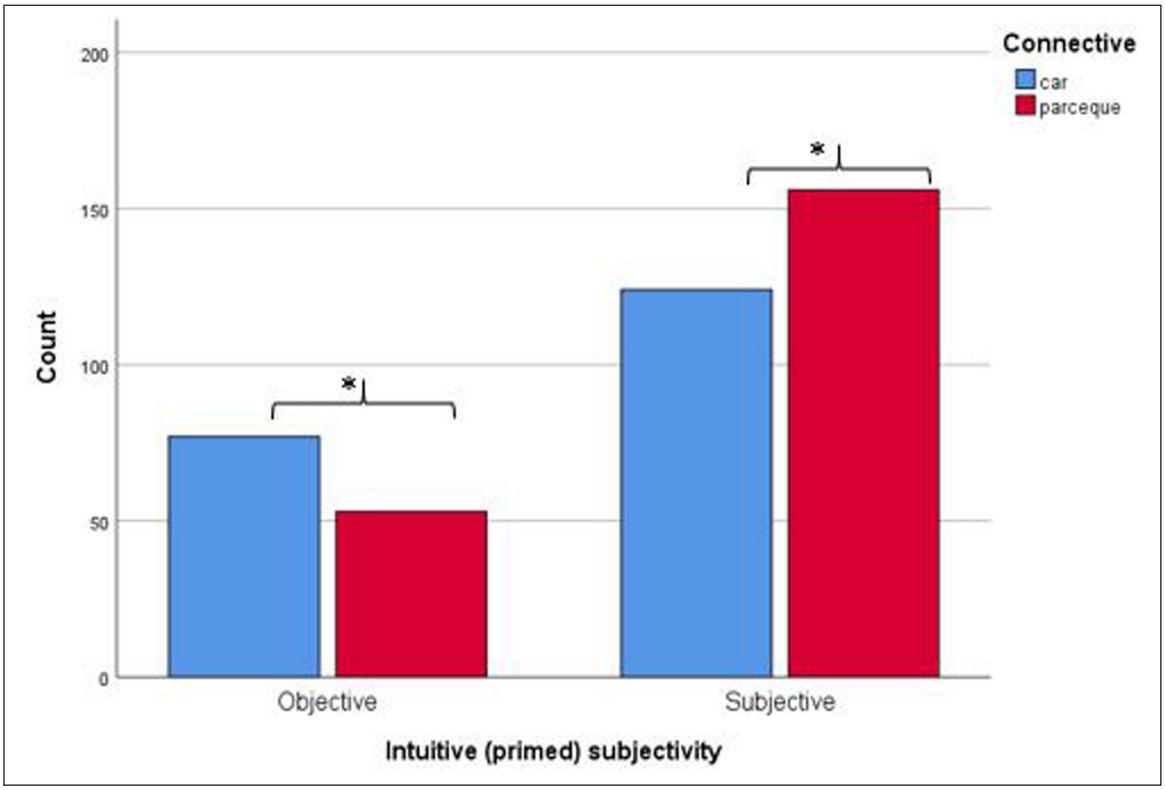

Figure 8: Connective and intuitive (primed) subjectivity.

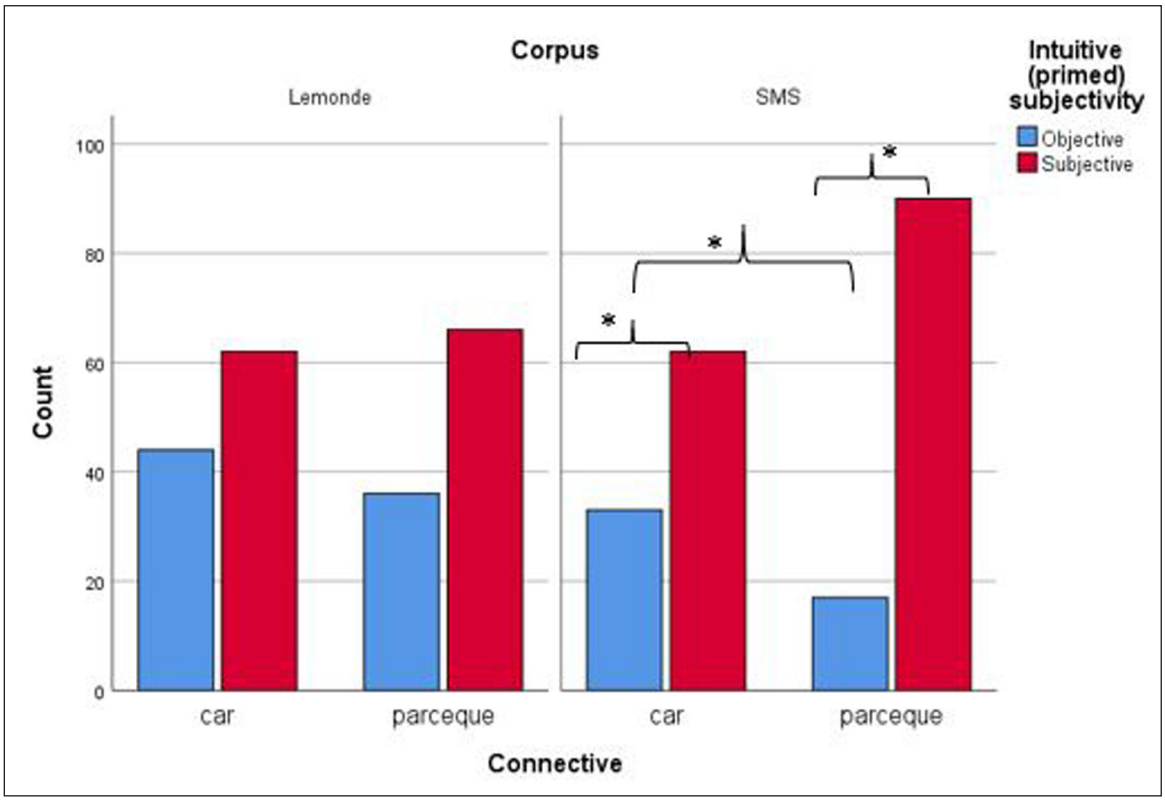

Figure 9: Connective and intuitive (primed) subjectivity by corpus.

Furthermore, Figure 9 presents the results for intuitive subjectivity, separating out the two types of corpus. As far as the *Le Monde* corpus is concerned, there are no significant differences either between subjective and objective uses of connectives (the left-hand

panel of Figure 9) or between connectives within the subjective/objective division ($\chi^2(1) = .848, p = .357$). By contrast, for the *SMS* corpus there are significant differences between subjective and objective uses of both connectives (the right-hand panel of Figure 9), as well as between both connectives within the subjective/objective division ($\chi^2(1) = 9.599, p = .002$). In particular, *car* and *parce que* in the *SMS* corpus were considered to be subjective at a significantly higher rate than they were considered to be objective. In the same corpus, *car* was also tagged as objective at a significantly higher rate than *parce que* was, while *parce que* was tagged as subjective more often than *car*.

4.4.5 Results of the non-primed annotation task

The inter-annotator agreement rate was 61%, which corresponds a Ķ value of 0.25. As in Study 1, this type of inter-annotator agreement corresponds in Grisot's (2017) scale to a highly context-dependent type of interpretation. In a second phase, we applied the double coding method: the two annotators discussed the disagreements during a free discussion, without the presence of the researcher, and could develop a cooperative annotation strategy. After this discussion, the two annotators were asked to explain how they resolved the disagreements. According to them, they interpreted as subjective when the sentence containing the connective (*car* or *parce que*) expressed an *opinion* or an *emotion*, when the verb was a mental state verb or a verb expressing an emotion, and when a person's direct discourse was reported using quotation marks. They interpreted the host sentence as objective when it expressed a fact. Due to the double coding method, the inter-annotator agreement rate increased to 100%. We took into account the agreements issued from the second phase. Figure 10 shows the total number of cases classified as subjective vs. objective by connective, regardless of the corpus. In total, there is an important, but non-significant, tendency to have a higher number of sentences classified as objective with *car* than with *parce que*, and more sentences were classified as subjective with *parce que* than with *car*, ($\chi^2(1) = 3.19, p = .074$).

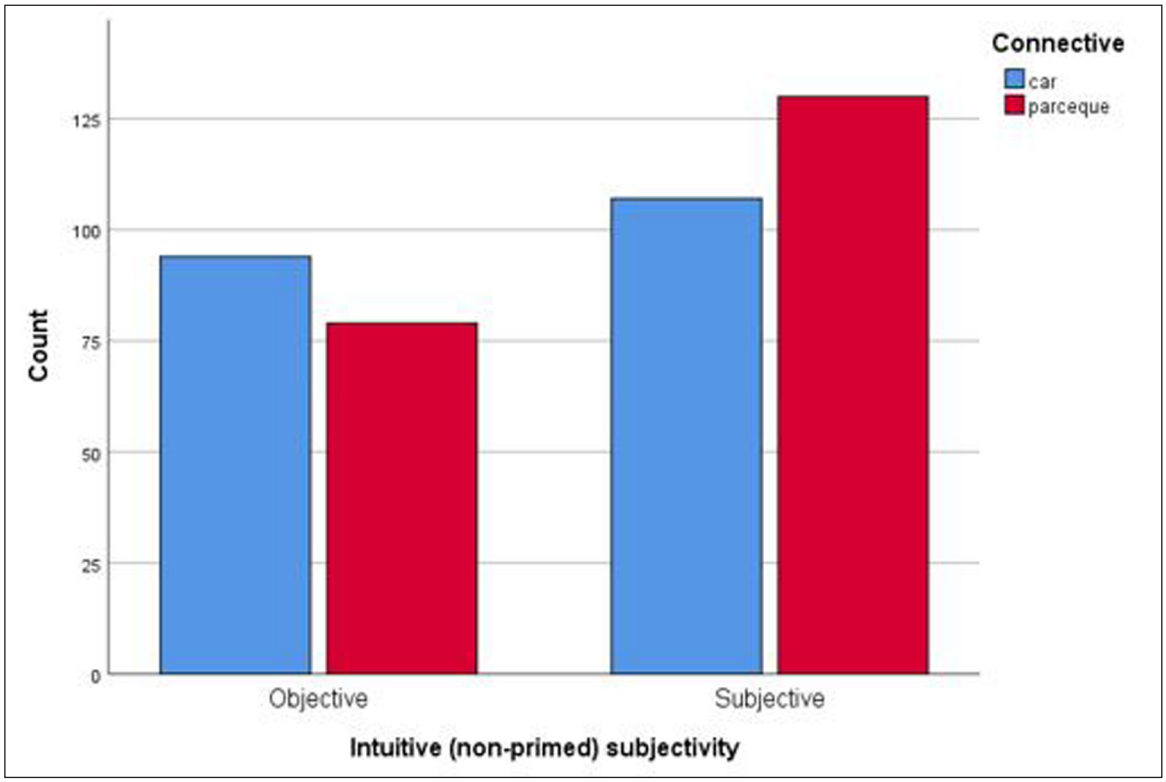

Figure 10: Connective and intuitive (non-primed) subjectivity.

