

Looking for the causal values of *as* and *since* in large corpora, and how these values compare with each other

*À la recherche de la valeur causale des conjonctions *as* et *since* dans de grands corpus, dans le but de les comparer*

Bénédicte Guillaume



Electronic version

URL: <http://journals.openedition.org/corpus/3628>

ISSN: 1765-3126

Publisher

Bases ; corpus et langage - UMR 6039

Electronic reference

Bénédicte Guillaume, « Looking for the causal values of *as* and *since* in large corpora, and how these values compare with each other », *Corpus* [Online], 18 | 2018, Online since 09 July 2018, connection on 08 September 2020. URL : <http://journals.openedition.org/corpus/3628>

This text was automatically generated on 8 September 2020.

© Tous droits réservés

Looking for the causal values of *as* and *since* in large corpora, and how these values compare with each other

*À la recherche de la valeur causale des conjonctions *as* et *since* dans de grands corpus, dans le but de les comparer*

Bénédicte Guillaume

This paper incorporates a few parts from a presentation given during the linguistics workshop at the annual Société des Anglicistes de l'Enseignement Supérieur Conference (Université du Sud-Toulon Var, 4th-6th June 2015). I wish to thank the organisers D. Boulonnais, L. Dufaye and G. Girard, and more particularly P. Miller for putting me in touch with anonymous reviewers regarding this very early stage of the present work, whose feedback has helped it evolve into the distinctly different present version. I also wish to thank my colleagues from the UMR 7320 Bases, Corpus, Langage CNRS research laboratory, J.-M. Merle and D. Mayaffre, for their advice and encouragement regarding the present paper. Still, I remain of course solely responsible for any mistakes or inconsistencies.

Introduction

- 1 Among the subordinators which are apt to introduce an adverbial clause expressing cause in English, it is not always easy to differentiate between *since* and *as* on a semantic level. For instance, in the following attested examples, which are taken from the *British National Corpus* (BNC),¹ replacing *as* by *since*, or the opposite, can hardly be said to alter the original meaning of the sentences in question in any perceptible way:²

(1) CER599 A cloud chamber is good for detecting low energy, electrically charged particles **as they leave behind a trail of liquid drops** as they pass through the mist in the chamber.

→ A cloud chamber is good for detecting low energy, electrically charged particles **since they leave behind a trail of liquid drops...**

(2) ASL1175 **Since the nucleus of each cell contains the same genetic information,** it is the reciprocal communication between nucleus and cytoplasm during development that determines which proteins are made.

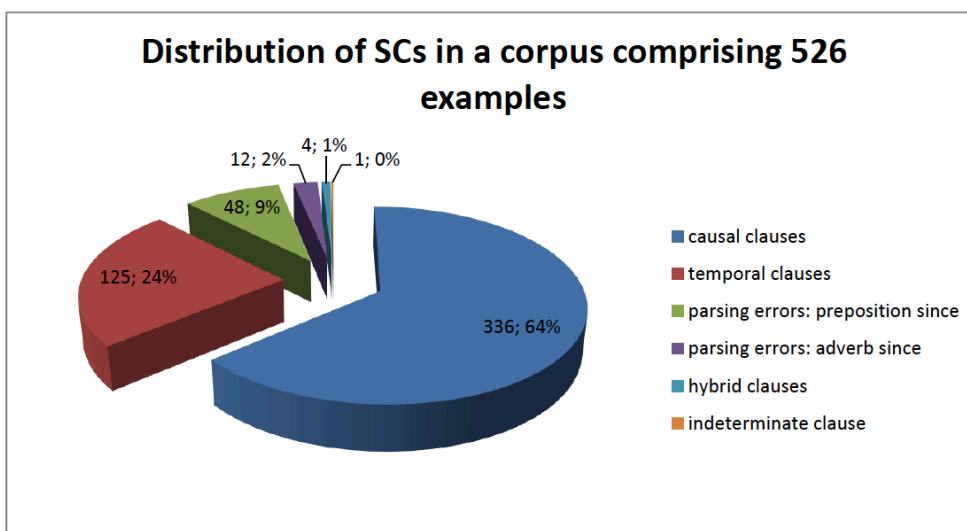
→ **As the nucleus of each cell contains the same genetic information,** it is the reciprocal communication...

