

Temporal relations at the sentence and text genre level: the role of linguistic cueing and non-linguistic biases. An annotation study of a bilingual corpus

Abstract

This study investigates the role of non-linguistic biases in the obligatory (verb tenses) and optional (discourse connectives) linguistic marking for inferring temporal relations at the sentence and the text genre levels. Specifically, we formulated and tested several assumptions: (i) the *linguistic cueing assumption* (verb tenses inform language users about the temporal relation), (ii) the *implicitness assumption* (highly expected relations need not be overtly marked), (iii) the *specialized connective assumption* (specialized connectives are more efficient than underspecified ones), (iv) the *text genre assumption* (language users' expectations of temporal relations are linked to the text genre), and (v) the *text status assumption* (information in translated texts tends to be more explicit than in original texts). We carried out an annotation study of a bilingual corpus (French-English) belonging to two different text genres: literary and journalistic. Our results challenge the *implicitness* and the *text status* assumptions while confirming the *linguistic cueing* and the *text genre* assumptions. So, we put forth an alternative view, according to which language users have equal expectations about all three types of temporal relations and are oriented to one relation or the other by linguistic cueing (obligatory and optional marking) as well as text genre.

Keywords: temporal relations; Uniform Information Density hypothesis; implicitness; verb tense and aspect; connectives; corpus data

1 Introduction

In this study, we investigate the role of obligatory and optional linguistic cueing, i.e. verbal tenses and discourse connectives, as well of non-linguistic biases, i.e. text genre and translation status, for expressing temporal relations. While these factors have been investigated individually, as we will discuss below, they have never been investigated all together and from the perspective of their role at the sentence level and the wider level of text genre. This is our ambition in the current study, in which we focus on two tensed languages: a Germanic language – English, and a Romance language – French.

In languages such as English or French, verb tenses are obligatorily marking elements and may cue language users towards inferring a chronological (i.e. time advances from situation 1 to situation 2), synchronous (i.e. time stagnates between situations 1 and 2) or backwards (i.e. time moves backwards from situation 1 to situation 2) temporal relation (Kamp & Rohrer 1983; Asher & Lascarides 2003). By contrast, temporal connectives (specialized connectives) are optional markers of temporal relations. They specifically instruct language users which temporal relation to infer, an instruction which may override the temporal relation signalled by verb tenses. Temporal relations may be marked by the underspecified connective *and*, and even left implicit. At the same time, prior work has suggested that there are other, non-linguistic biases, which drive comprehenders' inferences of temporal relations in the absence of overt linguistic marking (Hopper 1979; Dowty 1979, 1986; Croft 1990; Segal et al. 1991; Murray 1997). For example, a chronological relation would be the default temporal interpretation of a series of two sentences, generally in the Simple Past, while the synchronous and backwards interpretations arise in specific linguistic conditions, such as imperfective marking on the verb and the presence of a synchronous specialized connective (*at the same time*).

Furthermore, scholars working in the fields of discourse analysis (Adam 1992) and the cognitive approach to discourse relations (Sanders 1997; Kintsch 1998; Sagi 2006) argued that most linguistic and discourse phenomena – and thus temporal relations – display significant differences depending on the text genre (for example, narrative literary texts vs. descriptive or argumentative journalistic texts). Moreover, translation studies in corpus linguistics have also pointed out the issue of translation universals (Baker 1993), referring to features of translated texts which appear due to the translation process itself. Of interest for our study is the *explicitation* feature, which refers to the fact that information is more frequently expressed explicitly in translated texts than in original texts.

In this context, our research questions are the following. First, *are specific verb tenses and/or the marking status of a relation (implicit, marked with underspecified connectives or with specialized connectives) significant predictors of one type of temporal relation or another?* Second, *are chronological relations more frequently expressed implicitly than overtly marked with a discourse connective and, conversely, are synchronous and backwards relations more frequently expressed by overt marking than left implicit?* Third, *does the role of verb tenses and discourse connectives depend on text genre and on the text status (original vs. translated)?* To answer these questions, we carried out a bilingual annotation corpus study to examine whether the effects of obligatory and optional linguistic marking of temporal relations differ from one language to another. Below, we describe previous findings regarding the role of obligatory marking (verb tenses) and optional marking (discourse connectives) in confirming or reversing the temporal interpretation triggered by non-linguistic biases (section 2). Then, we introduce and report our bilingual (English and French) annotation corpus study (section 3). We discuss the results and conclude in sections 4 and 5.

2 Overview of the current state of the research

2.1 The role of English and French verb tenses

Prior work in discourse semantics (Kamp 1979; Kamp & Rohrer 1983; Kamp & Reyle 1993; Lascarides & Asher 1993; Asher & Lascarides 2003) has argued that the specification of temporal relations between situations described by adjacent sentences is a function of the verb tenses of the sentences. For example, two sentences expressed in the simple past (English Simple Past and French Passé Simple) are interpreted chronologically (the situation described in the second sentence follows and does not overlap with the first one), as in (1). On the other hand, when a sentence is in the imperfective (English Past Progressive and French Imparfait), it is interpreted as describing a situation temporally synchronous with that of the previous sentence, as in (2). Furthermore, a sentence in the pluperfect (English Pluperfect and French Plus-que-parfait) is interpreted as expressing a past event that has occurred before another past event, as in (3).

- (1) Mary went to the supermarket. Someone called her.
- (2) Mary went to the supermarket. Someone was calling her.
- (3) Mary went to the supermarket. Someone had called her.

Such contributions of verb tenses to the inferring of temporal relations is due, on the one hand, to their temporal content (i.e. tense) and, on the other hand, to their aspectual content (i.e. grammatical aspect). We will refer to the role of tense and grammatical aspect as the **linguistic cueing assumption**, according to which the presence of these linguistic cues language users' expectations about certain types of temporal relations. In what follows, we present the roles of tense and grammatical aspect in more details.

The role of tense in determining the temporal relation between two situations from contiguous sentences is linked to the presence or absence of *referential autonomy* of certain verbal tenses (Reichenbach 1947; Hinrichs 1986; Partee 1973, 1984; Nerbonne 1986; Klein 1994; cf. discussion in Grisot 2018). Precisely, the simple past (English Simple Past and French Passé Simple) is *referentially* autonomous, which means that each situation expressed with this verb tense introduces an R point in the discourse, as in (1). In contrast, the past imperfective (Past Progressive and Imparfait) is *referentially non-autonomous* or *anaphoric* (Ducrot 1979; Kamp and Rohrer 1983; Tasmowsky-De Ryck 1985; Molendijk 1990; Kleiber 2003), which means they cannot introduce a new R and are dependent on an R already existing in the context, as in (2). As such, a situation expressed with the Past Progressive/Imparfait is interpreted as simultaneous with another situation (in the Simple Past, for example) because the former is temporally anchored (via the R point) to the latter. Nevertheless, empirical studies have found important differences between the English Simple Past and the French Passé Simple with respect to their role for signalling temporal relations. For example, Grisot (2017) showed using a corpus annotation study that the English Simple Past signals chronological or backwards relations (*narrative usage* in Grisot's terminology) in 59% of the cases whereas it signals synchronous relations (*non-narrative usage* in Grisot's terminology) in 41% of the cases. In contrast, the French Passé Simple was annotated as having a narrative usage much more frequently, that is in 92% of the cases. Grisot (2017) studied the role of the French Imparfait for signalling temporal relations and she found that this verb tense signalled synchronous relations in 77.5% of the cases. No data is available for the English Past Progressive.

The pluperfect (Pluperfect and Plus-que-parfait) represents a slightly different situation than that of the simple past and the past imperfective, as its meaning already includes the temporal relation between two situations. For Reichenbach (1947), the pluperfect expresses that *the point of event E precedes the point of reference R which precedes the point of speech S*. The reference point may be provided by a past time temporal adverbial or by another past time event (expressed with the simple past or the past progressive), as in (3). This is why the situation *had called* in example (3) is understood as taking place before the situation *went to the supermarket*, and the temporal relation inferred between sentence 1 and sentence 2 is backwards.

For some scholars, it is the grammatical aspect of verb tenses, i.e. *perfective* or *imperfective* (Comrie 1976), which plays a role for specifying the temporal relation between two situations. According to Comrie (1976), the perfective aspect displays a situation as a single whole, without internal structure, and with highlighted boundaries. The imperfective aspect gives access to the internal structure of a situation, or to a moment other than the initial and the final boundaries. The perfect focuses on the relevance of the resultative state of a past situation. Through their meanings, perfective (e.g. the Simple Past, the Passé Simple and the Passé Composé in some of its uses) and imperfective (e.g. the Past Progressive and the Present Progressive, the Imparfait and the Présent) verb tenses affect the mental representations of situations which hearers or readers build during the comprehension process (Johnson-Laird 1983; Kintsch 1998; for a review see Zwaan & Radvansky 1998). Specifically, the perfective instructs comprehenders to build mental representations of completed situations, whereas the imperfective instructs comprehenders to build mental representations of ongoing situations. Consequently, two situations mentally viewed as completed are understood as having taken place in a chronological way, whereas two situations mentally viewed as ongoing are understood as having taken place synchronously.

Numerous experimental studies have tested whether this is the effect of the instructional meaning of the perfective and imperfective in the inference of temporal relations. Up to now, the findings from English are inconclusive. While some scholars (e.g. Magliano & Schleich 2000; Feller et al. 2019) have found that the imperfective necessarily favours synchronous interpretations of the situations described by two adjacent sentences, others (e.g. Madden & Zwaan 2003; Author1, under review) have found that the imperfective triggers synchronous and chronological interpretations with equal frequency. For Madden and Zwaan (2003:670), comprehenders may infer that the following situation unfolds simultaneously or chronologically with respect to the current situation, depending on how they represent the internal structure of a situation expressed in the imperfective.

As for French, Grisot & Blochowiak (2019) investigate the role of French verb tenses and of the presence and absence of temporal connectives, separately and in interaction, as processing instructions for chronological relations in French, by means of a self-paced reading experiment and an offline evaluation experiment. They test how two past time verb tenses (the *Passé Simple* and the *Passé Composé*) and the presence and absence of two specialized connectives (*puis* 'then' and *ensuite* 'then') influence the processing and acceptability of short narrative contexts. Following theoretical descriptions of the meanings of these verb tenses, the *Passé Simple* is thought to encode an instruction about inferring a chronological relation between the situations described in the short narrative context, whereas the *Passé Composé* is supposed to be neutral with respect to the temporal relation inferred (Kamp & Rohrer 1983; Moeschler 2000, 2002; Saussure 2003). The ensuing predictions in terms of processing time were that the short narratives expressed with the *Passé Simple* will be read faster than those expressed with the *Passé Composé*. The results of the self-paced reading experiment did not show significant processing differences between the two verb tenses, whereas the offline evaluation experiment revealed comprehenders' preference for the *Passé Composé*. These findings showed that the *Passé Composé* is as efficient as the *Passé Simple* for inferring chronological relations. The authors interpreted this result in terms of the aspectual component that these two verb tenses share (they are both perfective verb tenses) arguing that it is grammatical aspect which instructs the hearers about the inference of chronological relations. Furthermore, speakers consciously preferred the *Passé Composé* over the *Passé Simple* because of the former is much more frequent and has, in contemporary French, nearly replaced the latter.

In sum, up to now existing experimental and empirical studies focusing on individual verb tenses showed the following: (i) both the French *Passé Simple* et *Passé Composé* trigger more frequently chronological than synchronous relations (Grisot & Blochowiak 2019; Grisot 2017), (ii) the French *Imparfait* signals more frequently synchronous than chronological relations (Grisot 2017), (iii) the English *Simple Past* was found either to signal equally frequently chronological and synchronous relations (Grisot 2017) or to trigger more frequently chronological than synchronous relations (Magliano & Schleich 2000; Madden and Zwaan 2003; Feller et al. 2019; Author1, under review), and (iv) the English *Past Progressive* was found either to trigger equally frequently chronological and synchronous relations (Madden & Zwaan 2003; Author1, under review) or to trigger more frequently chronological than synchronous relations (Magliano & Schleich 2000; Feller et al. 2019). These results are insufficient to understand the role of verb tenses for temporal relations because they show neither a monolingual not cross-linguistic comprehensive picture of all verb tenses and because they seem to be contradictory.

2.2 Implicit and overtly marked temporal relations

Nevertheless, the above-mentioned patterns linked to the contribution of verb tenses for temporal relations do not seem to be exclusive. Sometimes, the temporal interpretation triggered by the verb tense is altered. For example, it may be modified by the use of specialized temporal connectives, as in (4), in which the second sentence describes an event that happened before the event from the first sentence. In (5), the use of the adverbial *two minutes later* changes the synchronous interpretation (triggered by the imperfective) into a chronological interpretation. Also, the temporal relation may be marked by the conjunction *and*, as in (6), which may be understood as either chronological or synchronous. Equally, despite the use of the pluperfect in (7), the temporal relation holding between the two situations changes from backwards – as in (3) – to chronological or even ambiguous, depending on how *and* is interpreted.

- (4) Mary went to the supermarket after someone called her.
- (5) Mary went to the supermarket. Two minutes later, someone was calling her.
- (6) Mary went to the supermarket and she picked the kids up from school.
- (7) Mary had gone to the supermarket and someone had called her.

These examples demonstrate that temporal relations may be either left implicit – and only inferred via verbal tenses – as in (1)-(3), or overtly marked by some lexical item, as in (4)-(7). Unlike verb tenses, overt marking is optional and may be performed using specialized connectives (such as *then* for chronological, *after* for backwards and *at the same time* for synchronous relations), the underspecified connective *and*, and other types of expressions such as noun phrases (*two minutes later* or *the next summer*). Even if most of the attention given to the marking of coherence relations has focused on discourse connectives, coherence relations may be also marked by other types of signals (such as the prepositional phrase *in the event that* from example (8), which signals a condition relation) (see the *Signalling Corpus* released by Das & Taboada 2013, 2018).