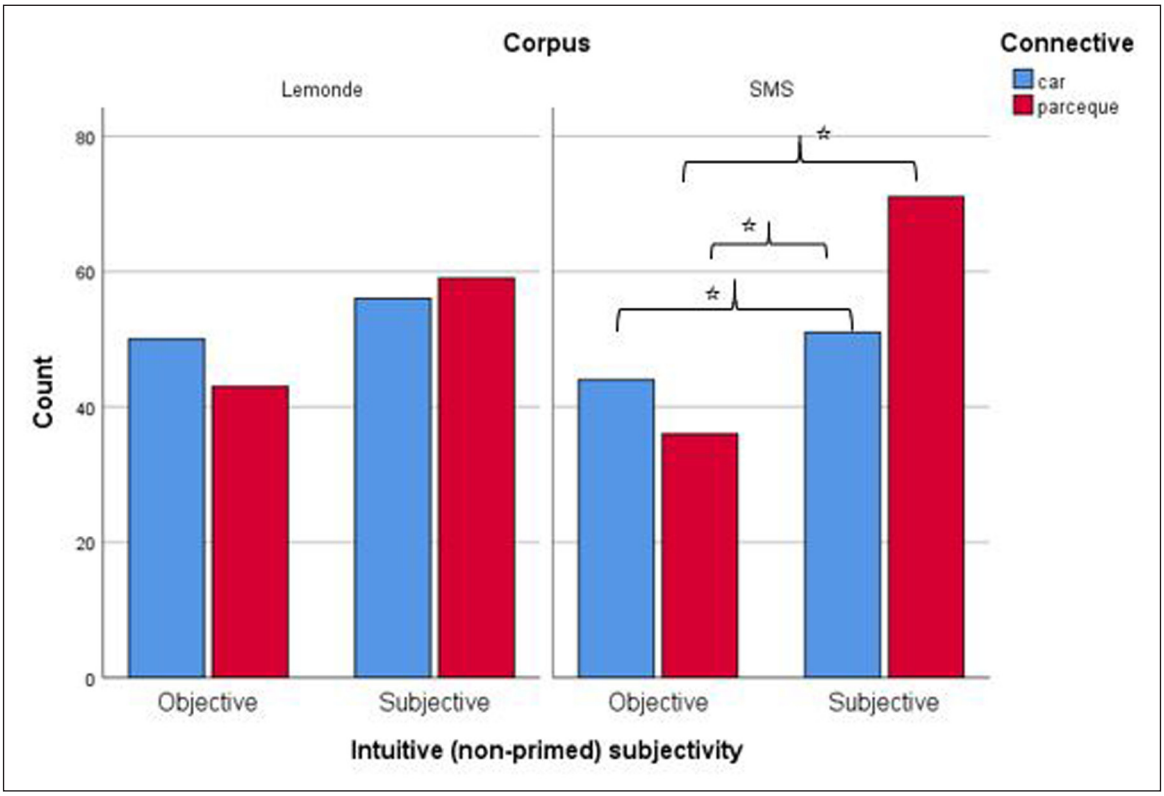

Figure 11: Connective and intuitive (non-primed) subjectivity by corpus.

Furthermore, Figure 11 presents the results for intuitive (non-primed) subjectivity, separating out the two types of corpus. As far as the *Le Monde* corpus is concerned, there are no significant differences either between subjective and objective uses of connectives or between connectives within the subjective/objective division (the left-hand panel of Figure 11) ($\chi^2(1) = 0.56$, $p = .453$). As regards the *SMS* corpus, the differences between subjective and objective uses of connectives almost reach significativity ($\chi^2(1) = 3.378$, $p = .066$). This indicates that, in the *SMS* corpus, *car* and *parce que* have the tendency to be considered as subjective more frequently than objective. In the same corpus, there was the tendency to tag *car* as objective at higher rate than *parce que* was, while *parce que* was tagged as subjective more often than *car*.

In sum, in Study 2 we found that, overall in our set of data, *car* is used more frequently to express objective relations than *parce que*. The difference was found statistically significant in the primed annotation task and it almost reached significativity in the non-primed annotation task. An analysis by type of corpus showed that the source of these differences is the *SMS* corpus. In both types of annotation task (primed and non-primed), there was no significant difference in the *Le Monde* corpus.

4.5 Further analyses

Two further analyses were carried out on the set of data: the first is the **intersection of the (primed) subjectivity classification and the fine-grained classification** from Study 1, and the second is the **intersection of the (non-primed) subjectivity classification and the fine-grained classification** from Study 1.

When the **intuitive (primed) subjectivity** classification is intersected with the fine-grained ontological classification, we obtain for each ontological category the number of occurrences classified as subjective and objective by the two annotators. Figure 12 presents the attribution of intuitive subjectivity for each category of the fine-grained ontological classification, by corpus and connective.

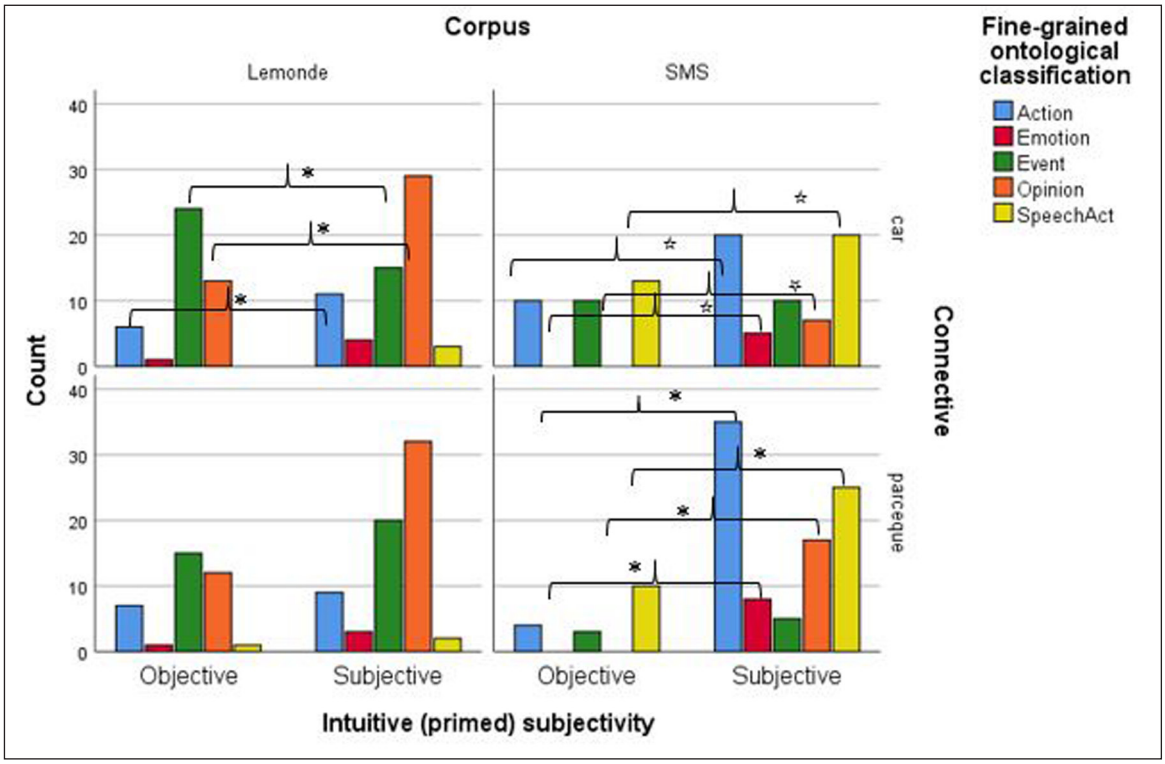

Figure 12: Intuitive (primed) subjectivity and fine-grained ontological classification by corpus and connective.

There is a significant difference in the *Le Monde* corpus, such that with *car events* were classified as more objective than subjective, *actions* and *opinions* as more subjective than objective ($\chi^2(4) = 11.724$, $p = .020$). In the *SMS* corpus, with *car actions*, *speech acts*, *opinions* and *emotions* have a tendency to be more often more subjective than objective ($\chi^2(4) = 8.784$, $p = .067$). In the *Le Monde* corpus, the differences among these five ontological categories when they are expressed with *parce que* were not found statistically significant ($\chi^2(4) = 2.808$, $p = .590$). In the *SMS* corpus, with *parce que actions*, *emotions*, *opinions* and *speech acts* were classified more subjective than objective ($\chi^2(4) = 12.657$, $p = .013$).