- 2 Despite the fact that the two markers are so close to each other in their causal uses, they have different origins and also differ in their other possible uses. I shall therefore first attempt to examine their causal use in the light of their other differences. For that purpose, I shall need to resort to a large number of attested examples, mostly causal, but possibly also examples of other values of these markers as well. However, the various acceptations of *since* and more particularly of *as* makes it time-consuming to harvest examples of one value in particular from large electronic corpora such as the BNC or the COCA (*Corpus of Contemporary American English*), hence the varied approaches advocated in this paper for finding relevant examples.
- 3 When it comes to positing a theoretical difference between the two markers, the chosen framework is A. Culioli's utterer-centred approach. The help of A. Culioli's linguistic tools as well as the observation of my corpus will lead to the formulation of a hypothesis regarding the genres in which each subordinator is more likely to be encountered, which I shall then put to the test in the Intercorp³ multilingual parallel corpus.

1. Collecting examples of the causal values of *since* and *as*

- 4 The closeness between the two markers in their causal use is not surprising, because *as* and *since* have quite a few things in common. Not only are they both apt to take on different interpretations according to contextual variations, but their causal use derives accidentally, so to speak, from their core value.
- 5 Regarding the various uses of both markers, *as* has many more possible uses than *since*, as we shall see. Indeed, *since* can have a causal or a temporal interpretation, with a few cases of overlap. As far as *since*-clauses (hereafter SCs) are concerned, I have sampled a corpus containing more than 500 examples from the BNC, 64% of which happen to be causal subordinates, while the rest essentially corresponds to temporal ones, if one leaves aside the parsing errors (*cf.* Guillaume 2014: 13):

A. Distribution of SCs in a corpus comprising 526 examples



- 6 *Since* was originally a temporal marker, not a causal one. R. Molencki explains that *since* is, among others, a cognate of German *seit*, which can be considered as one of the present-day equivalents of temporal *since* (2007: 98). The emergence of the causal use for the conjunction first appeared in Middle English, and gradually gained momentum until the 16th century, when the causal interpretation became prevalent, as compared to the temporal one (2007: 110). It is certainly because *since* primarily refers to the anteriority of an event that at some point it was used in order to express a cause and effect relationship.
- 7 Apart from the temporal and causal SCs, there is a third category which is worthy of interest, that of hybrid clauses, that is clauses in which the causal and the temporal meaning are so closely intertwined with each other that it becomes almost impossible to decide whether the clause should be classified as causal or as temporal, while its formal characteristics (such as the choice of tense, aspect and modality in the subordinate as well as the main clause, the presence of a comma...), which usually provide clues in order to differentiate between the two categories, are not determinative enough either (cf. Guillaume 2012):
- (3) EVS1821 The Catholic religion has marginalized women, in part, one might say, **since it was decided that God was male.**
- (4) H8H390 At fifty-three, her father was still an intensely busy man, head of an important software company, and lately dabbling in property speculation. He had the personality for it, strong, aggressive and with an eye to a bargain. Very few people ever saw his softer side, and she was one of them. He had lavished all his love on her, more so since her mother had died almost eighteen years ago, and, similarly, he was the only one who ever saw her vulnerable side.
- (5) She wanted to take over the housekeeper's post, you know. Apparently she had run your establishment and, I should imagine, well, and I'm sure your own staff obeyed and respected her, but you see, our people have worked under different rules for years, and I won't say they ignored her commands, but they went about things as they always had done. And they have managed very well, too, I would say exceedingly well since the staff was depleted some years ago. (Catherine Cookson, *The Black Candle*, London: Corgi Books, 1989: 257)
- 8 The above examples display a mixture of grammatical markers that can be traced back to the temporal value, blended with features more typical of the causal value,⁴ while