- (8) This notice must not be removed from the software, and in the event that the software is divided, it should be attached to every part. (Das & Taboada 2013: 256)

Problematically, the linguistic elements functioning as signals do not directly signal specific coherence relations and they are categorized or grouped with great difficulty (but see Hoek et al.'s 2018 proposal to categorise segment-internal elements as signals of coherence). In contrast, our knowledge about both discourse connectives, whether they are specialized or underspecified, is much more stable. Experimental studies show that when a discourse relation is overtly marked using a specialized connective, the processing and the integration of the information from the following segment is faster compared to when the relation is left implicit (e.g. Britton et al. 1982; Haberlandt 1982; Deaton & Gernsbacher 1997 for causal relations). For temporal relations, Grisot & Blochowiak's (2019) reading data showed that the specialized connectives *puis* and *ensuite* facilitated the processing of chronological relations, by comparison with contexts where the relation is left implicit. Their acceptability task experiment also revealed that comprehenders preferred chronological relations to be overtly marked rather than left implicit. Other studies investigated the role of the underspecified connective *and* to signal various discourse relations, among which are both synchronous and chronological relations (see for instance, Blakemore 1987; Carston 1993, 1998, 2002; Luscher & Moeschler 1990; Levinson, 2000). It is argued that the semantics of the connective *and* is minimal as it encodes a concept of logical conjunction (Blakemore 1987; Wilson & Sperber 1998; Blakemore & Carston 1999). This minimal semantic meaning can have cumulative or distributive readings (Saussure & Sthioul 2002) or it can refer to additivity (Spooren 1997; Zeevat & Jasinskaja 2007). In sum, scholars

agree that the meaning of *and* is underspecified and it can get more specific in the context depending on the type of eventualities described. Given the lack of specific meaning of some connectives, such as *and*, Sanders (2005) advances the *specialized connective* assumption, according to which the more specific the linguistic marker, the greater facilitating effect would be found on the segment following the linguistic marker. Thus, specialized connectives are more efficient than underspecified ones.

Among the three possible temporal relations (chronological, backwards and synchronous), prior work has argued that chronological relations have a special status: they represent the default case, whereas the other two types are non-default relations. This means that comprehenders interpret by default two contiguous sentences as conveying that the situations expressed unfolded chronologically (Grice 1975; Hopper 1979; Dowty 1979, 1986; Croft 1990; Segal et al. 1991; Murray 1997). This proposal has been linked to compliance during the comprehension process with one or more principles. Among these principles, some are said to be cognitive biases, such as the *continuity* hypothesis (Segal et al. 1991; Murray 1997); some are conversational principles, such as the *iconicity* hypothesis (Hopper 1979) or the *iconicity of sequence* (Croft 1990) and the *maxim of order* (Grice 1975; cf. also Levinson 2000); others are semantic principles, such as Dowty's *Temporal Discourse Interpretation Principle* (TDIP; Dowty 1979, 1986). The *continuity* hypothesis predicts that, when reading a text, readers are cognitively biased to interpret a situation described by a sentence as being chronologically or causally related to the previous one. The *iconicity of sequence* principle sets out that comprehenders should interpret the linear ordering of main and subordinate clauses as mirroring the sequential order in which the described events occurred. In Gricean pragmatics (Grice 1975), the conversational *maxim of order* states that cooperative speakers should narrate things in the (sequential) order in which they happened. Finally, Dowty's TDIP posits that, for a sequence of sentences, (i) each following sentence should be interpreted to be temporally consistent with a definite temporal adverbial, if there is one, and (ii) as temporally subsequent to the previous sentence otherwise. All these proposals share the core idea that a sequence of clauses describing a series of events should be interpreted as preserving the iconic order – that is, the chronological order in which the events happened in the world. In short, the chronological interpretation of a series events is the default one.

Other scholars have argued that, by default, coherence relations presenting a forward movement of time (i.e. chronological), are highly expected by comprehenders; their overt marking is thus not needed, as it would make the linguistic content more informative than necessary (Asr & Demberg 2012, 2015). This proposal may be explained by Grice's *maxim of quantity*, Horn's R and Q principles (1984) and Levy & Jaeger's *Uniform Information Density* hypothesis (UID) (Levy & Jaeger 2007). According to Grice, speakers are expected to make their contribution as informative as required for comprehenders to grasp the intended message, but no more informative than necessary. In Horn's pragmatic model, speakers should make their contributions sufficient (the Q principle) and necessary (the R principle). In other words, the Q principle reduces the hearer's effort to interpret an utterance while the R principle prevents the speaker from producing unnecessary linguistic content. The UID further specifies Grice's maxim and Horn's principles with respect to what is the most optimal way in which a speaker should communicate information: the speaker should distribute the information evenly across a text or utterance, thereby reducing or omitting redundant optional markers. As put by Hoek et al. (2018: 278), "a coherence relation should therefore be sufficiently marked so that the hearer will be able to construct the appropriate relations, but not overly or unnecessarily marked so as to limit the speaker's effort". As Asr & Demberg (2012: 2671) put it: "At the level of inter-sentential

relations, this would mean the presence of explicit connectives is necessary when the relation is unexpected, but that a connective may be implicit if the relation is predictable.” Where Grice’s maxim, Horn’s principles and the UID hypothesis initially refers to production, Asr & Demberg (2012, 2015) related production tendencies to comprehension biases and showed that corpus-based patterns provide support for the expectation hypothesis (Langacker 2000) (cf. also Grisot & Blochowiak 2019).

Asr & Demberg (2012: 2671) then formulated the **implicitness assumption**, according to which the implicitness of the discourse connective is “a sign of expectation of the discourse relation: if readers have a default preference to infer a specific relation in the text, this type of relation should appear without explicit markers.” Asr and Demberg (2012) provide evidence for *the implicitness assumption* from their study of the frequency of implicit and explicitly marked coherence discourse relations in the Penn Discourse Tree Bank (PDTB; Prasad et al. 2008). They find a pattern in the distribution of explicit and implicit relations: relations such as Conjunction, Contrast, Concession, Synchrony, Asynchrony and Condition are most often overtly marked, whereas relations such as Cause, Instantiation and Restatement are most often left implicit. Furthermore, Hoek et al. (2018) show, by means of a parallel corpus study, that the implicit vs. overt marking status of coherence relations depends not only on the relation itself but also on the cognitive complexity associated to each type of coherence relations (based on the Cognitive Complexity of Coherence Relations framework developed by Sanders et al. 1992, and later work¹), among other factors. For example, they find that coherence relations with basic order are more often implicit than relations with non-basic order, and that conditional relations are less often implicit than causal or additive ones (Hoek et al. 2018: 127).

In this precise study, Hoek et al. (2018) do not distinguish among the coherence relations that may be and relations that may not be temporally characterized (e.g. in some cases, additive relations may have a temporal interpretation, which can be either synchronous or chronological). Nevertheless, this specification is provided in Asr & Demberg (2012). For them, asynchronous temporal relations (i.e. chronological and backwards) and, to a point, synchronous relations “describe discontinuous events” (page 2678), where the *continuous/discontinuous* status refers to Murray’s (1997) *continuity hypothesis*. Compared to all the other coherence relations in the PDTB, they find that relations “classified as discontinuous are exactly the ones with lower [than the average] values of implicitness.” (page 2678). In other words, compared to all other discourse relations, those which may be characterized with one of the three types of temporal configurations are more frequently overtly marked than left implicit. Furthermore, Asr & Demberg compare the rates of explicit marking vs. implicit status of coherence relations presenting chronological and backwards movements of time. They find that “the forward relation [i.e. chronological] is associated to a higher degree of implicitness” (page 2679). So, in the PDTB corpus, coherence relations presenting a chronological order of events are significantly more frequently left implicit than overtly marked, whereas relations presenting a backwards order display the reverse distribution. What is of interest for the current paper is the

¹ In the CCR framework, coherence relations are annotated with respect to four types of primitives: *polarity* (causal relations are positive while contrast relations are negative), *basic operation* (a relation of implication between P and Q for causal relations and a relation of logical conjunction for additive relations), *source of coherence* (content, epistemic and speech act relations) and *order of segments* (the order of P and Q in implication relations). It is important to mention that Evers-Vermeul et al. (2017) add a fifth primitive to the CCR framework: temporality (coherence relations may be temporal or non-temporal).

implicitness assumption, which has been proposed for coherence relations in the PDTB framework and which we will apply to (chronological, synchronous and backwards) temporal relations.

2.3 Beyond sentence-level: the role of text genre and text status

Finally, two other factors that play a role in enhancing comprehenders expectations about the discourse to follow are **text genre** (Sanders 1997; Kintsch 1998; Sagi 2006; Author1, under review) and **text status**: if the text analysed is an original text or a translated text (Baker 1993; Colleague & Author1 2020).

Regarding text genre, it has been argued that the current distribution of discourse relations in a text (specific to each the text genre) influences the readers' expectation about what relations will appear more frequently later on in the text. Specifically, Sanders (1997) was the first to argue that other factors than the content of the discourse might influence the specifics of the relation inferred by the comprehender. For instance, he constructed an experiment to examine participants' sensitivity to the source of coherence (i.e. content or *semantic*, epistemic and speech act relations or *pragmatic*) in two text genres: argumentative and descriptive. He presented expert discourse analysts with discourses that included "chameleon" causal relations, whose source of coherence could be either semantic or pragmatic. While both sources of coherence were found to be equally frequent in argumentative texts, most participants agreed that in descriptive texts the source of coherence was most likely semantic. In addition, Sanders (1997) provides an analysis of these discourse genres and shows that the interpretations made by the participants in his study mirrored the distributions of relation types in the argumentative, and respectively, descriptive genre. One possible explanation of these results may be that discourse comprehension is directly affected by the perceived genre of the discourse (e.g. the *textual schemas* in Kintsch's 1998 terms). As put by Sagi (2006), this account implies that people differentiate among discourse genres and that such discourses are clearly identifiable types. Problematically, this is not always the case since there are situations in which several different genres could be applicable to the same text. Furthermore, in some cases a novel genre may be encountered, or even expected. In other words, readers have expectations about what discourse relations should be found in a certain text genre. For example, Author1 (under review) investigated how people infer temporal relations (chronological vs. synchronous) and assessed experimentally the role of coreference patterns (i.e. whether two situations are performed by the same agent or by different agents) in two fine-grained distinctions of text types: transfer-of-possession contexts (study 1) as in (9) and in short narratives (study 2), as in (10).

(9) John handed a book to Bob. He [wanted him to read it].

(10) Mary walked along the street. She gave her mother a call.

In study 1, participants had to provide continuations after having read the pronoun *he*, which were annotated after the study, by two independent annotators, with respect to temporal relations holding between the situation expressed in the first sentence and that expressed in the second sentence. In study 2, participants' task was to decide whether there was a chronological or a synchronous temporal relation between the situations expressed. The results revealed that temporal relation instantiated in cases of coreference is different in transfer-of-possession contexts and in short narratives. In the former, coreference triggered more continuations about the source of the transfer-of-possession verb, revealing the participants' expectation to have more information about the source. These cases were annotated most frequently as synchronous relations. In the latter, coreference triggered more chronological relations revealing their expectation to learn more about the sequence of the protagonist' actions. In sum, Author1's experimental study provides evidence that text genre raises

different expectations about what discourse relations will appear more frequently later in the text. We will call this the **text genre assumption**.

With respect to text status, translated texts may have specific features which originate in the translation process itself (cf. Gellerstam's 1996 *translationese* and Baker's 1993 hypothesis of *translation universals*). The advantages of translation corpora are that they provide objective linguistic data (Dyvik 1998) and that they are intended to express the same meaning and have the same discourse functions in the languages considered (Johansson 1998), hence inherently provide the terms of comparison. More specifically, one of the benefits of using translation corpora to investigate the issue of temporal relations and their implicitness is that we keep constant the content of the texts in the two languages we analyse, and thus eliminate the risk that the conceptual content expressed influences our findings for the two languages. The disadvantage of translation corpora is their very nature: the translation process can create a bias affecting the target text, such as the *explicitation* feature – that is, a higher rate of explicit information in translated texts than in original texts. Nevertheless, as recent research in translation studies has shown (House 2018), quantitative evidence supporting the thesis of translation universals has yet to be found. For example, in Colleague & Author1 (2020), we investigate the translation of texts between French, English and Mandarin Chinese with respect to temporal reference (i.e. inflected verbs for tense and aspect in English and French vs. non-inflected verbs in Mandarin), and did not find evidence confirming the existence of the two translation universals tested: *explicitation* and *normalization* (that is, the tendency to conform to patterns and practices which are typical of the TL, or even to exaggerate them). In the current study, to control for a potential translation bias, in each of the two languages investigated we compare the results we find when the texts were originally written in that language to the results we find when the texts were translated into that language. We will call this the **text status assumption**.

2.4 Hypotheses

In the previous sections, we introduced five assumptions: the *linguistic cueing* assumption, the *implicitness* assumption, the *specialized connective* assumption, the *text genre* assumption and the *text status* assumption. Table 1 provides the variables we test in our annotation study, the hypotheses we formulate on the basis of the above-mentioned assumptions issued from the state of the research, and their ensuing predictions for our empirical corpus study.

Table 1: Hypotheses and predictions

Variables tested	Assumptions	Predictions
Obligatory covert marking – referentially autonomous vs. non-autonomous VTs, perfective vs. imperfective verb tenses	– verb tenses inform language users about the temporal relation in a systematic and obligatory way (the <i>linguistic cueing assumption</i>)	– referentially autonomous and perfective VTs signal chronological relations – the pluperfect signals backwards relations – referentially non-autonomous and imperfective VTs signal synchronous relations or allow chronological and synchronous equally

Optional overt marking – implicit vs. overtly marked	– highly expected relations are most frequently left implicit, unexpected relations are most frequently overtly marked (the <i>implicitness assumption</i>)	– chronological relations are more frequently left implicit, synchronous and backwards relations are most frequently overtly marked
	– overtly marked chronological relations are processed faster than implicit ones (Grisot & Blochowiak 2019)	– chronological relations are more frequently overtly marked.
Optional overt marking – specialized and underspecified connectives	– specialized connectives are more efficient than underspecified ones (the <i>specialized connective assumption</i>)	– chronological, synchronous and backwards relations are more frequently overtly marked with specialized connectives than with underspecified connectives
Text genre – literary vs. journalistic	– literary texts increase the expectation of chronological relations, which thus need not be overtly marked (the <i>text genre assumption</i>)	– more chronological relations in literary texts than in journalistic texts (by definition) – in literary texts, chronological relations are overtly marked less frequently than they are left implicit
Text status – original vs. translated	– the rate of explicitation is higher in translated texts than in original texts (the <i>text status assumption</i>)	– more overtly marked temporal relations in translated texts than in original texts

The first variable tested is the obligatory marking via verb tenses (linked to the *linguistic cueing assumption*), and we focus on the distinctions between referentially autonomous vs. non-autonomous verb tenses on the one hand, and perfective and imperfective verb tenses on the other hand. The main hypothesis is that verb tenses inform language users in a systematic and obligatory way with respect to the temporal relation inferred. Following both theoretical linguistic and psycholinguistic studies, we expect referentially autonomous and perfective verb tenses to signal chronological relations, the pluperfect to signal backwards relations. Cross-linguistic differences are expected with respect to the role of the simple past: the English Simple Past might behave according to this pattern (due to its perspective aspect) or signal with equal frequency chronological and synchronous relations (as found by Grisot 2017). In contrast, the French Passé Simple should signal almost exclusively chronological relations. Referentially non-autonomous and imperfective verb tenses are expected to signal synchronous relations. While the French Imparfait is expected to behave according to this pattern (as found by Grisot 2017), the English Past Progressive might either follow it as well (as found by Magliano & Schleich 2000; Feller et al. 2019) or it might allow chronological and synchronous relations equally frequently (as found by Madden & Zwaan 2003; Author1, under review).