As above, when the **intuitive (non-primed) subjectivity** classification is intersected with the fine-grained ontological classification, we obtain for each ontological category the number of occurrences classified as subjective and objective by the two other annotators. Figure 13 indicates that the distribution of the five ontological categories with respect to subjective and objective interpretations are statistically significant in each of the two corpora and for each of the two connectives. There is a significant difference in the *Le Monde* corpus, such that with *car events* are classified as more objective than subjective, *opinions* as more subjective than objective ($\chi^2(4) = 12.321$, $p = .015$). In the *SMS* corpus, with *car actions* and *speech acts* are classified more as subjective than objective, whereas *events* are more objective than subjective ($\chi^2(4) = 16.133$, $p = .003$). No significant difference was found for *emotions* and *opinions*. In *Le Monde*, with *parce que actions* and *opinions* are classified as more subjective than objective, whereas *events* are more objective than subjective ($\chi^2(4) = 23.577$, $p = .000$). Finally, in *SMS*, with *parce que actions*, *emotions*, *opinions* and *speech acts* are classified as more subjective than objective ($\chi^2(4) = 11.829$, $p = .019$).

In sum, in this section we found similar results between intuitive primed and non-primed subjectivity when it was intersected with the fine-grained ontological classification (from Study 1). Recall that in Study 1, unexpectedly, we found that, in *SMS*, there are more

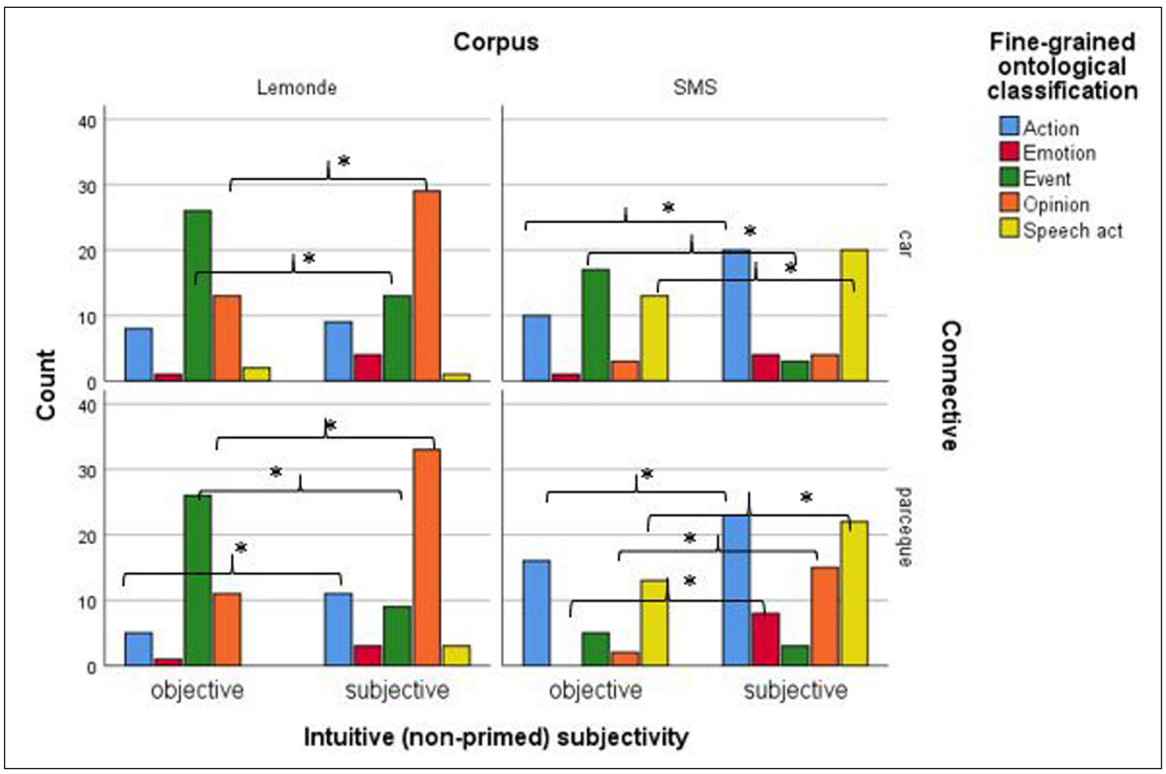

Figure 13: Intuitive (non-primed) subjectivity and fine-grained ontological classification by corpus and connective.

events with *car* than with *parce que*, and more *opinions* with *parce que* than with *car*. In the above mentioned analyses, we find that *events* are, with *car*, perceived as more frequently objective than subjective, and this in the two corpora. Regarding *parce que*, we find that in *Le Monde* objective *events* are significantly more frequent than subjective *events*; the difference is not significant in *SMS*. The opposite pattern is found for *opinions*, which are perceived as more frequently subjective than objective with *car* (in both corpora) and with *parce que* (only in *SMS*).

4.6 Discussion of results

4.6.1 Frequency data

The first important observation results from token frequency in the *SMS* corpus, where we see that *car* is frequently found although statistically less frequent than *parce que*. In Table 3, we report the results from our *SMS* corpus and compare them with the frequency results of the two connectives in written texts (novels and travel stories from the *Frantext* corpus) and in spoken data (formal and informal registers from the *Valibel* corpus) (Degand & Fagard 2012).

Table 3: Frequencies of *car* and *parce que* in two types of corpora.

Connective	Written texts (Degand & Fagard 2012) 3.7m words	*Le Monde* (our results) 7.5m words	*SMS* (our results) ~ 1m words	Spoken data (Degand & Fagard 2012) 3.9m words
car	1560 occurrences 0.43 per thousand	2907 occurrences 0.38 per thousand	1076 occurrences 1.09 per thousand	80 occurrences 0.02 per thousand
parce que	1655 occurrences 0.45 per thousand	1970 occurrences 0.25 per thousand	1494 occurrences 1.52 per thousand	13614 occurrences 3.7 per thousand

The spoken data analyzed by Degand & Fagard (2012) show a striking difference in the distribution of *car* and *parce que*. This difference levels out in written texts (novels and travel stories), which display similar distributions of the two connectives. This result stands in contrast to our journalistic written data extracted from the *Le Monde* corpus, in which *car* significantly outnumbers *parce que* ($p < .05$), where the *SMS* data again favor *parce que* ($p < .05$) (albeit with a lower proportion than in the spoken data). As we have seen, these findings contributed to the formulation of the hypothesis according to which *car* is reserved for high register language. However, our results from the *SMS* corpus do not corroborate this hypothesis. First, the fact that *car* does not have a dramatically lower frequency than *parce que* in a one million word *SMS* corpus (1.09 *car* per thousand compared to 1.52 *parce que* per thousand) indicates that *car* is widely present in everyday conversations, such as those using SMS. Note also that *car* is nearly three times as frequent in the *SMS* data as in the high register newspaper data (1.09 per thousand vs. 0.38 per thousand). A similar trend has been observed in chats (Véronis & Guimier de Neef 2006). Since neither text messaging nor chatting can be considered high register language, the fact that we found quite extensive use of *car* suggests that this connective is not, or at least not only, reserved to this specific register. If so, a question remains: to which type of content is the use of *car* specific? We will address this question in the last section of this paper.

4.6.2 Fine-grained classification data (Study 1)

The results from Study 1 regarding the fine-grained classification based on the ontological categories show differences between the two types of corpus.

In the *Le Monde* corpus, we observe that the two categories appearing most frequently regardless of the connective used are *opinions* and *events*, followed by *actions*. The two least frequent categories are *emotions* and *speech acts*, again for both connectives. Importantly, the two distributions (for *car* and for *parce que*) do not differ statistically (cf. Figure 2); thus, we can conclude that the distribution of the occurrences of the ontological categories remains similar for both connectives. This finding can simply be taken as reflecting the inherent nature of a journalistic type of corpus, in which reporting opinions or events is the most frequent activity, and talking about emotions or using speech acts is relatively infrequent, and punctuated by descriptions of actions; such a pattern is reflected in the use of both connectives, *parce que* and *car*.

Another picture emerges from the *SMS* corpus, where we found significant differences between the use of *car* and the use of *parce que*. In particular, *events* are expressed at a significantly higher frequency with the connective *car* than with *parce que*, whereas *opinions* are more frequent with *parce que* than with *car*. This finding is potentially puzzling with respect to the *hypothesis of subjective car* if we consider *events* to be prototypically related to objective uses of causal connectives and *opinions* to their subjective uses. We will come back to this result later, during the discussion of the type of subjectivity derived from the ontological categories. As for the use of the two connectives, each displays a similar pattern across the two types of corpus. Both *car* and *parce que* are used at a significantly higher frequency to talk about *actions* and *speech acts* and a significantly lower frequency for *events* and *opinions* in the *SMS* corpus than in the *Le Monde* corpus. Again, this result seems to reflect the distinct nature of each corpus: journalistic texts present more descriptions of events and opinions, whereas short messaging focuses more often on everyday content like descriptions of human actions, as well as expressing questions, orders, promises or requests conveyed by means of speech acts.

4.6.3 Derived subjectivity data (Study 1)

The results from the two derived subjectivity classifications conducted in Study 1 present certain differences. Concerning the first type, **Subjectivity1** (where objectivity was

derived from the ontological categories of *events* and *actions*, and all the other categories – *opinions*, *emotions* and *speech acts* – were taken as subjective), there were no significant differences, whether taking the data as a whole or splitting them by type of corpus. This means that this manner of merging the ontological categories does not allow us to establish any differentiation with respect to the subjective/objective use of *car* and *parce que*. Anticipating the discussion, we can say that the category of *actions* is not homogenous enough as far as the subjectivity/objectivity opposition is concerned. Regarding the second type, **Subjectivity2** (where only *events* remained objective, and all the other members of the ontological classification were counted as subjective), we found more subjective uses overall. This result might simply be due to the unequal distribution among the categories (one objective vs. four subjective); it might also reflect a predominant tendency in natural language to express subjective types of content. However, another very interesting result emerges at the general level where the data are not split by corpus: we found more objective cases expressed with *car* than with *parce que*. This is another finding which we take to indicate that the subjective nature of *car* might be not so pronounced. After division of the data by corpus, we found that there are more subjective cases in the *SMS* corpus, but no significant difference is found in the *Le Monde* corpus. The last result is probably just an echo of the nature of each corpus: the *SMS* type of text certainly conveys more subjective content than the journalistic type.