the event described in the subordinate corresponds to a starting-point which is also the direct cause of the state of affairs related in the main clause. Nevertheless, as can be seen in the above chart, there are only few such examples. Therefore, when resorting to large online corpora such as the BNC, there are good chances of easily finding lots of relevant examples of purely causal SCs, provided one specifies that the results should be limited to *since* used as a subordinator, not as a preposition or an adverb.⁵ It also turns out that my searches in the BNC returned twice as many examples of causal SCs as it did of temporal ones, but this may be due to the fact that, unlike the COCA for instance, the BNC is overwhelmingly a written corpus,⁶ which therefore favours the presence of causal SCs, a phenomenon connected to a rather elevated register.⁷

- 9 When it comes to finding examples of causal *as*- clauses (hereafter ACs), however, the difficulty of the task changes drastically, mostly because *as* has a much larger semantic range than *since*, which, after all, has only two main interpretations, with very few cases of ambiguity. As for *as*, it is first and foremost a marker of identity (cf. Deléchelle 1980; Culioli 1999), but C. Guimier (1997), for instance, lists no fewer than six main values, all linked to the core value of identification: identity of manner, comparison, concord, conformity, temporal value, and causal value. Even though the boundaries of the categories put forward in Guimier's classification can be questioned, especially since some of them seem to be extremely close to one another, this certainly gives a good idea of the large semantic scale covered by *as* as a subordinator. Besides, the different values tend to overlap a lot more than with *since* (cf. Claridge and Walker 2001: 40); in the same manner as the causal value of *since* derives from its capacity to signal anteriority, the causal value of *as* seems to be the logical outcome of other more common uses of this conjunction, such as identity and concomitance:

(6) I built it up, higher and higher. When I had done all I could I sat and waited for the time to come, **as I knew it must**. (Morpurgo Michael, *Kensuke's Kingdom*, Scholastic Inc., 2003 (1999): 91)

(7) **As women's roles have become broader and more encompassing**, their lives have become more powerful. (<https://science-forums.com/index.php?topic=243462.0>)

(8) **As saturation coverage of Russia-related leaks continues**, it seems likely that more and more congressional Republicans will either distance themselves from Trump himself or work with the Trump White House to find scapegoats. (*The Slate*, 15/02/2017, Time for Trump to Hit the Reset Button, Reihan Salam)

- 10 Whereas the first of these examples expresses identity and cannot be said to be causal, the knowledge that something is inevitable certainly influences the character's actions, which may provide a connection with the use of *as* in a proper reason clause. In the other two examples, which exemplify a very common use of *as*, the expression of temporal concomitance is certainly prevalent, but the juxtaposition of two particular events also attracts attention to a possible cause and effect relationship between what is described in the subordinate and what is described in the main clause.⁸ One can even find examples in which temporal simultaneity and cause and effect are closely intertwined, in the same manner as the temporal and the causal values in hybrid SCs. The following quotation by R. Quirk *et al.* describes this phenomenon, which is also illustrated by the attested examples that follow:

It is a testimony of the close and obvious connection between reason and temporal sequence that *as* and *since* are conjunctions of time as well as of cause. This dual function can give rise to ambiguity:

As he was standing near the door, he could hear the conversation in the kitchen.