The second variable tested is optional overt marking status – in particular, the implicit vs. overtly marked distinction. We consider two approaches which generate opposite predictions. First, we consider the *implicitness assumption*, according to which implicitness is a sign of expectation of the discourse relation, and that chronological relations are highly expected relations. As such, we expect to find that chronological relations are more frequently left implicit, whereas synchronous and backwards relations are most frequently overtly marked. Second, we consider Grisot & Blochowiak's

(2019) pragmatic approach, according to which chronological relations overtly marked with specialized connectives are processed faster and preferred by comprehenders over implicit relations. Thus, Grisot & Blochowiak's results contradict the implicit assumption. As such, we expect to find that chronological relations are more frequently overtly marked with specialized connectives than left implicit.

The third variable tested is a finer-grained view of optional overt marking status, i.e. the distinction between specialized and underspecified overt markers. We follow the *specialized connective* assumption, according to which specialized connectives are more efficient than underspecified ones. In pragmatics (Horn 1984; Sperber & Wilson 1986, 2012), the notion of efficiency is linked to the speakers' and hearers' search for reducing cognitive effort during utterance interpretation (cf. Horn's 1984 R and Q principle, and Sperber & Wilson's notion of *relevance*). In this context, we expect chronological, synchronous and backwards relations to be more frequently overtly marked with specialized connectives than with underspecified connectives because the former type of connectives allows them to satisfy their search for relevance.

The fourth variable tested is the text genre. We consider the hypothesis that text genre raises language users' expectations of discourse relations to follow. In particular, literary texts, which are frequently narrative texts, increase the expectation of inferring chronological relations more frequently than synchronous or backwards relations. Since chronological relations become more predictable due to the text genre, they need not be overtly marked. As such, we expect more chronological relations in literary texts than in journalistic texts (by definition); we also expect that chronological relations are less frequently overtly marked than left implicit in literary texts.

The fifth variable tested is the original vs. translated status of the text. Following Baker's (1993) hypothesis of translation universals, and more specifically the *explicitation* feature, we expect to find more overtly marked temporal relations (chronological, synchronous and backwards) in translated texts than in the original texts.

3 A bilingual corpus annotation study

3.1 Method

Annotation scheme. We used a three-category annotation scheme of types of relations holding between two situations: *chronological*, *backwards*, *synchronous*. Annotators could also use a 4th case, the *ambiguous* label, where they were not able to identify one of the first three cases uniquely. We calculated inter-annotator agreement rate using percentages - for a three-way classification, chance-level agreement is of 33.33%, and disagreements were not resolved². Further analysis was carried out only on the set of data for which annotators agreed.

Procedure. For each language, two annotators were presented with annotation guidelines, underwent a training phase, and then annotated the data independently. The annotation guidelines provided explanation and examples for each of the five possible cases of the classification scheme.

Annotators. Two native speakers of English (Study 1) from the University of xxx and two other native speakers of French (Study 2) from the University of xxx annotated the data and received financial compensation for their work. They were of similar background with respect to their age and level of education (linguistics undergraduates).

² As pointed out by an anonymous reviewer, it is likely that the difficult cases from the disagreement set of data included mixed signals and non-typical patterns of verb-tense / connective combination. As such, the results we discuss here reflect the most frequent and typical patterns. Future research should focus on the less frequent and non-typical cases.

Data analyses. English and French data were prepared for analysis and analysed in a similar manner. First, English and French sets of agreement data were coded with three surface features: the verb tense of segment 1 (VTS1) and segment 2 (VTS2); the presence or absence of an overt marker of the temporal relation (Marking); and whether the connective used was a specialized connective or the underspecified connectives *and* and *et*.

Two analyses were carried out on the English and French agreement sets of data. In the first analysis, we carried out quantitative descriptive analyses to explore which verbal tenses (from S1 and S2) and overt markers are most frequently used to express chronological, synchronous and backwards temporal relations when the temporal relation is not left implicit. In the second analysis, we fitted generalized mixed models in order to see which factors – i.e. obligatory marking (verb tense), optional marking (implicit, underspecified and specialized markers), register (literature and journalistic) and language status (original and translated) – played statistically significant roles in predicting chronological relations. The logistic mixed-effects models were fitted using the R software (R Development Core Team, 2010, version 3.1.2). Models were tested using the `glmer()` function of the `lmer4` package of R, and model comparisons were assessed using the `anova()` function, which calculates the chi-square value of the log-likelihood to evaluate the difference between models, following Baayen’s (2008) procedure. To fit the models, we coded the data to define a binary dependent variable and several independent variables, i.e. random and fixed factors, as shown in Table 2.

Table 2: Dependent and independent variables

Type of variable	Name	Categories in English	Categories in French
Dependent variable	<i>temporal relation</i>	chronological (coded as “1”) vs. synchronous (coded as “0”)	chronological (coded as “1”) vs. synchronous (coded as “0”)
Independent variables: fixed factors	<i>marking status</i>	implicit vs. <i>and</i> vs. specific connective vs. others	implicit vs. <i>et</i> vs. specific connective vs. others
	<i>VT from S1</i>	SP vs. <i>-ing</i> vs. others	PS vs. IMP vs. others
	<i>VT from S2</i>	SP vs. <i>-ing</i> vs. others	PS vs. IMP vs. others
	<i>corpus</i>	literary vs. journalistic	literary vs. journalistic
	<i>language status</i>	SL vs. TL	SL vs. TL
Independent variables: random factors	<i>infinitive verb from S1</i>	e.g. <i>run</i> vs. <i>be</i> vs. <i>climb</i>	e.g. <i>courir</i> vs. <i>être</i> vs. <i>grimper</i>
	<i>infinitive verb from S2</i>	e.g. <i>run</i> vs. <i>be</i> vs. <i>climb</i>	e.g. <i>courir</i> vs. <i>être</i> vs. <i>grimper</i>

Following Johnson (2008) and Field et al. (2014), we built the models by going from the simplest model to the model of interest. Into the simplest model, consisting of only the random structure, we incorporated the various fixed factors one after the other – the *marking status*, the *VT from S1* and from *S2*, the *corpus* and *language status*, as well as their possible interactions – and compared them using the `anova()` function.

3.2 Material

In this study, we used translation corpora between French and English belonging to two stylistic registers: literary and journalistic. In order to answer the research questions formulated in section 1, we studied the following sets of data. The data from the literary register consists in 125 randomly selected excerpts from two literary texts: 60 were from “A Christmas Carol” by Charles Dickens, originally written in English and professionally translated into French; 65 were from “Le comte de

Grisot, C., & Blochowiak, J. (2021). Temporal Relations at the Sentence and Text Genre Level: The Role of Linguistic Cueing and Non-linguistic Biases—An Annotation Study of a Bilingual Corpus. *Corpus Pragmatics*, 1-41.

Monte Cristo” by A. Dumas, originally written in French and professionally translated into English. The literary texts were accessed from the Project Gutenberg website (<https://www.gutenberg.org/>). The data from the journalistic register consists in 147 randomly selected excerpts from articles published in “Le monde diplomatique”, available on <http://cabal.rezo.net>: 77 were from articles originally written in English and professionally translated into French; 70 were originally written in French and professionally translated into English.

Regarding the English data, 125 literary and 147 journalistic corpus excerpts (i.e. chunks of text with an average length 30-40 words were used). In order to prepare the excerpts for the annotation task, we identified for each excerpt pairs of verbs expressing two situations which were temporally related. A total of 502 pairs of situations from the literary texts and 329 pairs of situations from the journalistic texts were annotated using the three-category annotation scheme. We observed full agreement at a rate much higher than chance: 78% (649 items). Disagreements were not resolved, and we considered for further analysis the 649 cases of agreement (224 cases in which English was the original language, and 425 in which English was the translated language), wherein there were 354 chronological relations, 236 synchronous relations and 59 backwards relations. The two ambiguous cases were considered cases of disagreement.

Regarding the French data, 129 literary and 147 journalistic corpus excerpts were used in this second annotation study. As for the English set of data, we identified for each excerpt pairs of verbs expressing two situations which were temporally related. A total of 514 pairs of situations from the literary texts and 345 pairs of situations from the journalistic texts were annotated using the same three-category annotation scheme. We observed full agreement at a rate much higher than chance, at 69% (597 items). Disagreements were not resolved, and we considered for further analysis the 597 cases of agreement (283 cases in which French was the original language and 314 in which French was the translated language), wherein there were 260 chronological relations, 286 synchronous relations and 51 backwards relations.

3.3 Results

3.3.1 English data

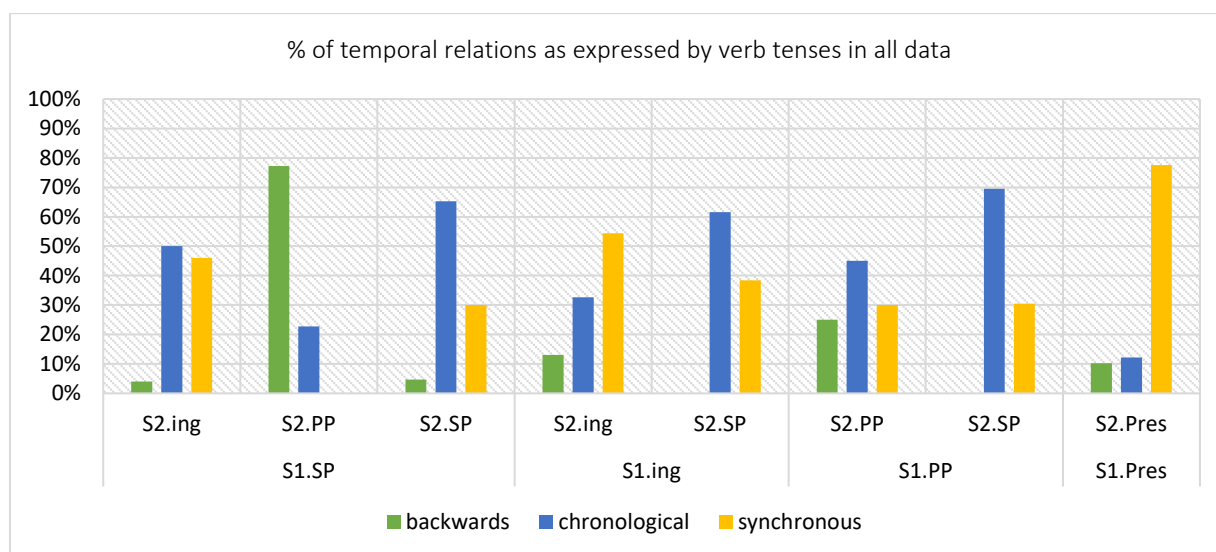
3.3.1.1 Descriptive quantitative analyses

On the set of 649 agreement data points, we carried out a quantitative exploratory analysis to show how verb tenses combine and co-occur with overt markers to express these chronological, synchronous and backwards temporal relations.

The role of VTs in expressing chronological, synchronous and backwards relations

Figure 1 shows the distribution of different verb tense combinations (from segment 1 and segment 2) with respect to chronological, synchronous and backwards relations. Recall that verb tenses, due to their referential autonomy or non-autonomy and to their perfective/imperfective nature, are expected strongly to influence the temporal relation inferred: for example, the progressive *-ing* is expected to signal synchronous relations more frequently than chronological or backwards relations, whereas the pluperfect is expected to signal backwards relations more frequently than chronological or synchronous relations. Note that here we focus not on the role of a verb tense taken individually but its combination with another verb tense.

Figure 1: Obligatory marking of temporal relations: verb tenses in S1 and in S2³



As Figure 1 illustrates, all verb tense combinations may generally be used to express all three types of temporal relations: chronological, synchronous and backwards relations.

First, the SP-SP combination is the most frequent verb tense combination (406 occurrences among the 649 verb pairs analysed), and most frequently expresses chronological relations (65%) (as in (11)), followed by synchronous relations (30%) (as in (12)(13)) and backwards relations (5%) (as in (13)). As expected, when the SP is followed by a PP, a backwards relation was annotated in 77% of cases (as in (14)) and a chronological relation in 23% (as in (15)); a synchronous relation was never annotated (0%). When the SP is followed by the *-ing* morpheme, annotators annotated almost equally frequently a chronological relation (50%) (as in (16)) or a synchronous relation (46%) (as in (17)); a backwards relation was annotated in only 4% of the cases (as in (18)). A Pearson's Chi-Square test performed on the data consisting of the SP in S1 (first left-hand panel of Figure 1) shows that this observed distribution is significantly different from a chance expected distribution ($\chi^2(4) = 151.27, p < .0001$).