As these two ways of deriving subjectivity show, the subjective/objective opposition is not clear-cut, as it seems to be dependent on the definition formulated by researchers. In a great majority of studies examining the subjective vs. objective uses of connectives, researchers provide annotators with a set of features indicating a subjective or objective use of a connective. While this method has many advantages in terms of the reliability of the coding process, it also has certain unwanted consequences. In particular, some features – like the use of epistemic attitude markers such as *I believe that…* – are used both to classify Sweetser's types of domain of use (content, epistemic or speech act) and to decide whether we are dealing with a subjective or objective use of the connective.

4.6.4 The intuitive classification data (Study 2)

For the reasons provided above, we decided to create a classification of subjectivity based on the readers' intuitions. More precisely, two types of intuitive classification were tested as mentioned in Section 4.1. First, we performed an intuitive (primed) classification, for which the two annotators did not have access to a definition of subjectivity while doing their annotations but had participated in Study 1. We assumed that the annotators acquired knowledge about the five ontological categories. Second, we also performed an intuitive (non-primed) classification with two new annotators, who did not participate in Study 1 and who were not exposed to the five ontological categories. Before presenting the results, it is useful to recall the cooperative strategy that the non-primed annotators put in place during the discussion of their disagreements (second phase). The annotators reported that after a few excerpts they decided that they would annotate as subjective sentences expressing emotions, opinions or containing mental state verbs or reporting a direct discourse with quotation marks whereas they would count as objective sentences describing facts. As we can observe, the non-primed annotators identified spontaneously a major part of the categories we defined for the guided classification. On the objective side, they pointed to facts, which quite straightforwardly correspond to our category of *events*, and on the subjective side they identified emotions and opinions leaving aside *speech acts* and *actions*. We think that this alignment of categories explains a fair share of the homogeneity in the results we discuss below.

The findings from the intuitive, both primed and non-primed, subjectivity classification reveal several interesting points. The first is that the increase in the annotator agreement rates from phase 2 compared to phase 1 indicates that there might be a spontaneous individual perception of subjectivity/objectivity, which is less open to variation when subjectivity is dealt with in a group (i.e. the cooperative strategy, as suggested by Spooren & Degand 2010). This increase was observed in both primed and non-primed subjectivity classifications (as reported in sections 4.3 and 4.4). This process is also similar to what studies found for problem solving and argumentation: individuals reason more efficaciously when they have to approach a problem in a group rather than when they are alone (see for instance Laughlin 2011; cf. Mercier & Sperber 2017).

Second, we note more subjective uses of both connectives regardless of the corpus across both intuitive primed and non-primed subjectivity classifications. This result is not unexpected, as it confirms that the expression of a subjective perspective is widely present in language (e.g. Benveniste 1966). Furthermore, this result mirrors the findings from the Subjectivity2 classification (in which only *events* were considered as objective), where we also observed significantly higher rates of overall subjective use. However, another puzzling observation can be made: if we compare according to type of relation expressed overall (subjective vs. objective), we find that the connective *car* is perceived to convey information significantly in a more objective way than *parce que*, while *parce que* is more often understood to transmit information in a more subjective way (as found in the primed annotation; cf. Table 4). Once more, this finding is in conflict with existing assumptions about the connective *car*, which is traditionally considered to be more subjective. So, the *hypothesis of subjective car* seems to be weakened by these findings. Interestingly, in the intuitive (non-primed) classification the results almost reach statistical significance. This difference might be due to the fact that the (non-primed) annotators identified only a subset of categories as subjective (*emotions*, *opinions* but not *speech acts* or *actions*), as opposed to the primed annotators (who had acquired knowledge about all five categories through their participation in Study 1).

The analysis of the data by corpus shows a more nuanced picture. The tendencies are the same, but in the *SMS* corpus the differences are statistically significant in the primed annotation and almost reach significance in the non-primed annotation. In other words,

Table 4: Summary of the results for the two types of subjectivity classifications.

	DERIVED SUBJECTIVITY FROM GUIDED ANNOTATION		**NON-DERIVED SUBJECTIVITY FROM NON-GUIDED ANNOTATION**	
	Subjectivity1	**Subjectivity2**	**Intuitive (primed) subjectivity**	**Intuitive (non-primed) subjectivity**
Overall	no difference	more subjective uses	more subjective uses	more subjective uses (tendency to significance)
		car more objective than *parce que*	*car* more objective than *parce que*	*car* more objective than *parce que* (tendency)
By corpus	no difference	more subjective uses than objective uses in *SMS*	more subjective uses than objective uses in *SMS*	more subjective uses than objective uses in *SMS* (tendency)
		no difference in *Le Monde*	no difference in *Le Monde*	no difference in *Le Monde*
		more objective uses with *car* in *SMS*	more objective uses with *car* in *SMS* more subjective uses with *parce que*	more objective uses with *car* in *SMS* (tendency) more subjective uses with *parce que* (tendency)
		no difference in *Le Monde*	no difference in *Le Monde*	no difference in *Le Monde*

the use of *car* in text messaging is intuitively perceived as more objective than the use of *parce que*. Once more, this finding invites us to revise the two hypotheses about *car*. First, the connective is not exclusively related to high register language, since we observe it in everyday use of French in SMS. Second, it is not mainly used to express subjective content: according to our results, it is intuitively evaluated as transmitting the content in a more objective way. Table 4 provides a summary of the results we obtained, with the four classifications we used to identify subjective and objective uses.

4.6.5 Intersection of fine-grained and intuitive classification data (Studies 1 and 2)

In the next stage, we intersected the ontological classification with the intuitive (primed and non-primed) subjectivity classifications, to check the homogeneity of the ontological categories with respect to subjectivity. Specifically, the goal was to verify the proportion of cases which will be classified in an intuitive way as subjective or objective within each of our five ontological categories. Overall, we found the same type of results in the primed and non-primed annotation. The first major finding is that the ontological categories are not homogenous with respect to the speakers' intuitive evaluation of the subjectivity/objectivity dimension. Regardless of the corpus and the connective, we found cases that were classified as objective and subjective in almost every category.

First, let's consider *events*, which correspond to the traditional Sweetserian content domain and which are commonly linked to the objective use of causal connectives. The results of the non-primed annotation confirm that events are more objective than subjective, and this with both *car* and *parce que*. However, the intriguing point is that if we follow the *subjectivity hypothesis of car*, *car* should not be used to express this objective ontological category. And this result is strengthened in the primed annotation, where surprisingly events were found as more objective than subjective only with *car*. Second, in the case of *opinions*, surprisingly we found many occurrences that were classified as objective with both connectives in both corpora. Yet most traditional classifications of epistemic uses of connectives – corresponding to our category of *opinions* – are strongly associated with a subjective character. Third, regarding the category of *speech acts* – which is also traditionally strongly related to the subjective dimension of language use – we found both subjective and objective qualifications. Fourth, the only two categories which followed the expected pattern were *actions* and *emotions*. *Emotions* were almost exclusively classified as subjective, where *actions* were sometimes classified as subjective and sometimes as objective. This mitigated character of *actions* probably explains the fact that the results from our first subjectivity classification (Subjectivity1) did not reveal any difference between objective and subjective uses of connectives in any corpus.

The second major finding concerns the connective *car* and its objective uses. The two statistically significant differences indeed confirm the trend we observed in previous analyses. On the one hand, the events reported with *car* were perceived by our annotators as more objective in the *Le Monde* corpus; on the other hand, *speech acts* performed with *car* found in the *SMS* corpus were again considered more objective. These findings once more lead us to believe that the *hypothesis of subjective car* should be revised; in particular, it draws our attention to a relationship that may exist between the connective *car* and objectivity. In the next section, we will try to put forth some hypotheses about the nature of this relationship.

5 Mismatching sources of objectivity and subjectivity

According to the traditional subjective/objective causality distinction we expected to find that the category of *events* (englobing both dynamic events and states) is tightly linked to objective causal relations and the category of *opinions* to subjective causal relations.

However, it was not infrequent to find subjectivity and objectivity with both *opinions* and *events*. In what follows, we will first discuss a selection of corpus examples showing the atypical categories associations and contrasting them with the more classical cases (section 5.1) and second we will propose a theoretical account of these findings within a semantic (section 5.2) and a relevance-theoretic framework (section 5.3).

5.1 Discussion of corpus examples

Let us illustrate two unusual clusters of categories (*subjective events* and *objective opinions*) with some examples from our corpora. We will contrast them with more classical configurations such as *objective events* and *subjective opinions*. We start with the category of *events* that was annotated as objective.

(10) Environ 500 résidents des environs de la centrale nucléaire de Fukushima, évacués après l'accident de mars 2011, vont pouvoir passer cinq jours dans leur maison pour le Nouvel An. Jusqu'alors les autorisations ne dépassaient pas quelques heures, **car** le niveau de radioactivité reste supérieur au niveau annuel admissible.
'About 500 residents of the vicinity of the Fukushima nuclear power plant, evacuated after the March 2011 accident, will be able to spend five days in their houses for the New Year. Until then the authorizations had not exceeded a few hours, because the level of radioactivity remains above the permissible annual level.'

(10) is an example that was considered as expressing causal relation between events in an objective manner in the *Le Monde* corpus. The context in which the causal relation is provided concerns the situation of the residents evacuated after the Fukushima nuclear accident in March 2011 who will be able to spend five days in their houses for the New Year. The explanation provided says that up to now the authorizations to stay in the contaminated zone were no longer than a few hours because ('car') the level of radioactivity was superior to the annual admissible level. As we can observe, the speaker who provides the explanation stays at the level of raw, here scientific, facts with no reference made to her personal viewpoint about them. However, it should be stressed that the perception of a description of a given event as objective by the hearer needs not to be linked to the scientific nature of the facts described, as the two examples below taken from our *SMS* corpus attest.

(11) Changement de programme on part à 18h15 **parce que** finalement les places ne sont pas numérotées ! Ça va pour vous?
'Change of plan we leave at 6:15 PM because finally the places are not numbered! Is this ok for you?'