['Since he was standing near the door...' or 'While he was standing near the door...']
(1985: 1105)

(9) They had to stop in the hallway **as a gurney was pushed out of an autopsy suite**. There was a body on it, wrapped in plastic. (Michael Connelly, *City of Bones*, 2002)]

(10) In the din of one of their several workshops were hundreds of prisoners all sitting before their machines. **As we passed**, they bowed their shaven heads. (*Newsweek* September 23, 1991, 22; quoted by Guimier 1997: 35)

- 11 Examples of ACs in which the simultaneity between two events brings about to some extent a cause and effect relationship, are reminiscent of examples of SCs which M. De Cola-Sekali calls mixed SCs (1992: 144; 154), in that the highlighting of the anteriority of an event also entails a causal relationship. Such examples are more common than proper hybrid examples (see Guillaume 2014: 116-8):

(11) C9M2427 **Ever since the Dutch version of Guitarist appeared on the shelves in Holland**, the English version has been withdrawn. We Dutch get a magazine approximately one quarter the size of the British *Guitarist*, with translations of the English reviews, but through the translation we lose the individuality and humour of the original article.

(12) HH39068 I was extremely interested to read your special issue on Cancer (NI 198). **Since my wife was diagnosed with the illness**, I have been researching alternative cures during which time I have come into contact with the 'Association stop au cancer' [...].

- 12 In examples (9) to (12), however, the temporal value of either marker is clearly predominant, which makes it difficult to replace *as* by *since* or the opposite, as their temporal values are distinct from each other, one expressing anteriority, while the other one describes simultaneousness. As a result, only examples of clear causal values of both markers should be examined in order to try and differentiate between *as* and *since* as causal subordinators.
- 13 I first intended to find such unambiguous examples of causal ACs in the BNC, in order to compare them with the SCs gleaned from that same source. However, in a sample of more than 100 examples of ACs extracted from the BNC, only about one in ten turned out to be causal, and that is even while excluding from the start the numerous *as if* examples. As my corpus of SCs contains more than 300 causal examples (see chart A), and even if I did not necessarily intend to work on a corpus of ACs which would be exactly the same size, it seemed that, in order to find a decent match for the sample of SCs using the BNC resources, I would need to collect and sort out several thousand examples of ACs, that is a particularly time-consuming way of going about it.
- 14 Instead of doing this, I decided to get help, and therefore set my Master's students the task of each finding ten to twenty examples of causal ACs as (a small) part of their assessment, on a voluntary basis. The topic of the Master's course was the use of corpora in linguistics. Eventually, 18 students of various nationalities took part, each collecting 15 examples on average. Still, only a little more than half of the examples (namely 165) could eventually be retained for the purpose of the current research, as many of the examples were not in fact causal (most students, if not all of them, were English majors, and yet many of them found it difficult to distinguish between the different values of *as*). Still, a ratio of about 60% valid examples was still a lot higher than what had been previously obtained with the attempts in the BNC.
- 15 As far as the origin of the examples is concerned, most of them were harvested on the internet, whether in on-line corpora (the BNC, the COCA, but also parallel corpora such

as the Russian National Corpus)⁹ or by having recourse to search engines, or again by browsing Google Books texts or newspaper websites. There were also examples from personal reading, more particularly from books that were on the Master's syllabus regarding other subjects. Finally, a few students looked up research articles on the subject, such as Meier (2002).

- 16 In the end, the corpus of causal ACs, to which I was able to add my BNC examples as well as examples found in personal reading, amounts to about 180 examples, that is a little more than half the number of SCs under consideration. Its advantage, however, is that on the whole it contains more recent examples than the BNC, which was finalised in the early 90s. Although the two corpora are not of comparable size and do not have the same origin, the number of examples in each of them and the varied sources represented (the BNC being itself made up of a miscellaneous collection of works of fiction, letters, press articles, diaries; legal texts, transcriptions of conversations...; see Kennedy 1998: 50-2) made it possible to start working on the comparison between the two markers.