- (11) The clerk in the tank involuntarily **applauded**: becoming immediately sensible of the impropriety, he **poked** the fire. [Literary corpus]
- (12) Scrooge never painted out Old Marley's name. There it **stood**, years afterwards, above the warehouse door: Scrooge and Marley. The firm **was known** as Scrooge and Marley. [Literary corpus]
- (13) Up Scrooge went, not caring a button for that: darkness is cheap, and Scrooge liked it. But before he **shut** his heavy door, he **walked** through his rooms to see that all was right. [Literary corpus]
- (14) This prospect of fresh festivity **redoubled** the hilarity of the guests to such a degree, that the elder Dantès, who, at the commencement of the repast, **had commented** upon the silence that prevailed. [Literary corpus]
- (15) Scrooge **lay** in this state until the chimes **had gone** three quarters more. [Literary corpus]
- (16) When the clock struck eleven, this domestic ball broke up. Mr. and Mrs. Fezziwig **took** their stations, one on either side the door, and **shaking** hands with every person individually as he or she went out, wished him or her a Merry Christmas. [Literary corpus]

³ *-ing* = all verb tenses ending with the *-ing* morpheme, PP = Pluperfect, SP = Simple Past, Spres = Simple Present, PresPerf = Present Perfect

- (17) Up Scrooge **went**, not **caring** a button for that: darkness is cheap, and Scrooge liked it. [Literary corpus]
- (18) The honorable, the king's attorney, is informed by a friend of the throne and religion, that one Edmond Dantès, mate of the ship Pharaon, **arrived** this morning from Smyrna, after **having touched** at Naples and Porto-Ferrajo. [Literary corpus]

Second, the combination of verbs presenting the *-ing* morpheme (i.e. the progressive or the gerund) most frequently expresses synchronous relations (54%) (as in (19)), followed by chronological (33%) (as in (20)) and backwards relations (13%) (as in (21)). When the *-ing* morpheme is followed by an SP, the two situations are more frequently interpreted as expressing a chronological relation (62%) (as in (22)), less frequently a synchronous relation (38%) (as in (23)), and never a backwards relation (0%). A Pearson's Chi-Square test showed that this distribution (second left-hand panel of Figure 1) is significantly different from a chance expected distribution ($\chi^2(2) = 7.48, p < .02$).

- (19) Facing head-on the dramatic collapse in exchange rates and the breathtakingly high rate of inflation, thousands of them - used to bending with the wind - **are** already **packing** their bags of their own accord. Some **are returning** to their home countries despite the endemic poverty and political risks: the Thai kingdom's undesirables include Shan and Karen refugees, members of ethnic minorities persecuted by the Burmese military junta. [Journalistic corpus]
- (20) "Ah," exclaimed the young girl, **blushing** with delight, and fairly leaping in excess of love, "you see he has not forgotten me, for here he is!" And **rushing** towards the door, she opened it, saying, "Here, Edmond, here I am!" [Literary corpus]
- (21) While French television's "Les Guignols" on Canal+ **were lampooning** Jamie Shea's press conferences (for two night the puppet Shea **was making** excuses for a missile attack on a bus), the British and Americans stayed po-faced to the end. [Journalistic corpus]
- (22) But how much greater was his horror, when the phantom **taking off** the bandage round its head, as if it were too warm to wear in-doors, its lower jaw **dropped** down upon its breast! [Literary corpus]
- (23) **Rising** to be life president of the De Beers mining conglomerate ("diamonds are for ever"), Barney Barnato **died** in mysterious circumstances at the age of 42. [Journalistic corpus]

Third, as we expected, when the PP is followed by an SP, a chronological (70%) (as in (24)) or a synchronous (30%) relation was annotated (as in (25)), and never a backwards relation (0%). When the PP verb combines with another PP verb, language users usually infer a chronological relation (45%) (as in (26)), followed by a synchronous relation (30%) (as in (27)), then a backwards relation (25%) (as in (28)). A Pearson's Chi-Square test showed that this distribution (third left-hand panel of Figure 1) is significantly different from a chance expected distribution ($\chi^2(2) = 6.86, p < .03$).

- (24) But he put his hand upon the key he **had relinquished**, **turned** it sturdily. [Literary corpus]
- (25) But they and their spirit voices faded together; and the night became as it **had been** when he **walked** home. [Literary corpus]
- (26) The ship drew on and **had** safely **passed** the strait, which some volcanic shock has made between the Calasareigne and Jaros islands; **had doubled** Pomègue. [Literary corpus]
- (27) The city **had** entirely **vanished**. Not a vestige of it was to be seen. The darkness and the mist **had vanished** with it. [Literary corpus]
- (28) Within hours the BBC **had promised** the Irishwoman air time to explain what she **had** actually **seen**. [Journalistic Corpus]

Fourth, when two present time verb tenses (Simple Present or Present Perfect) co-occur, annotators most frequently annotated a synchronous relation (78% when a PresPerf is used and 86%

when a Spres is used) (as in (29)), and less frequently a chronological (11% for the PresPerf and 14% for the Spres) (as in (30)) or backwards relation (11% for the PresPerf and 0% for the Spres) (as in (31)). A Pearson's Chi-Square test showed that these distributional differences (fourth left-hand panel of Figure 1) between the Spres and PresPerf are not statistically significant ($\chi^2(2) = 4.1, p = .12$), indicating that the two verb tenses have similar roles in expressing backwards, chronological and synchronous relations.

(29) This Sunday, the church **is full** and the doors **are open**. [Journalistic Corpus]

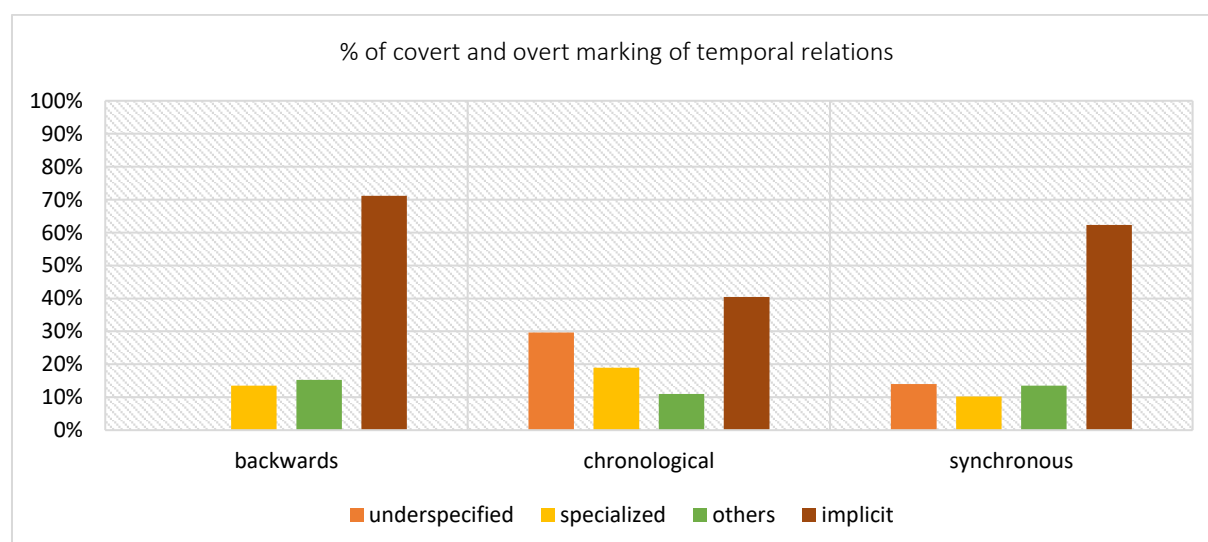
(30) In two years, the committee has dealt with some 50 cases, most of them women. They **leave** Asia, Africa or the Middle East for a promised Eldorado in France. When they arrive, they **are forced** to become skivvies.

(31) It **is** always an event at Marseilles for a ship to come into port, especially when this ship, like the Pharaon, **has been built**, rigged, and laden at the old Phocce docks, and belongs to an owner of the city. [Literary corpus]

The marking of chronological, synchronous and backwards relations

Figure 2 shows the distribution of chronological, synchronous and backwards relations according to their percentages of implicit cases, overt marking using the underspecified connective *and*, specialized connectives for expressing synchronous, chronological and backwards, and other linguistic markers such as NP or PPs.

Figure 2: Covert and overt marking of chronological, synchronous and backwards relations



First, backwards relations are most frequently expressed implicitly (71%), and are overtly marked in 29% of cases (with a specialized connective in 14% of cases, and other types of linguistic marker in

15% of cases). As already shown in Figure 1, the high percentage of implicit cases might be due to use of the PP in S2, preceded by a SP or an *-ing* form. Second, chronological relations are less frequently expressed implicitly (40%) than they are overtly marked (a total of 60%), distributed as follows: 30% with the underspecified *and*; 19% with a specialized connective, and 11% with other types of overt marker, such as NPs and PPs. Third, synchronous relations are also most frequently expressed implicitly (62%) rather than overtly marked (38%), distributed as follows: 14% with the underspecified *and*; 10% with a specialized connective, and 14% with other types of overt marker, such as NPs and PPs. A Pearson's Chi-Square test performed on the data for Figure 2 shows that this observed distribution is significantly different from a chance expected distribution ($\chi^2(6) = 56.96, p < .0001$).

From the perspective of the use of the underspecified vs. specialized connectives to express temporal relations, our data indicate a higher frequency of the underspecified connective *and* for overtly marking chronological (30%) and synchronous (14%) relations than specialized chronological (14%) and synchronous (10%) connectives. In contrast, no case was found when *and* is used to express backwards relations. Furthermore, among all cases where *and* is used, 76% correspond to a chronological relation, 24% to a synchronous relation and 0% to a backwards relation.

3.3.1.2 Statistical modelling

Data. After removing the cases of backwards chronological relations (59 occurrences) and the cases in which overt marking was performed using other types of linguistic marker than underspecified and specified connectives (41 occurrences), the final set of data on which we fitted the mixed models has 519 observations.

Results. To find the best random structure, we first tested two models: one model which only encompassed the infinitive verb from S1 as a random factor, and one which only encompassed the infinitive verb from S2 as a random factor. We compared the two models using the *anova()* function, which shows that the model which fits the data better is the model with the infinitive verb from S2 (the latter having an AIC score of 666.43, compared to 694.45 for the former).

Adding the *marking status* fixed factor to the simplest model (which only included the infinitive form of the verb from S2 as a random effect) significantly improved the model ($\Delta\chi^2 = 26.69, \Delta df = 1, p < .001$), as the probability of inferring a chronological relation significantly increases in the presence of overt marking such as *and* (33.3%) ($\beta=1.228, SE=0.279, p < .001$) or a specialized connective (21.3%) ($\beta=1.080, SE=0.313, p < .001$) compared to when no connective is used at all.

We then tested whether adding the VT from S1 and/or S2 significantly improved the model containing only one fixed factor (*marking status*). All three models (with VT from S1 only, with VT from S2 only and with the VTs from both S1 and S2) significantly improved the model. Nevertheless, their comparison using the *anova()* function showed that the best improvement originated in the model which only added the VT from S2 ($\Delta\chi^2 = 12.213, \Delta df = 1, p < .001$). This model showed that the presence of an overt marker such as *and* or of a specialized connective significantly increases the probability of chronological relations. Similarly, the presence of an SP – as opposed to *-ing* or other VTs – significantly increases the probability of chronological relations. Additionally, we tested a model that includes the interaction *marking status* and the VT from S2. The model with the interaction *marking status**the VT from S2 did not significantly improve the model without this interaction ($\Delta\chi^2 = 4.92, \Delta df = 4, p = .295$).

To the models with two fixed factors (*marking status* and the VT from S2), we added the interaction with *Corpus* (journalistic vs. literary) to check whether the role of *marking status* differs with respect to stylistic register. Adding *Corpus* significantly improved the model with *marking status* and the VT from S2 ($\Delta\chi^2 = 26.142, \Delta df = 3, p < .001$), showing a significant main effect of *Corpus*

($\beta=0.799$, $SE=0.312$, $p < .001$) according to which chronological relations have a high frequency in the literary register, as well as a significant interaction *marking status*Corpus* ($\beta=1.715$, $SE=0.662$, $p < .001$) according to which the use of *and* increases the rate of chronological relations more in the literary register (39%) than the journalistic register (14.9%). At the same time, the presence of a specialized connective and the absence of an overt marker significantly decreases the probability of chronological relations in literary texts (19.1% and 41.9% respectively) more so than it does in journalistic texts (28.4% and 56.8% respectively).

Finally, to the model with three fixed factors (*marking status*, the *VT* from *S2* and *Corpus*), we added the interaction with *language status* (English original vs. English translated) variable, to check whether the role of *marking status* differs with respect to whether the data analysed were originally written in English or translated into English. Adding the interaction *marking status*language status* did not significantly differ from the model with three fixed factors ($\Delta\chi^2 = 0$, $\Delta df = 1$, $p = 1$), as the rate of use of *and* is similar in texts originally written in English (48.6%) and texts translated into English (51.4%) ($\beta=-0.916$, $SE=0.606$, $p = .139$). The same pattern is found in the case of specialized connectives: the rate of use of specialized connectives is similar in texts originally written in English (61.2%) and texts translated into English (38.8%) ($\beta=0.088$, $SE=0.660$, $p = .892$).

3.3.1.3 Discussion of results

Both the exploratory quantitative analysis (section 3.3.1.1) and the statistical modelling (section 3.3.1.2) provided several fresh answers to our first two research questions, also confirming certain observations made by theoretical studies. Our two questions considered whether specific verb tenses and the presence or absence of overt markers are significant predictors for chronological, synchronous or backwards relations. We expected chronological relations to be significantly more frequently expressed with the Simple Past (due to its perfective aspect), synchronous relations to be more frequently expressed with the *-ing* morpheme (all verb tenses expressing the progressive meaning and the gerund), and backwards relations to be more frequently expressed with the Pluperfect. In our original analysis, we considered how the combination of two verb tenses and their co-occurrence with a specialized connective or an underspecified connective affects the temporal relation inferred. At the same time, we expected synchronous and backwards relations to be significantly more frequently overtly marked, and chronological relations to be left implicit. This prediction was formulated in the context of the *implicitness assumption* that highly expected discourse relations, such as chronological relations, are most frequently left implicit whereas synchronous and backwards relations which are not highly expected require overt marking. The analyses carried out on our English data partially confirmed the *linguistic cueing assumption* (regarding verb tenses), and did not validate the predictions formulated on the basis of the *implicitness assumption*, as we discuss below.

First, we found proof partially supporting the idea that backwards temporal relations are most frequently expressed in the presence of the Pluperfect – in particular, by the SP-PP and PP-PP combinations. Nevertheless, backwards temporal interpretations also arise in the presence of SP-SP and other less frequent combinations, as SP-*ing* and Spres_PresPerf-PresPerf. Furthermore, we found that backwards relations are never overtly marked using the underspecified connective *and*; they may be marked using a specialized connective, but are most frequently left implicit. These results confirm the predictions formulated on the basis of the previous theoretical studies about the role of verb tenses, which suggested that the PP, via its semantics, is a strong cue for backwards relations. In contrast, they do not support the proposal that backwards relations are unexpected and should be overtly marked. One might suggest that backwards relations are left implicit as the PP is already a

strong cue, and hence that their predictability increases. However, backwards relations are also implicit when they are expressed by a SP-SP pair of verb tenses. Unfortunately, due to the limited number of backwards relations in this study, they were not included in the statistical modelling analysis carried out whereby we tested the roles of stylistic register and language status in addition to verb tense and marking status. We believe that more data with backwards relations are needed to reveal more completely which factors predict a backwards temporal relation and whether they are, as in our limited set of data points, most frequently left implicit.