(12) […] Pas cours de grammaire **car** "assemblée générale" sur la grève. […]
'No grammar course because "general assembly" on the strike.'

In both cases, the SMS writer provides some simple factual information, which is not of a scientific nature. Rather, in (11), the writer informs the receivers of her SMS that what will cause their departure at 18:15 (probably for a cultural event, like a concert) is that the places will not be numbered (and they thought initially that the places would be numbered). In (12), the SMS writer informs her receiver that the classes will not be held and the cause of this cancellation is that there will be a general assembly concerning the strike.

We now turn to the puzzling fusion of the category of *events* and the category subjective with the example (13).

(13) Et c'est **parce que** nous avons construit une citoyenneté européenne que les ressortissants des pays membres de l'Union européenne ont obtenu, en 1992, un droit de vote, hélas restreint, mais un droit de vote lié à cette citoyenneté. […]
'And it is because we built a European citizenship that the nationals of the member countries of the European Union obtained, in 1992, a right to vote, unfortunately restricted, but a right to vote related to this citizenship.'

The sentence (13) reports the existence of a causal relationship between two events: citizens from the European Union got a right to vote in 1992 because we [politics] constructed a European citizenship. At this stage, we deal with a simple explanation which addresses the question of why citizens, members of the EU, obtained the possibility to vote. In addition, we can clearly observe that the speaker providing this explanation adds in her personal stance, by interjecting her viewpoint concerning the range of the vote which she judges *hélas restreint* ('unfortunately restricted'). In other words, the speaker describes a situation which is a factual description of the events and she joins her personal evaluation of the situation.[6]

Another example of an event that was annotated as *subjective* from the *Le Monde* corpus is provided below.

(14) Les nouveaux indicateurs, eux, sont prêts, et assez ambitieux, avec notamment la création d'une mesure des « atteintes à la tranquillité publique ». Mais, premier obstacle, les informaticiens du ministère ont fini par s'apercevoir qu'ils ne pourront être rétroactifs. […] De plus, l'objectif d'une mise en œuvre en janvier ne pourra être tenu, **car** le nouveau logiciel de rédaction de procédures de la police – véritable serpent de mer depuis dix ans – a encore pris du retard. Il est maintenant prévu pour fin 2014.
'The new indicators, they are ready, and quite ambitious, including the creation of a measure of "attacks on public tranquility". But, the first obstacle, the computer scientists of the ministry have come to realize that they cannot be retroactive. […] Moreover, the goal of implementation in January cannot be maintained, because the new software for writing police procedures – true *sea serpent* for ten years – is still behind schedule. It is now scheduled for late 2014.'

In (14), again, the ironic expression *serpent de mer* – which is generally used by journalists to talk about a subject that often comes back as news but never really comes true – marks the speaker's distance towards an explanation provided by the computer scientists as to why the set-up of the new software for the police would be delivered later. Indeed, the reporter of the explanation interjects that the new software is a true *serpent de mer* for ten years now, which indicates that the reporter has some doubts about a final delivery of the software.

The next example from *Le Monde* provides another type of *subjective event*.

[6] Note that in this particular case, the clefting of the *parce que* clause rules out the use of *car* for syntactic reasons.

(15) […] Lundi, sur France Info, le premier secrétaire du PS a tenté de relativiser la défaite dans « trois fiefs traditionnels de la droite ». Mais, au-delà du contexte local, la majorité n'a pas réussi à rassembler. « La mobilisation a été beaucoup plus forte à droite qu'à gauche **parce que** le peuple de gauche regarde, attend, observe ce que fait le gouvernement », a estimé, sur Europe 1, le poids lourd du PS Jean-Christophe Cambadélis.
'Monday, on France Info, the first secretary of the PS tried to relativize the defeat in "three traditional strongholds of the right". But, beyond the local context, the majority failed to gather. "The mobilization has been much stronger on the right than on the left because the people of the left watch, wait, observe what the government does," said, on Europe 1, the heavyweight of PS Jean-Christophe Cambadélis.'

Example (15) constitutes an answer provided by a leader of the French Socialist Party (*PS*) to a question concerning the low mobilization of voters of the left compared to the voters of the right. In this context the answer provided by the politician is considered to be subjective first because the answer itself is elusive and the source of this answer, the politician himself, is biased as one of the most important leaders (*poids lourd*) of the Socialist Party.

To sum up, when the emergence of a personal viewpoint evaluating the situation described pops up while the speaker is establishing a causal relation between these events, the explanation provided is considered as subjective. The source of this subjectivity might be the use of an expression marking speaker's distance or it might be the speaker himself who is not considered as a reliable person in a given situation.

Now, let us turn to *opinions*, which were annotated sometimes as subjective and sometimes as objective. Here are two examples of *subjective opinions* from the *Le Monde* corpus.

(16) Mais la réponse est aussi, hélas, négative, **parce qu'**une politique monétaire plus agressive confirmerait les craintes allemandes de voir la BCE se transformer en Banca d'Italia.
'But the answer is, unfortunately, negative, because a more aggressive monetary policy would confirm the German fears of the ECB becoming Banca d'Italia.'

(17) D'autres membres de ce conseil estiment au contraire que le « plus d'austérité » n'est pas d'actualité, **car** il risquerait de gripper encore davantage la croissance – de 2013 à 2014, les mesures d'austérité coûteront déjà 1,4 point de croissance. […]
'Other members of the board on the contrary believe that the "more austerity" is not relevant, because it could further squeeze growth – from 2013 to 2014, the austerity measures will cost already 1.4 growth point.'

In example (16), the speaker provides a justification for an answer which was negative, and independently of the fact that *opinions* can be perceived as subjective per se, in this example, there is another linguistic element which adds to the flavor of subjectivity, namely, the interjection *hélas* ('unfortunately'). By using this expression, the speaker provides in addition her depreciatory evaluation concerning this negative answer. The opinion provided in example (17) appears subjective because of its non-unanimity status. It is reported that 'other members of the board judge on the contrary that…' which presupposes that the remaining members of the board have the opposite opinion. This is a typical example of a case in which there is no consensus opinion concerning some topic,

and therefore, it is difficult to acknowledge that the opinion of one part of the members of the board is more objective than the opinion of the other part. We find a confirmation of this observation in the example below, which reports an opinion that was annotated this time as objective.

(18) Depuis 2007, quelque 210 000 entreprises – des PME dans leur immense majorité – ont fermé leurs portes, dont 47 000 au cours des douze derniers mois. Si la révolte collective n'est pas d'actualité, c'est, de l'avis général, **parce que** le secteur informel représente toujours un quart de l'activité du pays […].
'Since 2007, some 210 000 enterprises – the vast majority of them being SMEs – have closed, including 47 000 in the last twelve months. If the collective revolt is not of actuality, it is, in the general opinion, because the informal sector still represents a quarter of the country's activity.'

In (18), we note the presence of the expression *de l'avis général* ('according to a general opinion') implicating that a general consensus exists, therefore, a given opinion can be considered as objective. Another example of an objective opinion is provided below.

(19) Un avocat, spécialiste du droit des affaires, considère que la loi organise le défaut d'impartialité des tribunaux de commerce **car** ceux-ci sont composés de juges non professionnels. […]
'A lawyer, specialist in business law, considers that the law organizes the lack of impartiality of commercial courts because they are composed of non-professional judges.'

In (19), what makes the reported opinion more objective is that the person who is the source of this opinion, here a lawyer, is presented as a specialist of the topic under discussion. Therefore, we can conclude that an opinion would be perceived as more objective if it represents a general consensus or if it is provided by a specialist of the subject.

5.2 Kinds of subjectivity related to descriptive and expressive use of language

In this section, our goal is to explicate the kind of 'objectivity' associated with the use of *car* in our study drawing on semantic approaches to language. For this, we suggest to distinguish between a traditional approach of subjectivity related to causal relations, which is truth-conditional and applies to the descriptive content of language, and a second approach related to the expressive use of language, which is use-conditional (Gutzmann 2015) and applies to the manner in which speakers convey the content of their utterances, as we saw in the corpus examples (13)–(17). We will call the former *d_subjectivity* (for *descriptive* subjectivity) and the latter *e_subjectivity* (for *expressive* subjectivity, and more precisely, *evaluative* subjectivity in the case of causal connectives).

An idea we would like to put forth is that the subjective/objective distinction perceived by the annotators in the non-guided annotation task (the intuitive classification) does not apply to bare causal relations, but instead to more complex language structures containing them – namely, explanations and justifications. In other words, we think that we can better account for our results if we adopt the claim that sentences with causal connectives refer to explanations and justifications, rather than to mere causal relations. A similar claim has been made by Beebee (2004), based on considerations of the interaction of negation and descriptions of events in causal situations (see also Nølke (1995), who appeals to explanation and justification to analyze the causal connectives). This solution

might, prima facie, appear to be only cosmetic and terminological; however, as we will see, it leads to a clarifying generalization.

First let us analyze a simple coined example of a causal relation: the relation whose relata are events (more broadly, eventualities), as in example (20)a. One possible analysis would simply state that the connective *because* signals the presence of a causal relation between the two events, as is formally expressed in the neo-Davidsonian style of formalization in (20)b, which successfully captures a primary intuition we have about the meaning of (20)a, that is, there exist (∃) two events, *e* and *e'*, and the event *e'* caused the event *e* (for the causal relation, we use the operator CAUSE as defined in Dowty 1979).

(20) a. Caesar died because Brutus stabbed him.
b. ∃e ∃e' [e = Caesar died, e' = Brutus stabbed Caesar ∧ CAUSE (e', e)]

However, there are cases like (21)a where a simple application of an operator CAUSE does not seem to be enough to capture all the components of meaning, as in (21)b. It is obvious that we are still dealing with events and the causal relations between them. Yet in parallel, there is the speaker's point of view, resurfacing in the expressive *nasty*, and we have a strong intuition that this particular viewpoint of the action performed by Brutus is not to be considered as an integral part of the causal relation. Thus, we need to add the speaker *s* and the operator *BEL* which refers to the speaker's epistemic attitude of believing. This is made explicit in (21)c.

(21) a. Caesar died because the nasty Brutus stabbed him.
b. ?? ∃e ∃e' [e = Caesar died, e' = Brutus stabbed Caesar ∧ Agent of e' was nasty ∧ CAUSE (e', e)]
c. ∃e ∃e' ∃s [e = Caesar died, e' = Brutus stabbed Caesar ∧ CAUSE (e', e) ∧ BEL (s, Agent of e' was nasty)]

This simple example serves also to illustrate where the flavor of e_subjectivity with the category of *events* may arise, namely, from the manner in which situations are reported. This is why we suggest that sentences with causal connectives do not indicate bare causal relations, but convey causal explanations which – even if they appeal to causal relations – have more layers of meaning upon which the e_subjective/objective distinction may hinge.[7]

Concerning the category of *opinions*, we propose a semantic ascent similar to that for the category of *events*. First, the basic relation involved is not a relation of *cause* linking events but the relation of *reason* linking propositions and the speaker holding them as true or probably true. The reasons for holding that a given proposition is true are of a different nature, and are usually thought of in terms of evidentiality or epistemicity. An epistemic relation of *reason* may pertain to various types of evidence which speakers can gather to ground their beliefs, claims or knowledge.[8] What is important here is that we can see the relation of justification that indicates the relation of reason, analogous to the relation of

[7] This is to some extent a reformulation of Pander Maat & Degand's (2001: 230) constraint on the use of causal connectives that "[a] connective encodes a certain speaker-involvement level, which it contributes to the interpretation of its discourse environment. When this level is too low or too high to be combined with the level allowed for by the discourse environment, the use of the connective is inappropriate". In other words, the use of the connective needs to be compatible with its discourse environment.

[8] The subject related to epistemic relations (i.e. relations between a speaker *s*, a proposition *p* and a reason *r* that *s* has to hold *p*) has been under investigation since at least von Wright 1951 and Hintikka 1962, and has given rise to various formal models in epistemic logic. There is also an abundance of linguistic literature which identifies various types of evidence (visual, non-visual, inference, hearsay, etc.) which then can serve as a reason for speakers' opinions (see Aikhenvald 2004 for an overview). We will not expand on these notions in this paper.

explanation of events referring to a more basic causal relation. Like an explanation, a justification can also be perceived and evaluated by the hearer as more or less e_subjective.

We have so far associated *events* with explanations and *opinions* with justifications, in the sense that *events* are subject to explanations and *opinions* are subject to justifications (see Blochowiak 2014 for more details). When it comes to the remaining categories, the puzzling point is that emotions, actions and speech acts (to the extent that they can be seen as a sub-category of actions) can be both explained and justified. These form a grey area which philosophers have painstakingly tried to illuminate (starting with Davidson 1980) and to which laymen pay little attention. What is important for a linguist working on causal connectives is that some types of sentences with connectives including these categories can be ambiguous between explanations and justifications – namely, those which talk about actions, emotions or speech acts. In the case of explanations, the speaker would just explain why a given action or speech act was performed, or why a given emotion was experienced by herself or somebody else, whereas in the case of justification, the speaker would add another dimension (such as an axiological or a bouletic dimension). This topic will not be developed here: what is important for the subject of causal connectives is the fact that some of the ontological categories can be both explained and justified, whereas others can only be explained (*events*) or justified (*opinions*).

In essence, our proposal is that a communicative act of explaining or justifying is constructed around (at least) two axes or dimensions: (i) the ontological dimension, referring to the object of an explanation or justification (i.e. *what* we are talking about, that is, the categories of our ontological classification); and (ii) the evaluative dimension indicating how an explanation or justification is provided (i.e. *how* we talk about what we are talking about, that is, the e_subjective/objective dimension).

5.3 A relevance-theoretic account of subjectivity

Now we will determine how all these elements can be systematized within a pragmatic framework of language comprehension, such as Relevance Theory. However, it should be stressed that other theoretical frameworks could also accommodate these ideas. In a nutshell, according to Relevance Theory, an utterance expresses a proposition whose basic explicature (the full-fledged propositional form of the utterance) is susceptible of being true or false. In addition, it can convey higher order explicatures, as its propositional force. Moreover, there is some implicitly communicated content which can be added, such as various kinds of implicature. In recent developments, relevance-theoretic theorists have identified other types of effect which are notoriously difficult to pinpoint: the non-propositional effects (Wilson & Carston 2019). Their main characteristics are that they are difficult to paraphrase (with different addressees coming up with different paraphrases), they are open-ended, and they often activate different perceptual, sensimotor or perceptual mechanisms. In example (22), borrowed from Wilson & Carston (2019), by answering "I enjoyed some of it," and knowing that Jack had put a lot of effort into cooking a special meal to surprise her on her birthday, Sue might communicate a wide array of non-propositional effects about her attitude not only to the meal but to her relationship with Jack.

(22) a. Jack: Did you enjoy the meal?
b. Sue: I enjoyed some of it.

Interestingly enough, expressives (such as the expression *nasty* in (21)a) which are certainly relevant to subjective/objective distinction have been analyzed in terms of non-propositional effects (Wharton 2016). A closer look at these provides our proposal with some clues about the expressive-subjective/objective distinction of the *evaluative* kind.

Usually, it is claimed that expressives have a special kind of status, and various approaches have sought to characterize this phenomenon. For instance, Kaplan (1999) proposes that expressives have a special meaning because of their non-descriptive content (see also Potts 2007; Gutzmann 2015; 2019). In Relevance Theory, expressives are analyzed as typical expressions leading to non-propositional effects and exhibiting three main features: they are independent of the proposition expressed, and so have non-truth-conditional content; they are descriptively ineffable; and they present interesting parallels with non-verbal behaviors in general (Wharton 2016). In Gutzmann's (2015) *hybrid semantics* approach, expressives have use-conditional meaning often in addition to descriptive meaning.

We propose that the e_subjectivity/objectivity dimension of language is an evaluative property[9] which (i) emerges from various factors, which are mainly non-propositional effects and (ii) applies at the level of the utterance. This means that e_subjectivity is not necessarily in itself a non-propositional effect, but is built on the basis of non-propositional effects, triggered for example by expressives. In the case of causal connectives, the proposition expressed by an utterance can contain a causal relation, but whether a given utterance (for instance, a causal explanation) is perceived as e_subjective or e_objective does not enter into the calculation of the truth-conditions of the proposition expressed by that utterance. This is illustrated in (21)c, where the subjective viewpoint of the speaker (e_subjectivity) is treated as a part of the utterance meaning distinct from the causal relation itself. Example (23) below serves to illustrate this proposal further: (23)A articulates a question under discussion (which can be stated overtly or left implicit), and (23)B gives an answer to it.

(23) A: Why did Caesar die?
B: Caesar died because the nasty Brutus stabbed him.

Basic explicature:	CAUSE (Brutus stabbed Caesar, Caesar died)
Higher order explicature:	B EXPLAINED to A why Caesar died
Non-propositional effects:	B thinks that Brutus was nasty
	B thinks that the action performed by Brutus was unfair
	B has some positive sentiment towards Caesar
	…

The proposition expressed (basic explicature) has a property of being true/false. The utterance of explanation (higher order explicature) has a property of being e_subjective to a degree *d* which is determined by the non-propositional effects.[11]

We focused here on expressives as one possible factor contributing to the evaluation of a given utterance as more or less e_subjective. Various factors can certainly affect the perception of e_subjectivity. Among these are not only traditional features such as the use of modals or evidentials, but also other more elusive and often linguistically non-

[9] We follow here Kerbrat-Orecchioni's (1999) distinction between *deictic subjectivity* (expressed through a broad class of deictic elements, such as personal pronouns, demonstratives, verbal tenses, grammatical aspect and temporal adverbs, among others) and *affective-evaluative subjectivity* (referring to "the individual usages of the common [language] code" and mainly expressed through the evaluative and affective lexicon) (p. 80).

[10] The distinction between the content of basic explicature and higher-order explicatures could also be couched in Gutzmann's hybrid semantics framework, where causal relations would constitute the descriptive truth-conditional part of the meaning whereas explanations/justifications would correspond to the expressive use-conditional part of the meaning. In other words, the e_subjective/objective distinction we try to pin down here is a non-truth-conditional (or use-conditional in Gutzmann's terms) property of utterances or communicative acts in general, realized as explanations or justifications in the case of causal connectives.

marked elements, such as the hearer's trust in the speaker's competence and benevolence (cf. epistemic vigilance mechanisms, Sperber et al., 2010), considerations of the speaker's and hearer's backgrounds, shared knowledge and shared practices, prosody, gestures, and others beside. The aim here is not to list exhaustively the elements which can impact the perception of a given utterance or communicative act as e_subjective or e_objective, but to indicate that they are quite numerous, come from various sources, and may vary from one person to another.[11]

Finally, it should be stressed that our proposal has many further-reaching consequences, as it is applicable to other types of more complex language or discourse constructions. For instance, we can think of arguments which can be valid or non-valid (analogously to propositions being true or false), and which could also be evaluated by the addressees as more or less persuasive, convincing and also subjective. Like e_subjectivity, persuasiveness or convincingness can be considered non-truth-conditional (or use-conditional) properties of arguments emerging from the non-propositional effects accompanying them. In addition, the factors contributing to these non-propositional effects will certainly overlap these different properties, a topic we will leave for future development.