Second, thanks to the higher number of data points, we were able to investigate the predictors of chronological and synchronous relations using logistic mixed-effects models. The first fixed factor which we added to the simplest model including only the random structure was *marking status*. The model with this predictor confirmed the results of the exploratory quantitative analysis: chronological relations are significantly more frequently overtly marked than they are left implicit. In particular, to mark overtly a chronological relation, the underspecified connective *and* is used more frequently than a specialized connective. Conversely, synchronous relations are significantly more frequently left implicit than they are overtly marked. As was found for chronological relations, synchronous relations are overtly marked with the underspecified connective *and* more frequently than with a specialized connective. This finding does not confirm the prediction made on the basis of *implicitness assumption*. As we will discuss in section 4, we believe that this finding raises two possible issues: (i) chronological relations are not necessarily highly expected relations, whereas synchronous relations are not necessarily less predictable than chronological ones; and (ii) highly expected relations are not necessarily left implicit whereas less predictable relations need not be overtly marked.

The second fixed factor we added to model including the random structure and the *marking status* was the VT from S1 and/or from S2. The model comparison showed that it is the VT from S2 that is the most relevant for predicting temporal relations. In particular, we found evidence supporting the hypothesis that the SP significantly increases the rate of chronological relations. By contrast, we did not find evidence in favour of the hypothesis that verb tenses containing the *-ing* morpheme favour the inference of synchronous relations: in our data, this morpheme was equally associated with synchronous and chronological relations. In English, the verb tenses containing *-ing* may correspond either to the progressive (Past, Present or Future progressives) or the gerund. To understand our results better, we made this fine-grained distinction with respect to the verb tense from S2. Among the 64 occurrences of verb tenses containing *-ing*, we found that 23 corresponded to the genuine progressive aspect and 41 to the gerund. In the case of the 23 occurrences of the progressive *-ing*, the temporal relation expressed was synchronous (78%) more frequently than chronological (22%). On the other hand, in the case of the 41 occurrences of the gerund *-ing*, the temporal relation expressed was synchronous (53%) and chronological (47%) with equal frequency.

So, our results reveal that progressive *-ing* more frequently signals a synchronous relation, as found by Magliano and Schleich (2000) and Feller et al. (2019) rather than Madden and Zwaan (2003) and Author1 (under review). However, when we consider the usage of the *-ing* morpheme more widely (corresponding to both the progressive and the gerund), its presence signals chronological and synchronous relations with equal frequency (as the results of the mixed-effects model show). Author1 (under review) investigates transfer-of-possession contexts (e.g. *John handed/was handing a book to Bob. He spotted another one on the shelf.*), suggesting that the progressive aspect (*was handing*) diminishes the higher rate of chronological relations triggered by the perfective (*handed*) (from 63% with the perfective to 46% with the progressive). Our corpus results provide complementary evidence

that the role of the English *-ing* morpheme for the inference of a chronological or a synchronous relation depends on whether it corresponds to the progressive aspect or the gerund.

The third fixed factor we added to the model, consisting of the random structure, the *marking status* and the *VT* from S2, was *Corpus* and its interaction with *marking status*. We found a main effect of *Corpus* according to which there are more chronological relations in the literary texts than in the journalistic texts. This result is not unexpected, and we consider it a validation of our statistical modelling: the essence of literary texts is that there is a story being narrated. The interaction *Corpus*marking status* showed that there are significant interactions with the type of corpus. First, we found that the use of *and* in the literary register increases the rate of chronological relations more than it does in the journalistic register. Second, the presence of a specialized connective, as well as the absence of an overt marker, significantly increases the probability of chronological relations in journalistic texts. In other words, chronological relations are frequently overtly marked with the underspecified *and* in literary texts, and with specialized temporal connectives in journalistic texts.

Finally, the fourth fixed factor we added to the ongoing model was the interaction *language status*marking status*, to check whether the role of *marking status* differs with respect to whether the data analyzed were originally written in English or translated into English. Adding this new factor did not significantly improve the model, as we did not find a significant main effect of *language status*, and the behavior of chronological relations with regard to their implicit vs. overt marking using the underspecified or the specialized connectives did not vary according to whether the texts were originally written in English or translated into English. With this result, our data do not support the hypothesis of *translation universals*, and in particular the *explicitation* feature of translated texts. If this translation universal were confirmed in our data, we should have found significantly more cases of overt marking of chronological relations in the texts translated into English than in the texts originally written in English.

3.3.2 French data

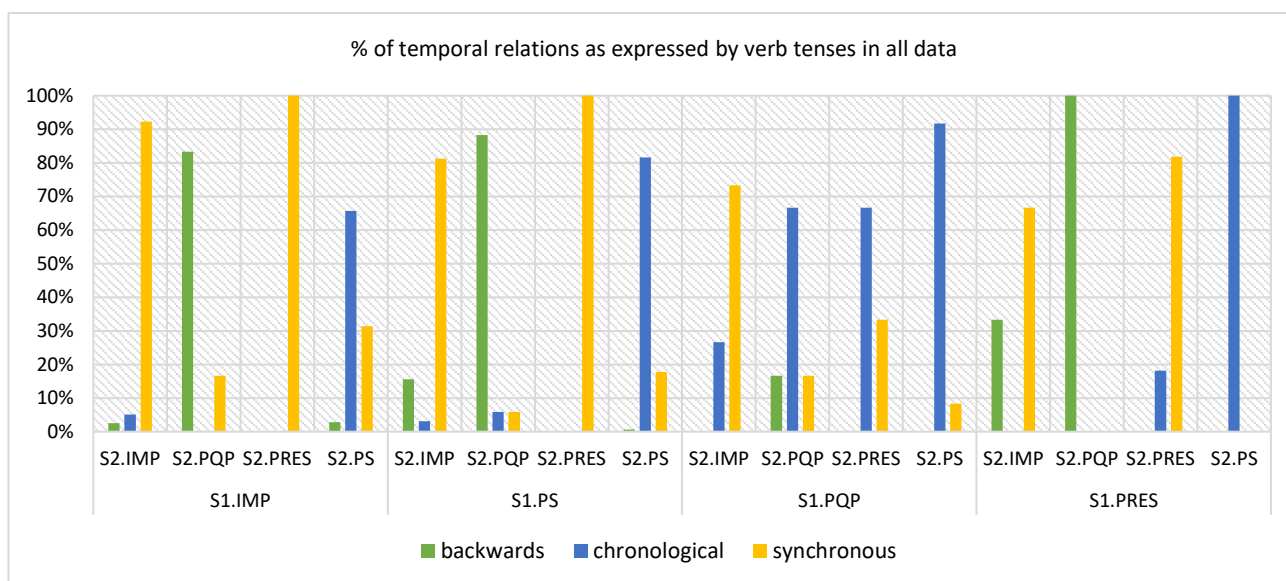
3.3.2.1 Descriptive quantitative analyses

As in Study 1, we carried out a quantitative exploratory analysis to show how verbal tenses combine and co-occur with overt markers to express these chronological, synchronous and backwards temporal relations for the set of 597 agreement data points.

The role of VTs to express chronological, synchronous and backwards relations

Figure 3 shows the distribution of different verb tense combinations (from segment 1 and segment 2) with respect to chronological, synchronous and backwards relations. In general, all verb tense combinations may be used to express all three types of temporal relations examined.

Figure 3: Obligatory marking of temporal relations: verb tenses in S1 and in S2⁴



First, following theoretical studies, the use of the IMP is expected to systematically signal synchronous relations. In our set of data, this is the case when an IMP (from S1) combines with another IMP (92%) (as in (32)) or a PRES (100%) (as in (33)). When an IMP is followed by a PQP, the temporal interpretation is that of a backwards relation (83%) (as in (34)). Finally, the IMP-PS combination signals a chronological interpretation in 66% of cases (as in (35)), a synchronous relation in 31% of cases (as in (36)), and a backwards relation in 3% of cases (as in (37)). A Pearson's Chi-Square test performed on this set of data shows that this observed distribution is significantly different from a chance expected distribution ($\chi^2(6) = 114.47, p < .0001$).

- (32) Près du pilote [...] était un jeune homme au geste rapide et à l'œil actif, qui **surveillait** chaque mouvement du navire et **répétait** chaque ordre du pilote. [Literary corpus]
 'Near the pilot [...] was a young man with a quick gesture and an active eye, who monitored every movement of the ship and repeated every order of the pilot.'
- (33) Un immense bonheur les isolait du monde, et ils ne **parlaient** que par ces mots entrecoupés qui **sont** les élans d'une joie si vive qu'ils semblent l'expression de la douleur. [Literary corpus]
 'A great happiness isolated them from the world, and they spoke only with these interspersed words which are the impulses of a joy so vivid that they seem the expression of pain.'
- (34) Il se tourna vers l'esprit, et voyant qu'il le regardait avec un visage dans lequel [...] **se retrouvaient** des traits épars de tous les visages qu'il lui **avait montrés**, il se jeta sur lui. [Literary corpus]
 'He turned to the spirit, and seeing that he looked at him with a face in which [...] there were scattered features of all the faces he had shown him, he threw himself upon him.'
- (35) Le père Pamphile **venait de voir passer** Dantès il n'y avait pas dix minutes. Certain que Dantès était aux Catalans, ils **s'assirent** sous le feuillage naissant des platanes. [Literary corpus]
 'Father Pamphile had just seen Dantes pass less than ten minutes ago. Sure that Dantes was to the Catalans, they sat under the nascent foliage of the plane trees.'
- (36) En ce moment, le jeune homme **revenait**; Danglars **s'éloigna**. [Literary corpus]
 'At that moment the young man was coming back; Danglars walked away.'
- (37) Scrooge **savait-il** qu'il [Marley] **fût mort**? Sans contredit. [Literary corpus]
 'Did Scrooge know that he [Marley] was dead? For sure'.

⁴ IMP = Imparfait, PQP = Plus-que-parfait, PS = Passé Simple, PRES = Présent

Second, theoretical studies predict that the use of the PS signals systematically chronological relations. This is confirmed when the PS combines with another PS (82%) (as in (38)). However, when it combines with an IMP (73%) (as in (40)) or a PRES (100%) (as in (41)), the relation inferred is a synchronous one. Finally, when the PS is followed by a PQP, annotators most frequently annotated a backwards relation (88%) (as in (42)). A Pearson's Chi-Square test performed on this set of data shows that this observed distribution is significantly different from a chance expected distribution ($\chi^2(6) = 288.7, p < .0001$).

- (38) Et, sans abandonner Mercédès, dont il tenait la main serrée dans une des siennes, il **tendit** avec un mouvement de cordialité son autre main au Catalan. Mais Fernand [...] **resta** muet et immobile comme une statue. [Literary corpus]
 'And, without abandoning Mercedes, whose hand he held tightly in one of his own, he extended his other hand to the Catalan with a movement of cordiality. But Fernand [...] remained silent and motionless like a statue.'
- (39) Mercédès **rougit**. Fernand **se tourmentait** sur sa chaise, tressaillait au moindre bruit. [Literary corpus]
 'Mercedes blushed. Fernand tormented himself in his chair, trembling at the slightest sound.'
- (40) Athéisme il y **eut**, pourtant. Le Parti communiste cubain **se définissait** à l'époque de la révolution, en 1959, comme athée, proclamant l'inexistence de Dieu. [Journalistic corpus]
 'Atheism there was, however. The Cuban Communist Party defined itself at the time of the revolution, in 1959, as atheist, proclaiming the non-existence of God.'
- (41) En octobre 1992, Maale Adumim **devint** la première colonie à obtenir le statut de ville. Elle **compte**, en 1999, 25 000 habitants. [Journalistic corpus]
 'In October 1992, Maale Adumim became the first colony to be granted city status. It had a population of 25,000 in 1999.'
- (42) «Oui, tout serait dit», **s'écria** Caderousse, qui par un dernier effort d'intelligence **avait suivi** la lecture [...]. [Literary corpus]
 "'Yes, everything would be said," cried Caderousse, who by a last effort of intelligence had followed the reading [...].'

Third, the PQP-IMP combination of verb tenses most frequently signals synchronous relations (73%) (as in (43)), followed by chronological relations (27%) (as in (44)). The PQP-PS, PQP-PRES and PQP-PQP combinations most frequently signal chronological relations (92%, 67% and 67% respectively) (as in (45), (46)(46), and respectively (47)). A Pearson's Chi-Square test performed on this set of data shows that this observed distribution is significantly different from a chance expected distribution ($\chi^2(6) = 18.21, p < .005$).

- (43) Retournant en hâte à Pristina, elle déclara à un reporter radio de la BBC qu'elle **avait vu** un corps à Vucitrn, mais que le coin **était** « plein de policiers serbes ». [Journalistic corpus]
 'Rushing back to Pristina, she told a BBC radio reporter that she had seen a body in Vucitrn, but that the area was "full of Serbian policemen".'
- (44) La demande lui fut accordée, et trois mois après, autour des douze ou quinze bâtiments qui **avaient amené** ces bohémiens de la mer, un petit village **s'élevait**. [Literary corpus]
 'The request was granted to him, and three months later, around the twelve or fifteen ships that had brought these gypsies from the sea, a small village rose.'
- (45) Mais il mit la main sur la clef, qu'il **avait lâchée** d'abord, la **tourna** brusquement, entra et alluma sa chandelle. [Literary corpus]
 'But he put his hand on the key, which he had dropped at first, turned it abruptly, entered and lit his candle.'
- (46) Ceux qui **avaient** le mieux **réussi** à enfourcher les « dragons » **sont** en voie d'être terrassés. [Journalistic corpus]

‘Those who had been most successful in riding the "dragons" are on their way to being defeated.’

- (47) A la reprise économique, elle **n’avait pu** les récupérer : relativement bien formés en Malaisie, ils **avaient trouvé** mieux au Japon et à Taïwan. [Journalistic corpus]

‘When the economy recovered, it had not been able to recover them: relatively well trained in Malaysia, they had found better in Japan and Taiwan.’