6 Conclusion

In this paper, we presented a corpus study with annotation tasks with the aim of verifying two widely attested hypotheses concerning the meaning and use of *parce que* and *car*: *the hypothesis of subjective car* and *the hypothesis of high register car*. The results of our investigation do not adhere to the prevailing hypotheses issued from the literature. The findings from our corpus-based study consistently indicate that the French connective *car* is not used to predominantly express subjectivity, but that it is in fact employed to convey the message in a more objective way. This finding is not isolated. To the best of our knowledge, two other studies point in the same direction. On the one hand, our results echo Nazarenko's (2000) study, whose findings report a number of cases of objective *car* in contemporary written French. On the other hand, results from recent experimental studies have also raised doubts about the subjective character of *car*. Zufferey et al. (2018) conclude: "In French 'car' is not strongly associated with subjective relations, which implies that French-speaking readers do not use 'car' to infer the presence of a subjective relation during reading, an observation that is in line with the fact that in corpus data 'car' is not strongly associated with subjective relations, and that participants do not have a strong tendency to choose it for subjective relations." (Zufferey et al. 2018: 100) An alternative hypothesis proposed by Zufferey et al. (2018) is that the two connectives are used for different register purposes, with *car* particularly associated with high register language. Even though this hypothesis has been confirmed in their study, it probably does not provide a full explanation of *car*. Indeed, we have shown in our corpus study that *car* is used quite often in *SMS*, and Véronis & Guimier de Neef (2006) also find frequent uses of *car* in chats. Since these two modes of communications are typically associated with low register language, *car* cannot be restricted to high register style.

In an original manner, we performed two types of annotation tasks: one guided (i.e. with annotation guidelines) in Study 1 and one non-guided (i.e. without annotation guidelines) in Study 2. In Study 1, on the basis of the fine-grained classification of ontological categories, we derived two types of subjectivity: Subjectivity1 and Subjectivity2. In Study 2, we

[11] There remains an open question about the role of the connective itself: is the connective a trigger of a subjective or of an objective interpretation of an utterance, or do connectives of certain types appear more frequently in subjective or in objective contexts. As pointed out by one of the anonymous reviewers, our annotations seem to support the latter option. However, we believe further investigations are required to specifically address this question.

annotated subjectivity as language comprehenders approach it in an intuitive way, and we distinguished between the intuitive primed and non-primed types of subjectivity (cf. section 4.1). Subjectivity2 and intuitive (primed and non-primed) subjectivity resulted in the same significant finding that *car* was used to express more objective relations. The crucial point was to observe that the results of the derived and the intuitive classifications of subjectivity did not align, they overlapped only partially. We took this as indicating that the two types of classification (derived vs. intuitive) revealed two types of subjectivity. The derived classifications (Subjectivity1 and Subjectivity2) were based on ontological categories, such as *events*, *actions* or *opinions*, and thus echoed the more traditional views on subjectivity, which relates the subjective/objective divide to the type of causal relations involved (e.g. objective causality is typically related to *events* whereas subjective causality to *opinions*). We called this descriptive subjectivity (d_subjectivity). The intuitive (primed and non-primed) subjectivity was based on the spontaneous understanding of the subjective/objective division by the comprehender, which turned out to target the evaluative dimension of the use of *car* and *parce que*. We called this evaluative subjectivity (e_subjectivity).

To capture the difference between d_subjectivity and e_subjectivity, we put forward a relevance-theoretic account of these phenomena. Firstly, we proposed that the sentences with causal connectives be considered to indicate explanations or justifications (see also Nølke 1995) rather than to denote bare causal relations (see Beebee 2004 for a similar claim proposed for a different, independent reason). As a consequence, in the specific case of connectives, the property of e_subjectivity modifies the speaker's explanation or justification formulated with a connective, whereas the causal relation itself can be considered as pertaining to d_subjectivity.

Secondly, we proposed that the e_subjective/objective dimension be seen as an evaluative property of speaker's utterance(s) (see Kerbrat-Orecchioni 1999 for distinction between *deictic* and *affective-evaluative* subjectivity). This evaluative property hinges on a series of factors originating in various sources, such as the speaker's use of certain linguistic expressions (e.g. expressives, modals and other expressions indicating speaker's involvement), the speaker's use of non-linguistic cues (e.g. prosody, gestures and mimicry), considerations of interlocutors' epistemic vigilance (Sperber et al. 2010), background, shared knowledge, and so forth.

These factors are intended to feed non-truth-conditional properties of speaker's utterances, such as e_subjectivity, which arise mostly from the non-propositional effects advocated by relevance-theorists (cf. Wharton 2016 for expressives; Wilson & Carston 2019 for a general approach). An utterance with a causal connective can be broken down as follows: its basic explicature is constituted by a proposition containing a basic relation (*cause* or *reason*), and its higher order explicature refers to a higher order relation (explanation for *cause* and justification for *reason*). The non-propositional effects issued from a variety of factors make the hearer perceive and evaluate the speaker's explanation or justification as having the property of being more or less e_subjective. In other words, the evaluative notion of subjectivity proposed here is the property which applies to speaker's utterances at the level of higher order explicatures. Importantly, a given explanation of an event can be considered as objective on a descriptive truth-conditional level (causal relations between external events are considered d_objective) whereas on the expressive use-conditional level (e.g. explanations) it can be considered as e_subjective.

To conclude, the results of our investigation concerning the difference between the two French causal connectives *parce que* and *car* did not confirm the traditional thesis according to which *car* is more subjective, neither in terms of d_subjectivity (see also Zufferey et al.

2018 for similar result) nor in terms of e_subjectivity. In the light of our theoretical proposal, we interpret this result as indicating that the use of the connective *car* contributes to the perception that the explanations or justifications in which it appears are more e_objective, in the sense of the evaluative kind of subjective/objective distinction we put forth in this paper.

Abbreviations

IMP = Imperfective tense; PRES = Present tense; PC = Compound past; REFL = Reflexive

Acknowledgements

We would like to thank the three anonymous reviewers for very fruitful comments on the previous versions of this paper. All remaining errors are ours.

Funding Information

The present study was a part of the Move-in Louvain project co-funded by the Marie Curie Actions awarded to Joanna Blochowiak and conducted at the Université catholique de Louvain (promotor Liesbeth Degand), and of the project VTS (n° 100015_170008/1) funded by the Swiss National Science Foundation and carried out at the University of Geneva by Cristina Grisot (PI Jacques Moeschler).

Competing Interests

The authors have no competing interests to declare.

References

Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford: Oxford University Press.

Anscombre, Jean-Claude & Oswald Ducrot. 1983. *L'argumentation dans la langue*. Bruxelles: Editions Mardaga.

Artstein, Ron & Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics* 34(4). 555–596. DOI: https://doi.org/10.1162/coli.07-034-R2

Beebee, Helen. 2004. Causing and nothingness. In L. A. Paul, E. J. Hall, & J. Collins (eds.), *Counterfactuals*, 291–308. Cambridge MA: MIT Press.

Bentolila, Fernand. 1986. Car en français écrit. *La Linguistique* 22. 95–115.

Bertin, Annie. 1997. *L'expression de la cause en ancien français*. Genève: Librairie Droz.

Blochowiak, Joanna. 2014 A theoretical approach to the quest for understanding. Semantics and pragmatics of *whys* and *becauses*. Phd thesis, University of Geneva.

Canestrelli, Anneloes R., Willem M. Mak & Ted J. M. Sanders. 2013. Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes* 28(9). 1394–1413. DOI: https://doi.org/10.1080/01690965.2012.685885

Canestrelli, Anneloes, Ted Sanders & Willem Mak. 2010. Comparing because to want; How connectives affect the processing of causal relations. In *MAD 2010*. Moissac.

Carletta, Jean. 1996. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics* 22(2). 249–254.

Cohen, Jacob. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1). 37–46. DOI: https://doi.org/10.1177/001316446002000104

Cougnon, Louise-Amélie. 2012. *L'écrit sms: variations lexicale et syntaxique en francophonie*. Université Catholique de Louvain dissertation.

Crible, Ludivine & Degand Liesbeth. 2019. Reliability vs. granularity in discourse annotation: What is the trade-off? *Corpus Linguistics and Linguistic Theory* 15(1). 71–99. DOI: https://doi.org/10.1515/cllt-2016-0046

Davidson, Donald. 1967. Causal relations. *The Journal of Philosophy* 64(21). 691–703. DOI: https://doi.org/10.2307/2023853

Davidson, Donald. 1980. *Essays on actions and events*. Oxford: Clarendon Press.

Debaisieux, Jeanne-Marie. 2002. Le fonctionnement de parce que en français contemporain: étude quantitative. In C. Pusch & W. Raible (eds.), *Romanistische Korpuslinguistik – Romance Corpus linguistics*, 349–362. Tübingen: Gunter Narr Verlag.

Debaisieux, Jeanne-Marie. 2004. Les conjonctions de subordination: mots de grammaire ou mots du discours? Le cas de parce que. *Revue de Sémantique et Pragmatique* 15(16). 51–67.

Degand, Liesbeth. 1996. Causation in Dutch and French. *Functional Descriptions: Theory in Practice*. Catholic University of Louvain. DOI: https://doi.org/10.1075/cilt.121.08deg

Degand, Liesbeth & Fagard, Benjamin. 2012. Competing connectives in the causal domain: French car and parce que. *Journal of Pragmatics* 44(2). 154–168. DOI: https://doi.org/10.1016/j.pragma.2011.12.009

Degand, Liesbeth, & Pander Maat, Henk. 2003. A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale. *LOT Occasional Series* 1. 175–199.

Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht: Kluwer Academic Publishers. DOI: https://doi.org/10.1007/978-94-009-9473-7

Ducrot, Oswald. 1983. Opérateurs argumentatifs et visée argumentative. *Cahiers de Linguistique Française* 5. 7–36.

Fagard, Benjamin & Degand, Liesbeth. 2008. La fortune des mots: grandeur et décadence de 'car'. In H. Durand & B. Laks (eds.), *Congrès Mondial de Linguistique Française*, 211–223. Paris: Institut de Linguistique Française. DOI: https://doi.org/10.1051/cmlf08213

Ferrari, Angela. 1992. Encore à propos de parce que, à la lumière des structures linguistiques de la séquence causale. *Cahiers de Linguistique Française* 13. 183–214.

Frei, Henri. 1982. *La grammaire des fautes*. Genève: Slatkine Publishers.

Grisot, Cristina. 2017. A quantitative approach to conceptual, procedural and pragmatic meaning: Evidence from inter-annotator agreement. *Journal of Pragmatics* 117. 245–263. DOI: https://doi.org/10.1016/j.pragma.2017.06.020

Groupe-λ-1. 1975. Car, parce que, puisque. *Revue Romane* 10(2). 248–280.

Gutzmann, Daniel. 2015. *Use-Conditional Meaning: Studies in Multidimensional Semantics*. Oxford, New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780198723820.001.0001

Gutzmann, Daniel. 2019. *The Grammar of Expressivity*. Oxford, New York: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198812128.001.0001

Hintikka, Jaakko. 1962. *Knowledge and Belief*. Ithaca: Cornell University Press.

Iordanskaja, Lidija. 1993. Pour une description lexicographique des conjonctions du français contemporain. *Le Français Moderne* 2. 159–190.

Kaplan, David. 1999. *The meaning of ouch and oops. Explorations in the theory of Meaning as Use*.

Kerbrat-Orecchioni, Catherine. 1999. *L'énonciation. De la subjectivité dans le langage*. Paris: Armand Colin.

Lambrecht, Knud, Julia Bordeaux & Robert Reichle. 2006. Cognitive constraints on assertion scope. The case of spoken French parce que. In Chiyo Nishida & Jean-Pierre Y. Montreuil (eds.), *New Perspectives on Romance Linguistics: Morphology, syntax, semantics and pragmatics*, 143–154. Amsterdam/Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.275.12lam

Laughlin, Patrick R. 2011. *Group problem solving*. Princeton, NJ, US: Princeton University Press. DOI: https://doi.org/10.1515/9781400836673

Leech, Geoffrey. 2005. Adding linguistic annotation. In M. Wynne (ed.), *Developing linguistic corpora: a guide to good practice*, 17–29. Oxford: Oxbow Books.

Mercier, Hugo & Dan Sperber. 2017. *The enigma of reason*. Cambridge, MA: Harvard University Press. DOI: https://doi.org/10.4159/9780674977860

Millis, K. K. & M. A. Just. 1994. The Influence of Connectives on Sentence Comprehension. *Journal of Memory and Language* 33(1). 128–147. DOI: https://doi.org/10.1006/jmla.1994.1007

Moeschler, Jacques. 1987. Trois emplois de parce que en conversation in Nouvelles approches des connecteurs argumentatifs, temporels et reformulatifs. *Cahiers de Linguistique Française* 8. 97–110.

Moeschler, Jacques. 2003. Causality, lexicon, and discourse meaning. *Rivista Di Linguistica* 15(2). 277–303.

Moeschler, Jacques. 2005. Connecteurs pragmatiques, inférences directionnelles et représentations mentales. *Cahiers Chronos* 12. 35–50.

Nazarenko, Adeline. 2000. *La cause et son expression en Français*. Paris: Ophrys. Retrieved from https://hal.archives-ouvertes.fr/hal-00619264

Nølke, Henning. 1995. Contrastive and argumentative linguistic analysis of the French connectors' donc'and'car'. *Leuvense Bijdragen* 84(3). 313–328.

Noordman, Leo G. M. & Femke De Blijzer. 2000. On the processing of causal relations. *Cause, Condition, Concession and Contrast: Cognitive and Discourse Perspectives*, 35–56. DOI: https://doi.org/10.1515/9783110219043.1.35

Noordman, Leo G. M. & Wietske Vonk. 1997. The different functions of a conjunction in constructing a representation of the discourse. In J. Costermans & M. Fayol (eds.), *Processing interclausal relationships: Studies in the production and comprehension of text*, 75–93. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Pander Maat, Henk & Ted Sanders. 2001. Subjectivity in causal connectives: An empirical study of language in use. *Cognitive Linguistics* 12(3). 247–274. DOI: https://doi.org/10.1515/cogl.2002.003

Pander Maat, Henk & Liesbeth Degand. 2001. Scaling causal relations and connectives in terms of speaker involvement. *Cognitive Linguistics* 12(3). 211–246. DOI: https://doi.org/10.1515/cogl.2002.002

Pasch, Renate. 1983. Die Kausalkonjunktionen da, denn und weil: drei Konjunktionen-drei lexikalische Klassen. *Deutsch Als Fremdsprache* 20(6). 332–337.

Pit, Mirna. 2003. *How to express yourself with a causal connective: subjectivity and causal connectives in Dutch, German and French*. Amsterdam: Rodopi.

Pit, Mirna. 2006. Determining subjectivity in text: The case of backward causal connectives in Dutch. *Discourse Processes* 41(2). 151–174. DOI: https://doi.org/10.1207/s15326950dp4102_3

Pit, Mirna. 2007. Cross-linguistic analyses of backward causal connectives in Dutch, German and French. *Languages in Contrast* 7(1). 53–82. DOI: https://doi.org/10.1075/lic.7.1.04pit

Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33(2). 165–198. DOI: https://doi.org/10.1515/TL.2007.011

Roulet, Eddy, A. Auchlin & J. Moeschler. 1985. *L'articulation du discours en français contemporain*. Berne: Peter Lang.

Sanders, Ted J. M. & Leo G. M. Noordman. 2000. The role of coherence relations and their linguistic markers in text processing. *Discourse Processes* 29(1). 37–60. DOI: https://doi.org/10.1207/S15326950dp2901_3

Sanders, Ted J. M. & Wilbert P. M. Spooren. 2015. Causality and subjectivity in discourse: The meaning and use of causal connectives in spontaneous conversation, chat interactions and written text. *Linguistics* 53(1). 53–92. DOI: https://doi.org/10.1515/ling-2014-0034

Sanders, Ted & Wilbert Spooren. 2009. Causal categories in discourse–Converging evidence from language use. In Ted Sanders & Eve Sweetser (eds.), *Causal categories in discourse and cognition*, 205–246. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110224429.205

Sanders, Ted & Eve Sweetser. 2009. *Causal categories in discourse and cognition*. Berlin: Walter de Gruyter. DOI: https://doi.org/10.1515/9783110224429

Sperber, Dan, Fabrice Clément, Christophe Heintz, Olivier Mascaro, Hugo Mercier, Gloria Origgi & Deirdre Wilson. 2010. Epistemic vigilance. *Mind & Language* 25(4). 359–393. DOI: https://doi.org/10.1111/j.1468-0017.2010.01394.x

Sperber, Dan & Deirdre Wilson. 2015. Beyond speaker's meaning. *Croatian Journal of Philosophy* 15(44). 117–149.

Spooren, Wilbert & Liesbeth Degand. 2010. Coding coherence relations: reliability and validity. *Corpus Linguistics and Linguistic Theory* 6(2). 241–266. DOI: https://doi.org/10.1515/cllt.2010.009

Stukker, Ninke & Ted Sanders. 2012. Subjectivity and prototype structure in causal connectives: A cross-linguistic perspective. *Journal of Pragmatics* 44(2). 169–190. DOI: https://doi.org/10.1016/j.pragma.2011.06.011

Sweetser, Eve. 1990. *From etymology to pragmatics: The mind-body metaphor in semantic structure and semantic change*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511620904

Traxler, Matthew J., Anthony J. Sanford, Joy P. Aked & Linda M. Moxey. 1997. Processing causal and diagnostic statements in discourse. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23(1). 88. DOI: https://doi.org/10.1037/0278-7393.23.1.88

Traxler, Matthew J., Michael D. Bybee & Martin J. Pickering. 1997. Influence of connectives on language comprehension: eye tracking evidence for incremental interpretation. *The Quarterly Journal of Experimental Psychology: Section A* 50(3). 481–497. DOI: https://doi.org/10.1080/027249897391982

Van Belle, William. 1989. Want, omdat en aangezien: een argumentatieve analyse. *Leuvense Bijdragen/Leuven Contributions to Linguistics and Philology* 78. 435–456.

Véronis, Jean & Emilie Guimier de Neef. 2006. Le traitement des nouvelles formes de communication écrite. *Compréhension Automatique Des Langues et Interaction*, 227–248.

Von Wright, Georg Henrik. 1951. *An essay in modal logic*. Amsterdam: North-Holland Pub. Co.

Wharton, Tim. 2016. That bloody so-and-so has retired: expressives revisited. *Lingua* 175. 20–35. DOI: https://doi.org/10.1016/j.lingua.2015.08.004

Wilson, Deirdre & Robyn Carston. 2019. Pragmatics and the challenge of 'non-propositional' effects. *Journal of Pragmatics* 145. 31–38. DOI: https://doi.org/10.1016/j.pragma.2019.01.005

Zufferey, Sandrine. 2012. Car, parce que, puisque revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics* 44(2). 138–153. DOI: https://doi.org/10.1016/j.pragma.2011.09.018

Zufferey, Sandrine. 2014. Givenness, procedural meaning and connectives. The case of French puisque. *Journal of Pragmatics* 62. 121–135. DOI: https://doi.org/10.1016/j.pragma.2013.09.022

Zufferey, Sandrine, Willem M. Mak & Ted J. M. Sanders. 2015. A cross-linguistic perspective on the acquisition of causal connectives and relations. *International Review of Pragmatics* 7(1). 22–39. DOI: https://doi.org/10.1163/18773109-00701002

Zufferey, Sandrine, Willem Mak, Sara Verbrugge & Ted Sanders. 2018. Usage and processing of the French causal connectives 'car'and 'parce que'. *Journal of French Language Studies* 28(1). 85–112. DOI: https://doi.org/10.1017/S0959269517000084

How to cite this article: Blochowiak, Joanna, Cristina Grisot and Liesbeth Degand. 2020. What type of subjectivity lies behind French causal connectives? A corpus-based comparative investigation of *car* and *parce que*. *Glossa: a journal of general linguistics* 5(1): 50. 1–36. DOI: https://doi.org/10.5334/gjgl.1077

Submitted: 29 August 2019 **Accepted:** 16 March 2020 **Published:** 02 June 2020

]u[*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press. **OPEN ACCESS**