Fourth, when the PRES was followed by the IMP, annotators annotated synchronous relations (67%) (as in (48)), followed by backwards relations (33%) (as in (49)); they never annotated chronological relations (0%). Backwards relations are also the preferred interpretation in the case of the PRES-PQP combination (100%) (as in (50)). The PRES-PRES combination of verb tenses signals synchronous relations in 82% of cases (as in (51)) and chronological relations in 18% of cases (as in (52)); the PRES-PS combination systematically signals chronological relations (100%). A Pearson’s Chi-Square test performed on this set of data shows that this observed distribution is significantly different from a chance expected distribution ($\chi^2(6) = 60.14, p < .0001$).

- (48) L’obscurité ne **coûte** pas cher, c’est pour cela que Scrooge ne la **détestait** pas. [Literary corpus]

‘Darkness is cheap, which is why Scrooge did not hate it.’

- (49) Ils **veulent** adopter les techniques développées par le privé pour maximiser le taux d’écoute. L’éthique initiale de Pacifica lui **intimait** au contraire de procurer une voix à ceux qui n’ont rien [...]. [Journalistic corpus]

‘They want to adopt the techniques developed by the private sector to maximize the listening rate. Instead, Pacifica’s original ethics meant that he provided a voice for those who had nothing [...].’

- (50) Elle **se souvient** d’un fâcheux précédent au milieu des années 80. Elle **avait** alors **expulsé** 200 000 immigrants pour cause de difficultés passagères. [Journalistic corpus]

‘She remembers an unfortunate precedent in the mid-1980s. At the time, she had deported 200,000 immigrants because of temporary difficulties.’

- (51) Je **dis** ce que je veux dans mon église, affirme le curé, mais je **sais** que tout est surveillé. [Journalistic corpus]

‘I say what I want in my church, says the priest, but I know that everything is monitored.’

- (52) Le grand tournant se produit en novembre dernier lorsque la Corée du Sud [...] **vient** quémander l’aide du FMI. Le président américain **change** alors de ton. [Journalistic corpus]

‘The big turning point came last November when South Korea [...] begged for IMF assistance. The U.S. president then changes his tone.’

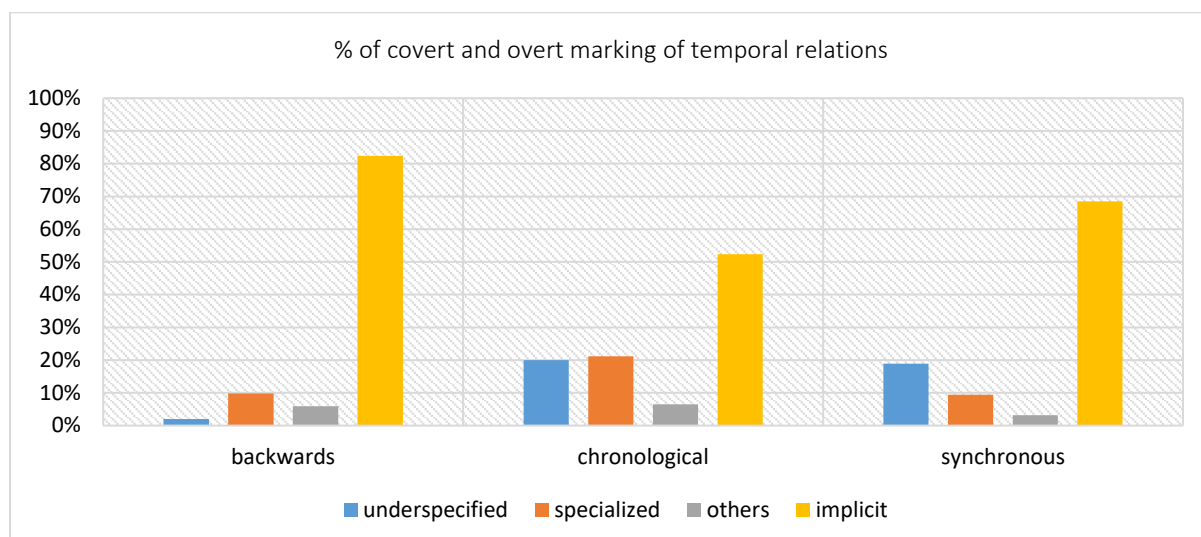
- (53) Un immense bonheur les isolait du monde, et ils ne parlaient que par ces mots entrecoupés qui **sont** les élans d’une joie si vive qu’ils semblent l’expression de la douleur. Tout à coup, Edmond **aperçut** la figure sombre de Fernand. [Literary corpus]

‘A great happiness isolated them from the world, and they spoke only with these interspersed words which are the impulses of a joy so vivid that they seem the expression of pain. Suddenly Edmond saw the dark figure of Fernand.’

The covert and overt marking of chronological, synchronous and backwards relations

Figure 4 displays the distribution of chronological, synchronous and backwards relations according to their percentages of implicit cases, overt marking using the underspecified connective *and*, specialized connectives for expressing synchronous, chronological or backwards relations, and other linguistic markers such as NP or PPs.

Figure 4: Covert and overt marking of chronological, synchronous and backwards relations



First, in our set of data, backwards relations are most frequently left implicit (82%), and only very rarely overtly marked (18%). When they are overtly marked, they are expressed with the underspecified connective *et* (2%), with a specialized connective (10%), and with other types of overt markers (6%). As already shown in Figure 3, the high percentage of implicit cases might be due to the use of the PQP in S2, following an IMP, a PS or a PRES. Second, chronological relations are left implicit in 52% of cases and overtly marked in 48% of cases, using the underspecified *et* (20%), a specialized connective (21%), and other types of overt markers (7%). Third, synchronous relations are left implicit in 69% of cases and overtly marked in 31% of cases, using the underspecified *et* (19%), a specialized connective (9%), and other types of overt markers (3%). A Pearson's Chi-Square test performed on the data underlying Figure 4 shows that this observed distribution is significantly different from a chance expected distribution ($\chi^2(6) = 34.24, p < .0001$).

In terms of the use of underspecified vs. specialized connectives to express temporal relations, our data shows that – in the case of chronological relations – the underspecified and specialized connectives are used with similar frequencies (20% vs. 19%). For synchronous relations, underspecified *et* is used more frequently (20%) than specialized connectives (9%). Furthermore, the underspecified *et* is used with equal frequency to express chronological (49%) and synchronous relations (50%), and almost never used to express backwards relations (1%).

3.3.2.2 Statistical modelling

Data. After removing the 51 cases of backwards chronological relations and the 26 cases in which overt marking made use of types of linguistic markers other than underspecified and specified connectives, the final set of data on which we fitted the mixed models had 528 observations.

Results. To find the best random structure, we first tested two models: one model which only encompassed the infinitive verb from S1 as a random factor; and a model that only encompassed the infinitive verb from S2 as a random factor. We compared the two models using the *anova()* function, which shows that the model that fits the data better is the model with the infinitive verb from S2 (the latter having an AIC score of 711.34, compared to 718.51 for the former).

Adding the *marking status* fixed factor to the simplest model significantly improved the model ($\Delta\chi^2 = 15.91, \Delta df = 1, p < .001$), as the presence of a specialized connective triggers a chronological relation (67%) more frequently than a synchronous relation (33%) ($\beta=1.080, SE=0.313, p < .001$). In contrast, the rate of chronological relations does not change according to underspecified *et* ($\beta=0.312$,

$SE=0.254$, $p = .219$), as it is used equally frequently to express chronological (49.1%) and synchronous (50.9%) relations.

We then tested whether adding the *VT* from *S1* and/or from *S2* significantly improved the model containing only one fixed factor (*marking status*). All three models (with *VT* from *S1* only, with *VT* from *S2* only, and with the *VTs* from both *S1* and *S2*) significantly improved the model. Nevertheless, their comparison using the *anova()* function showed that the best improvement originated in the model which added only the *VT* from *S2* ($\Delta\chi^2 = 106.836$, $\Delta df = 1$, $p < .001$). This model showed that the presence of a specialized connective significantly increases the probability of chronological relations ($\beta=1.226$, $SE=0.330$, $p < .001$). Similarly, the presence of a *PS*, compared to other *VTs*, significantly increases the probability of chronological relations ($\beta=1.870$, $SE=0.261$, $p < .001$). In contrast, the presence of an *IMP* (9.1%) ($\beta=-1.846$, $SE=0.356$, $p < .001$) decreases the probability of chronological relations. Additionally, we tested a model that includes the interaction *marking status* and the *VT* from *S2*. The model with the interaction *marking status**the *VT* from *S2* did not significantly improve the model without this interaction ($\Delta\chi^2 = 0.596$, $\Delta df = 1$, $p = .440$). Following the Occam's razor principle of avoiding unusable complexity, we kept the model without the interaction for further analysis.

To the models with two fixed factors (*marking status* and the *VT* from *S2*) we added the interaction with *Corpus* (journalistic vs. literary). Adding *Corpus* significantly improved the model with *marking status* and the *VT* from *S2* ($\Delta\chi^2 = 9.944$, $\Delta df = 3$, $p < .001$), showing a significant main effect of *Corpus* ($\beta=0.806$, $SE=0.300$, $p < .001$) according to which the literary register significantly increases the rate of chronological relations. The interaction *marking status***Corpus* tends towards statistical significance ($\beta=-1.145$, $SE=0.661$, $p = .083$), according to which the presence of a specialized connective decreases the probability of chronological relations in literary texts (18.2%) more than it does in journalistic texts (32.1%). As for the presence of the underspecified connective *et* to signal chronological relations, the difference between literary (15.4%) and journalistic texts (24.2%) is not statistically significant ($\beta=0.079$, $SE=0.547$, $p = .889$).

Finally, to the model with three fixed factors (*marking status*, *VT* from *S2* and *Corpus*), we added the interaction with *language status* variable (French original vs. French translated) to check whether the role of *marking status* differs with respect to whether the data analysed were originally written in French or translated into French. Adding this interaction did not improve the model with *marking status*, *VT* from *S2* and *Corpus* as fixed factors ($\Delta\chi^2 = 0$, $\Delta df = 1$, $p = 1$). Precisely, the use of specialized connectives to mark chronological relations is not found to be statistically different in texts translated into French (26.5%) or originally written in French (19%) ($\beta=0.088$, $SE=0.660$, $p = 0.893$). Also, the rates of use of *et* to mark chronological relations is not found to be statistically different in texts originally written in French (28%) or translated into French (14.5%) ($\beta=-0.833$, $SE=0.581$, $p = .151$).

3.3.2.3 Discussion of results

We expected chronological relations to be significantly more frequently expressed with the *Passé Simple*, synchronous relations to be more frequently expressed with the *Imparfait*, and backwards relations to be more frequently expressed with the *Plus-que-parfait*. Furthermore, we also examined how the combination of two verb tenses and their co-occurrence with a specialized or underspecified connective affects the temporal relation inferred. Following the *UID hypothesis* and the *implicitness assumption*, we expected synchronous and backwards relations to be significantly more frequently overtly marked, where chronological relations are left implicit. The analyses carried out on our French data confirmed the *linguistic cueing assumption* (regarding the role of verb tenses), and did not confirm the *implicitness assumption*, as we discuss below.

First, the descriptive quantitative analysis validated the theoretical predictions that each of the three temporal relations investigated is closely linked to the usage of a particular verb tense: the PS for chronological relations, the IMP for synchronous relations, and the PQP for backwards relations. As predicted, the presence of a PS significantly signals a chronological relation, especially when the PS is preceded by another PS or by other verb tenses, such as IMP, a PQP or a PRES. However, chronological relations are also frequently signalled by tense combinations in which the PQP is followed by the PQP or other verb tenses, such as the PRES or the PS. These results are to be explained by the semantics of the PQP: it locates an event or state prior to a past time reference point (which may be given by a temporal adverbial or by another past time event or state).

Synchronous relations were most frequently inferred in the presence of two verb tenses: an IMP (in the IMP-IMP, IMP-PS and IMP-PQP combinations) and a PRES (in the IMP-PRES, PS-PRES, PQP-PRES, PRES-IMP and PRES-PRES combinations). These findings strongly confirm the prediction that the IMP plays a significant role in expressing synchronicity of situations, and highlight the role played by the PRES.

Finally, in our data, backward temporal relations are most frequently expressed in the presence of the PQP, when it was preceded by an IMP, a PS or a PRES form. As mentioned above, the PQP-PQP combination did not signal backwards relations most frequently, but chronological relations. Backwards relations are less frequently expressed by other than PQP verb tense combinations, such as PS-IMP or PRES-IMP. These results strongly support the significant role of the PQP to express backwards relations, which may also explain the high rate of implicit backwards relations found (Figure 4). As was the case for our English data, the fact that backwards relations number fewer than chronological and synchronous relations prevented us from including them in the statistical modelling analysis, in which we tested the roles of verb tense, marking status, stylistic register and language status in predicting chronological vs. synchronous relations.

The first factor we tested in the data with a binary outcome (chronological vs. synchronous relation) was *marking status*. Recall that the application of the *implicitness assumption* to temporal relations triggers the prediction that chronological relations are more frequently left implicit (because they are highly expected) whereas synchronous relations are more frequently overtly marked (because they are not). Our results do not confirm this prediction, as the explicit vs. implicit status of chronological relations depends on which overt markers are taken into consideration. On the one hand, we found that overall chronological relations are overtly marked in 52% of cases and left implicit in 48% of cases (Figure 4). On the other hand, the mixed effects model (section 3.3.2.2) showed that the presence of an overt marker, and more specifically of a specialized connective, triggers chronological relations more frequently. The underspecified *et* is used to express chronological and synchronous relations with equal frequency. Figure 4 shows that specialized connectives are used in 19% of cases to overtly mark a chronological relation vs. 0% in the case of synchronous relations. Since the underspecified *et* is ambiguous between chronological and synchronous relations, the authentic rate of overtly marked chronological relations is revealed by the rate of specialized connectives employed to express these relations.

The second factor we tested was the role of the verb tense from S1 and S2 to signal chronological vs. synchronous relations. We found that the verb tense from S2 influences the temporal relation to be inferred: the PS occurs frequently with chronological relations, whereas the IMP occurs frequently with synchronous relations. This finding confirms the results shown by the descriptive quantitative analyses we carried out, which showed that chronological, synchronous and backwards temporal

relations are favoured by certain verb tense combinations, but that these combinations are not exclusive. As stated above, chronological relations not only occur frequently in the absence of the PS but also in the presence of the PQP or the PRES when they are preceded by a PQP. Likewise, synchronous relations are frequently inferred in the absence of an IMP but in the presence of a PRES (use in S1 or in S2).

Finally, we tested whether the role of obligatory and optional overt marking of temporal relations depends on the text genre (literary vs. journalistic) and the language status (French original vs. French translated). As a control for our study, we found there were more chronological relations in the literary corpus than in the journalistic one, the former involving by definition the narration of events in a chronological manner. Our data showed that the status of optional overt marking of temporal relations did not significantly change between literary and journalistic texts, with one exception: there are fewer cases of chronological relations overtly marked with a specialized connective in literary texts than there are in journalistic texts. We believe that this result is again explained by the typicality of literary texts: readers know that a story is being told, and that situations are narrated in a chronological way. As such, the need to mark them overtly with a specialized connective is less pressing than in journalistic texts (which do not raise this expectation). Furthermore, we also found that the rate of overtly marked chronological relations is similar in texts translated into French and in texts originally written in French. This result defies the prediction of the translation universals, according to which there is a higher rate of *explicitation* (hence, overt marked relations) in translated texts than original texts.

4 General discussion

Our purpose in this study was to investigate the role of linguistic cueing (obligatory and optional linguistic overt marking) of temporal relations in the light of non-linguistic biases. Building on previous research, we expected to find that verb tenses inform language users about the temporal relation expressed, and that the implicitness or absence of overt marking of the relation is a sign of expectation of that relation: if language users have a default preference for chronological relations, they should tend to appear without optional overt markers. In contrast, relations which are not highly expected (synchronous and backwards) should tend to appear with optional overt markers, such as connectives. Moreover, we expected that text genre affects the inference of temporal relations: chronological relations are more frequently expected in narrative texts than in journalistic texts. Finally, if translators tend to explicitate more in translated texts than in original texts, all three temporal relations should tend to be overtly marked in translated texts more frequently than in original texts. Table 3 summarizes the results of the descriptive quantitative and statistical modelling analyses carried out on the English and French sets of data with respect to these predictions (also formulated in Table 1).

Table 3: Summary of results from English and French data in terms of frequency

Predictions	Results in English	Results in French
VT		
– referentially autonomous and perfective VTs signal chronological relations	✓ more chronological with SP in S2 and the SP-SP combination	✓ more chronological with PS in S2 or PQP in S1
– referentially non-autonomous and imperfective VTs signal synchronous relations	✓ more synchronous with the progressive <i>-ing</i> , Spres or PresPerf in S1 or in S2	✓ more synchronous with IMP or PRES in S1 or S2
– pluperfect signals backwards relations	✓ more backwards with PP in S2	✓ more backwards with PQP in S2 or PRES-IMP

Marking status – chronological relations are overtly marked with specialized connectives more frequently than with underspecified connectives	✗ chronological most frequently marked with <i>and</i>	✓ chronological most frequently marked with specialized connectives
Text genre – more chronological relations in literary texts than in journalistic texts – in literary texts, chronological relations are overtly marked less frequently than left implicit	✓ more chronological in literary texts ✓ more chronological overtly marked with <i>and</i> in literary texts than in journalistic texts ✓ less chronological overtly marked with specialized connectives or left implicit in literary texts than in journalistic texts	✓ more chronological relations in literary texts ✗ similar frequency of chronological overtly marked with <i>and</i> or left implicit in literary texts and in journalistic texts ✓ less chronological overtly marked with specialized connectives or left implicit in literary texts than in journalistic
Text status – more overtly marked chronological relations in translated texts than in original texts	✗ similar rates of overt chronological in translated and original English texts ✗ similar rates of chronological relations overtly marked with <i>and</i> and with specialized connectives in translated and original English texts	✗ similar rates of overt chronological relations in translated texts than in original French texts ✗ similar rates of chronological relations overtly marked with <i>et</i> and with specialized connectives in translated and original French texts

With respect to the role of verb tenses, the descriptive quantitative analyses showed that patterns do emerge, despite the great variability when it comes to which verb tense combination expresses which temporal relation. These patterns were either confirmed or refined by the statistical modelling analysis. Our results confirm that verb tenses inform language users about the temporal relation expressed in the text. Our original data reveal that, in both English and French, it is the verb tense from S2 which plays a significant role. We find that the presence of a pluperfect (English PP and French PQP) in S2 signals the presence of backwards relations; that the presence of a simple past (English SP and French PS) in S2 signals chronological relations; and that the presence of a present time verb tense (English Spres or PresPerf and French PRES) signals towards synchronous relations. Finally, we find that the imperfective (English progressive *-ing* and French IMP) signals the presence of synchronous relations. Importantly, our corpus analysis revealed that not all verb tenses containing the *-ing* morpheme signal synchronous relations. In particular, this is true only of the progressive *-ing*; the gerund *-ing* signals chronological and synchronous relations with equal frequency. So, cross-linguistically speaking, imperfective verb tenses (English progressive *-ing* and French IMP) are significantly more frequently used to express synchronous relations (78% for the former and 91% for the latter) than they are used to express chronological relations (22% for the former and 9% for the latter).

Regarding the role of overt marking status, our results from the English and French sets of data do not confirm the prediction of the **implicitness assumption** applied to temporal relations. In fact, we find the opposite result: chronological relations are more frequently overtly marked with a specialized connective, whereas synchronous and backwards relations are more frequently left implicit. Our study reveals that specialized and non-specialized connectives do not have the same “weight” when it comes

to overtly marking a temporal relation, and that this is language-dependent. Chronological relations in English are most frequently overtly marked with the underspecified *and* (and the specialized connectives are less frequent); in French, they are most frequently overtly marked with the specialized connectives (and the underspecified connective *et* is used to express chronological and synchronous relations with equal frequency). These results confirm Grisot and Blochowiak's (2019) experimental findings: processing effort of short narrative texts was smaller when the chronological relation was overtly marked with a specialized connective than when it was left implicit, and comprehenders found overt marking of the chronological relation more acceptable than leaving it implicit. Furthermore, these results reveal that the semantics and pragmatics of the conjunction differ between English and French (cf. also Blochowiak & Castelain's 2018 experimental study on French *et*).

Given our corpus-based results and related experimental results from the literature (discussed in sections 2.1, 2.2 and 2.3), three possible explanations can be formulated:

- If the implicitness assumption (i.e. the *implicitness of the discourse connective as a sign of expectation of the discourse relations*) applies to temporal relations, our results would suggest that the temporal relations, which are highly expected, are synchronous and backwards relations, and not chronological relations.
- Chronological relations are not different from synchronous and backwards relations with respect to their “highly expected” status. Because of the obligatory marking (i.e. verb tenses and aspect), language users have equal expectations and are oriented to one relation or the other by the optional linguistic cues they encounter (specialized and underspecified connectives, and other types of overt marking) or by their world knowledge. So, at the sentence level, there is no expectation other than what is triggered by the verb tense and aspect.
- Given the particular status of temporal relations, language users only develop expectations at the level of text genre and not at the sentence level. As such, our results are compatible with the text genre assumption.

The first explanation of our results is nonetheless unlikely from experimental and frequency perspectives. Indeed, there are no experimental studies which would support such a view. In fact, these studies show that linguistic factors and world knowledge drive comprehenders' processes of inferring temporal relations. Moreover, this view is not compatible with frequency data (as found in our study; cf. also Author1, under review) demonstrating that backwards relations are significantly less frequent than synchronous and chronological ones. Thus, it would be unfounded to assume that backwards relations are highly expected, as this would mean that infrequent relations (like backwards) are highly expected by language users in the interpretation process.

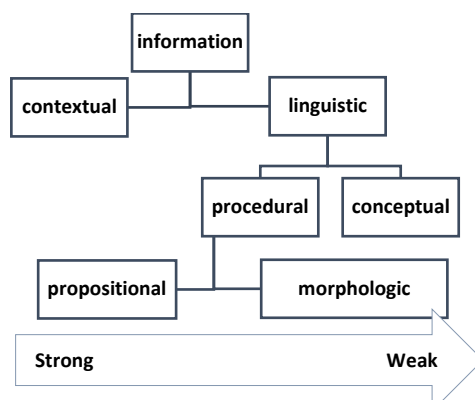
The second explanation draws on Grisot & Blochowiak's (2019) proposal which states that “the more accurate way to explain the processing of temporal relations is to take into consideration a cluster of factors including world knowledge (rules, laws and schemes) and linguistic knowledge (connectives, tenses, grammatical and lexical aspect)” (2019: section 5.2). In short, Grisot & Blochowiak suggest that comprehenders' world knowledge of a chronological order of the events described (for instance, in *The plane landed. The passengers got off*) is what drives their interpretation of this sequence of sentences. If we apply this idea to the other types of temporal relation, the synchronous interpretation of a series of sentences (as in *Mary vacuumed and listened to pop music*) also draws principally on our world knowledge that people often vacuum while listening to music. The same principle applies to the backwards interpretation of *Mary arrived late at work. Her car did not*

start. These temporal relations (i.e. chronological, synchronous and backwards respectively) are the most relevant ones when interpreting these three series of utterances. Here, the notion of *relevant interpretation* is understood in the relevance-theoretic pragmatic framework as the interpretation which keeps the balance between (low) cognitive effort and (high) cognitive effects. Importantly, in all three examples, the situations are expressed in the simple past, which, according to our results, should impose a chronological interpretation. What these examples show is that linguistic cues, i.e. verb tense and grammatical aspect, may be overridden by world knowledge (see also Lascarides & Asher 1993; Moeschler 2002). At the same time, the interpretation drawing on world knowledge could be changed in very specific contexts — if the plane had technical problems, and the passengers had to get off in an unusual manner before the plane's actual landing (such as via the slides). The deviation from this interpretation may be marked with an optional overt marker (e.g. *before* or *an hour ago* for a backwards relation, and *at the same time* or *when* for a synchronous relation). Most importantly, these alternative interpretations need to be supported by the context, in which comprehenders formulate the contextual hypotheses that will validate the most relevant interpretation.

As such, we suggest that language users have equal expectations of temporal relations, and lean towards one or the other according to the optional linguistic cues they encounter (specialized and underspecified connectives, and other types of overt marking) or their world knowledge. This is due to the obligatory marking (i.e. verb tense and aspect), which provides a baseline for inferring temporal relations: the presence of a simple past most frequently signals chronological relations; a pluperfect tense most frequently signals backwards relations; and the imperfective (the progressive *-ing* and the French IMP) most frequently signals synchronous relations (as shown by the probabilities found in the mixed models fitted on our two sets of data). When no specific rules, laws or schemas (from the comprehenders' world knowledge) apply, and in the absence of temporal overt markers, the lexical aspect of the verb phrase (*states*, *activities*, *achievements* and *accomplishments*, as defined by Vendler 1957) plays a significant role in inferring the temporal relation (as suggested by Partee 1984; Dowty 1986; Kamp & Reyle 1993, and confirmed experimentally by Gennari 2004; Therriault & Raney 2007; Becker et al. 2013; Dery & Koenig 2015; Author1, under review). In particular, comprehenders most frequently infer synchronous relations with states (Gennari 2004) and activities (Becker et al. 2013; Author1, under review), and chronological relations with events and accomplishments (Becker et al. 2013; Author1, under review).

Therefore, it is possible that language users have equal expectations of the three types of temporal relations, which they infer, during the incremental comprehension process (Altmann 1988; Crain & Steedman 1985; Kehler & Rohde 2017), on the basis of their world knowledge and the linguistic cues they are exposed to (the semantics of the verb, the verb tense, grammatical aspect, temporal overt markers and the lexical aspect of the verb phrase). Nevertheless, a definite order in which these various factors come into play has not been established experimentally. We have seen that there are cases when world knowledge overrides linguistic information, and cases when the different types of linguistic information override world knowledge. In a relevance-theoretic pragmatic framework, Moeschler (2002) makes an interesting and potentially very useful theoretical proposal on the scale of strength of these factors (provided in Figure 5).

Figure 5 Hierarchy of non-linguistic and linguistic information (Moeschler 2002)



Specifically, he distinguishes between *contextual information* (i.e. contextual hypotheses triggered by comprehenders' world knowledge), linguistic cues encoding *propositional procedural* information (such as connectives), linguistic cues encoding morphologic procedural information (such as grammatical aspect) and linguistic cues encoding conceptual information (such as lexical aspect). For him, world knowledge is stronger than and overrides linguistic information, wherein linguistic procedural information is stronger and overrides linguistic conceptual information. Within procedural information, propositional information is stronger and overrides morphologic information. We believe that this pragmatic model has the potential to refine further the cluster of features relevant for inferring temporal relations, and to explain the corpus distribution of obligatory and optional markers found in the current study. We leave this refinement for future work.

The third explanation focuses on expectations that language users may have about temporal relations that come from text genre. Earlier research has found that text genre increases expectations regarding the discourse relations to be inferred (Sanders 1997; Sagi 2006); our investigation shows that these expectations also apply to temporal relations. In our data, annotators found more chronological relations in literary texts than in journalistic texts did (as shown in Table 3), in English and French alike. Unconfirmed, however, was the general prediction that chronological relations are more highly expected in literary texts (due to the text genre bias) and should thus be most frequently left implicit. Instead, we found finer-grained differences with respect to the distinction between specialized and underspecified connectives. In both English and French data, we found that the presence of a specialized connective correlates with higher rates of chronological relations. Crosslinguistic dissimilarities were found with respect to the role of the underspecified connective: the use of English *and* increases the rate of chronological relations more in literary texts than journalistic texts, whereas in the use of French *et* to mark chronological vs. synchronous relations is similar in literary and journalistic texts. Additionally, Author1's (under review) experimental study showed that text type (short narratives vs. transfer-of-possession contexts) influences comprehenders' expectations of which temporal relation to infer (chronological vs. synchronous) between two situations performed by the same agent or by different agents. In particular, Author1 (under review) finds that situations performed by the same agent are more frequently interpreted chronologically in short narratives, but synchronously in transfer-of-possession contexts.

Furthermore, we tested whether obligatory and optional overt marking of temporal relations varies with respect to the translation process. Our data do not provide evidence to support Baker's (1993) translation universal regarding the higher degree of explicitation in translated texts than in

original texts. Both in the English and the French sets of data, we find similar rates of overt marking (with underspecified or specialized connectives) in translated and original texts.

5 Conclusion

In this study, we addressed three questions. The first was *whether specific verb tenses and the marking status of a relation (implicit, marked with underspecified connectives, or marked with specialized connectives) are significant predictors for one type of temporal relation or another*. We answer this question in the affirmative. First, we show that there are patterns linking certain verb tenses with certain temporal relations (i.e. the simple past is linked to chronological relations, the imperfective to synchronous relations and the pluperfect to backwards relations); we also show that the relation between a verb tense and a specific temporal relation is not a one-to-one relation. This means that the presence of a certain verb tense does not exclusively signal one type of temporal relation (e.g. the presence of a simple past does not always signal a chronological relation), and a certain type of temporal relation may also arise in the absence of its “preferred” verb tense (e.g. chronological relations also arise in the absence of a simple past). Second, the distinction between implicit vs. overt marking of temporal relations is too coarse grained. In particular, we found differences with respect to whether the temporal relation is overtly marked with the underspecified conjunction, such as English *and* and French *et*, or with specialized connectives. These differences are language-specific. On the one hand, English *and* has a strong preference for chronological relations; French *et* is less restrictive, allowing chronological and synchronous relations with equal frequency. On the other hand, underspecified *and* is the favoured means of marking chronological relations in English, whereas specialized connectives take the upper hand over the conjunction *et* in French.

The second question was *whether chronological relations are more frequently expressed implicitly than overtly marked with a discourse connective, and, conversely, whether synchronous and backwards relations are more frequently expressed with overt marking than left implicit*. We answer this question in the negative, and suggest that chronological relations do not have a different status from synchronous or backwards ones. In fact, we did not find empirical evidence that chronological relations are more frequently left implicit because they are highly expected due to cognitive- or conversation-related principles, as stated in the implicitness assumption. We believe that language users rely on linguistic cues (both obligatory and optional), as well as on their world knowledge, to infer the temporal relations expressed. Future work is required to investigate this issue by means of experiments.

The third question was *whether the answers to the first two questions depend on the text genre and the text status*. We answer the first part of this question in the affirmative, endorsing the text genre assumption, as we find that *text genre* had an effect on the distribution of the different temporal relations investigated. In the absence of default expectations of chronological relations arising at the sentence level, our study shows that it is at the level of text genre (literary vs. journalistic) that chronological relations are more frequent, and thus, maybe more expected by language users. The question of language users’ expectations should be further refined and tested in an experimental study. With respect to the second part, our data did not provide evidence supporting the text-status assumption, according to which temporal relations should have been more frequently overtly marked than left implicit in the translated texts compared to the original texts. Furthermore, the analysis of text genre confirmed the above-mentioned dissimilarities between the role of English and French underspecified and specialized connectives to mark temporal relations overtly.

Grisot, C., & Blochowiak, J. (2021). Temporal Relations at the Sentence and Text Genre Level: The Role of Linguistic Cueing and Non-linguistic Biases—An Annotation Study of a Bilingual Corpus. *Corpus Pragmatics*, 1-41.

The authors do not have any conflict of interest to declare.

References

- Adam, J.-M. (1992). *Les Textes: textes et prototypes*. Paris: Nathan-Université.
- Altmann, G. (1988). Interaction with context during human sentence processing. *Cognition*, (30), 191–328.
- Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
- Asr, F. A. T. O., & Demberg, V. E. (2012). Implicitness of discourse relations. In *Proceedings of COLING* (pp. 2669–2684). Presented at the COLING 2012, Mumbai.
- Asr, F. T., & Demberg, V. (2015). Uniform surprisal at the level of discourse relations: Negation markers and discourse connective omission. In *Proceedings of the 11th international conference on computational semantics* (pp. 118–128).
- Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge: Cambridge University Press.
- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 233–253). Amsterdam: John Benjamins.
- Becker, R. B., Ferretti, T. R., & Madden-Lombardi, C. J. (2013). Grammatical aspect, lexical aspect, and event duration constrain the availability of events in narratives. *Cognition*, 129(2), 212–220.
- Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford: Blackwell. Accessed 8 October 2014
- Blakemore, D., & Carston, R. (1999). The pragmatics of and-conjunctions: The non-narrative cases. *CILISC1 (L'économie dans les structures, les computations et l'utilisation du langage)*, 12–15.
- Blochowiak, J., & Castelain, T. (2018). How logical is natural language conjunction? An experimental investigation of the French conjunction et. In P. Saint-Germier (Ed.), *Language, Evolution and Mind: Essays in Honour of Anne Reboul*. (pp. 97–125). College Publications.
- Britton, B. K., Glynn, S. M., Meyer, B. J., & Penland, M. J. (1982). Effects of text structure on use of cognitive capacity during reading. *Journal of Educational Psychology*, 74(1), 51.
- Carston, R. (1993). Conjunction, explanation and relevance. *Lingua*, 90(1), 27–48.
- Carston, R. (1998). Conjunction and pragmatic effects. In R. Asher (Ed.), *The Encyclopedia of Language and Linguistics* (pp. 692–698). Oxford: Pergamon Press.
- Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.
- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the syntactic processor. In D. R. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational and theoretical perspectives* (pp. 320–358). Cambridge: Cambridge University Press.
- Croft, W. (1990). *Typology and Universals*. Cambridge: Cambridge University Press.
- Dalgaard, P. (2010). R Development Core Team (2010): R: A language and environment for statistical computing. <https://research.cbs.dk/en/publications/r-development-core-team-2010-r-a-language-and-environment-for-sta>.
- Das, D., & Taboada, M. (2013). Explicit and implicit coherence relations: A corpus study. In *Proceedings of the 2013 annual conference of the Canadian Linguistic Association*. University of Victoria Victoria.
- Das, D., & Taboada, M. (2018). RST Signalling Corpus: A corpus of signals of coherence relations. *Language Resources and Evaluation*, 52(1), 149–184.
- Deaton, J. A., & Gernsbacher, M. A. (1997). Causal conjunctions and implicit causality cue mapping in sentence comprehension. *Journal of Memory and Language*, 33, 128–147.
- Dery, J. E., & Koenig, J.-P. (2015). A narrative-expectation-based approach to temporal update in discourse comprehension. *Discourse Processes*, 52(7), 559–584.

- Grisot, C., & Blochowiak, J. (2021). Temporal Relations at the Sentence and Text Genre Level: The Role of Linguistic Cueing and Non-linguistic Biases—An Annotation Study of a Bilingual Corpus. *Corpus Pragmatics*, 1-41.
- Dowty, D. 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.
- Dowty, D. R. (1986). The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and philosophy*, 9(1), 37–61.
- Ducrot, O. (1979). L'imparfait en français. *Linguistische Berichte Braunschweig*, (60), 1–23.
- Dyvik, H. (1998). A translational basis for semantics. In S. Johansson & S. Oksefjell (Eds.), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies* (pp. 51–86). Amsterdam: Rodopi.
- Evers-Vermeul, J., Hoek, J., & Scholman, M. C. (2017). On Temporality in Discourse Annotation: Theoretical and Practical Considerations. *Dialogue & Discourse*, 8(2), 1–20.
- Feller, D. P., Eerland, A., Ferretti, T., & Magliano, J. P. (2019). Aspect and narrative event segmentation. *Collabra: Psychology*, 5(1).
- Field, A., Miles, J., & Field, Z. (2014). *Discovering Statistics Using R* (2nd ed.). New York: SAGE Publications.
- Gellerstam, M. (1996). Translations as a source for cross-linguistic studies. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.), *Languages in contrast: papers from a symposium on text-based cross-linguistic studies* (pp. 53–62). Lund: Lund University Press.
- Gennari, S. P. (2004). Temporal references and temporal relations in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(4), 877.
- Grice, H. P. (1975). Logic and Conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics* (Vol. 3: Speech Acts). New York: Academic Press.
- Grisot, C. (2017). Tense, Grammatical Aspect and Subjectivity: An Experimental Study Using Inter-Annotator Agreement Rates and Corpus-Based Data. *Corpus Pragmatics*, 1–29.
- Grisot, C. (2018). *Cohesion, coherence and temporal reference from an experimental corpus pragmatics perspective*. Cham: Springer.
- Grisot, C., & Blochowiak, J. (2019). Temporal connectives and verbal tenses as processing instructions: Evidence from French. *Pragmatics & Cognition*, 24(3), 404–440.
- Haberland, K. (1982). Reader expectations in text comprehension. In J. F. Le Ny & W. Kintsch (Eds.), *Language and language comprehension* (pp. 239–249). Amsterdam: North-Holland.
- Hinrichs, E. (1986). Temporal anaphora in discourses of English. *Linguistics and philosophy*, 9(1), 63–82.
- Hoek, J., Zufferey, S., Evers-Vermeul, J., & Sanders, T. J. M. (2018). The linguistic marking of coherence relations: Interactions between connectives and segment-internal elements. *Pragmatics & Cognition*, 25(2), 276–309. <https://doi.org/10.1075/pc.18016.hoe>
- Hopper, P. J. (1979). Aspect and foregrounding in discourse. In T. Givón (Ed.), *Discourse and syntax* (pp. 213–243). New York: Academic Press.
- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context* (pp. 11–42). Washington DC: Georgetown University Press.
- House, J. (2018). *Translation: The Basics*. New York: Routledge.
- Johansson, S. (1998). On the role of corpora in cross-linguistic research. In S. Johansson & S. Oksefjell (Eds.), *Corpora and Cross-linguistic Research: Theory, Method and Case Studies* (pp. 3–24). Amsterdam: Rodopi.
- Johnson, K. (2008). *Quantitative Methods In Linguistics*. Malden MA: Wiley-Blackwell.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge MA: Harvard University Press.
- Kamp, H. (1979). Events, instants and temporal reference. In R. Bauerle, U. Egli, & A. von Stechow (Eds.), *Semantics from different points of view* (pp. 376–418). Amsterdam: Springer.
- Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Amsterdam: Springer.
- Kamp, H., & Rohrer, C. (1983). Tense in Texts. In R. Bauerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language* (pp. 250–269). Berlin: Walter de Gruyter.

- Grisot, C., & Blochowiak, J. (2021). Temporal Relations at the Sentence and Text Genre Level: The Role of Linguistic Cueing and Non-linguistic Biases—An Annotation Study of a Bilingual Corpus. *Corpus Pragmatics*, 1-41.
- Kehler, A., & Rohde, H. (2017). Evaluating an expectation-driven question-under-discussion model of discourse interpretation. *Discourse Processes*, 54(3), 219–238.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kleiber, G. (2003). Entre les deux mon cœur balance ou l'imparfait entre aspect et anaphore. *Langue française*, 8–19.
- Klein, W. (1994). *Time in language*. London and New York: Routledge.
- Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-Based Models of Language*. Stanford: CSLI Publications.
- Lascarides, A., & Asher, N. (1993). Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and philosophy*, 16(5), 437–493.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge MA: MIT Press. Accessed 9 February 2015
- Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. *Advances in neural information processing systems*, 19, 849.
- Luscher, J.-M., & Moeschler, J. (1990). Approches dérivationnelles et procédurales des opérateurs et connecteurs temporels: les exemples de et et de enfin. *Cahiers de linguistique française*, 11, 77–104.
- Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & cognition*, 31(5), 663–672.
- Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse processes*, 29(2), 83–112.
- Moeschler, J. (2000). Le modèle des inférences directionnelles. *Cahiers de linguistique française*, 22, 57–100.
- Moeschler, J. (2002). Economy and pragmatic optimality: the case of directional inferences. *Generative Grammar Geneva*, 3, 1–20.
- Molendijk, A. L. (1990). *Le passé simple et l'imparfait: une approche reichenbachienne*. Amsterdam: Rodopi.
- Murray, J. D. (1997). Connectives and narrative text: The role of continuity. *Memory & Cognition*, 25(2), 227–236.
- Nerbonne, J. (1986). Reference Time and Time in Narration. *Linguistics and Philosophy*, 9(1), 83–95.
- Partee, B. (1984). Nominal and Temporal Anaphora. *Linguistics and Philosophy*, 7(3), 243–286.
- Partee, B. (1973). Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy*, 601–609.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A. K., & Webber, B. L. (2008). The Penn Discourse TreeBank 2.0. In *Proceedings of the 6th International conference on Language Resources and Evaluation*. Presented at the LREC 6, Marrakech. <http://www.lrec-conf.org/proceedings/lrec2008/>.
- Reichenbach, H. (1947). *Elements of symbolic logic*. New York: Mcmillan.
- Sagi, E. (2006). Context and the processing of discourse: Priming and genre effects on discourse comprehension. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 28).
- Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse processes*, 24(1), 119–147.
- Sanders, T. (2005). Coherence, causality and cognitive complexity in discourse. In *Proceedings SEM-05, First International Symposium on the exploration and modelling of meaning* (pp. 105–114).
- Sanders, T. J., Spooren, W. P., & Noordman, L. G. (1992). Toward a taxonomy of coherence relations. *Discourse processes*, 15(1), 1–35.
- Saussure de, L. (2003). *Temps et pertinence: éléments de pragmatique cognitive du temps*. Bruxelles: De Boeck Duculot.
- Saussure, L. de, & Sthioul, B. (2002). Interprétations cumulative et distributive du connecteur et: temps, argumentation, séquencement. *Cahiers de linguistique française*, 24, 293–314.

- Grisot, C., & Blochowiak, J. (2021). Temporal Relations at the Sentence and Text Genre Level: The Role of Linguistic Cueing and Non-linguistic Biases—An Annotation Study of a Bilingual Corpus. *Corpus Pragmatics*, 1-41.
- Segal, E. M., Duchan, J. F., & Scott, P. J. (1991). The role of interclausal connectives in narrative structuring: Evidence from adults' interpretations of simple stories. *Discourse processes*, 14(1), 27–54.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.
- Spooren, W. (1997). The processing of underspecified coherence relations. *Discourse processes*, 24(1), 149–168.
- Tasmowski-De Ryck, L. (1985). L'imparfait avec et sans rupture. *Langue française*, 59–77.
- Therriault, D. J., & Raney, G. E. (2007). Processing and representing temporal information in narrative text. *Discourse Processes*, 43(2), 173–200.
- Vendler, Z. (1957). Verbs and times. *The philosophical review*, 143–160.
- Wilson, D., & Sperber, D. (1998). Pragmatics and time. In R. Carston & S. Uchida (Eds.), *Relevance Theory: Applications and implications* (pp. 1–22). Amsterdam: John Benjamins.
- Wilson, D., & Sperber, D. (2012). *Meaning and relevance*. Cambridge: Cambridge University Press.
- Zeevat, H., & Jasinskaja, K. (2007). And as an additive particle. *Language, representation and reasoning. Memorial volume to Isabel Gómez Txurruka*, 315–340.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, 123(2), 162.