2. The concept of *presupposition* and the placement of causal *as* and *since* clauses

- 17 One point is often mentioned in textbooks and research papers concerning causal *since*, which is not usually made regarding *as*, and it may therefore be a criterion that could help differentiate their causal uses. Because the causal value of *since* stems from the expression of anteriority, the content of a causal SC is often described as presupposed (cf. Rissanen 1999: 305), especially in contrast with the content of a temporal SC. B. Dancygier and E. Sweetser, for instance, characterise causal SCs as “representing ‘factual’ or presupposed information”¹⁰ (2000: 126). In the field of linguistics, presupposition can be understood as the quality of what can be deduced implicitly from the content of a given clause or a given sentence: for instance, *Close the door!* implies that there is a door, that it is open and also that the circumstances make it possible to close it (cf. Dubois *et al.* 1999: 378-9). But presupposition can also correspond to the opposition between given (old) and new information, as defined by M.A.K. Halliday:¹¹

We can now see more clearly what the terms Given¹² and New actually mean. The significant variable is: information that is presented by the speaker as recoverable (Given) or not recoverable (New) to the listener. What is treated as recoverable may be so because it has been mentioned before; but that is not the only possibility. It may be something that is in the situation, like *I* and *you*; or in the air, so to speak; or something that is not around at all but that the speaker wants to present as Given for rhetorical purposes. The meaning is: this is not news. Likewise, what is treated as non-recoverable may be something that has not been mentioned; but it may be something unexpected, whether previously mentioned or not. The meaning is: attend to this; this is news. (1985: 277)

- 18 As a result, causal *since* is apt to describe things which are obvious in the situation, but that the speaker needs to recall explicitly, as they entail something else; such a context may be ironic, as in the second example:

(13) AR3567 ‘It has served me perfectly well thus far as it is, Miss Kenton, though I appreciate your thoughts. In fact, **since you are here**, there was a certain matter I wished to raise with you.’ ‘Oh, really, Mr Stevens.’

(14) (*Jane Doe wanted to join the FBI team in interrogating a suspect, but the FBI agent in charge, Kurt Weller, refuses point blank. Nevertheless, when he realises that the suspect only*

speaks Chinese, he decides to fetch Jane from the car in which she has been told to wait, as she is the only one who speaks this language among the people present.)

Jane: Forget something?

Weller: Could you please come upstairs with me for a minute, M'am?

Jane (ironically): **Since you ask so nicely...** (*Blind Spot*, season 1, episode 1)

- 19 Because *as* is intimately linked with the expression of identity, and, as a result, temporal simultaneity, the cause and effect relationship which it brings about is more difficult to characterise as presupposed. One might even suppose that causal ACs are more prone to imparting 'new' information in the sense of Halliday than 'old'. The context of example (15) is particularly clear regarding the fact that the information given in the subordinate is new (*now*), as it is contrasted with a previous one:

(15) ANM 1148 Pram Seat; Could whoever it was who wanted one please contact me **as I now know of one for sale.**

- 20 Still, an AC can also very well impart old information, as in the following examples:

(16) It is often also argued that the Queen is an ideal figure to represent Britain **as she is a neutral figure above the political arena and one who can represent the nation as a whole without carrying any political baggage.** (*Interstate - Journal of international affairs* vol. 1998/1999 No. 1: 1, James T. Williams, "The Future of The British Monarchy")

(17) The young pigs were piebald, and **as Napoleon was the only boar on the farm**, it was possible to guess at their parentage. (George Orwell, *Animal Farm*, 1945, chapter IX)

- 21 In (16), the essayist upholds constitutional monarchy in a manner which is not particularly original, while the fact that Napoleon is the only boar in Orwell's famous *Animal Farm* is not news to the reader at this stage of the novel. Besides, the old vs. new information distinction may be a relevant theoretical criterion, but, in practice, it often turns out to be difficult to decide whether an AC imparts old or new information, especially if there is only little context:

(18) **As Harry Potter action figures are relatively new to the market**, there is not an overabundance of information available. This book was written to help fill that niche. (Silvester William, *Harry Potter Collector's Handbook*, F+W Media, Inc, 2010, 17)

- 22 If this article is read in its context (which is available online on Google Books), it may be considered as given information, because the author has just explained that such action figures only became available in 2000. At the same time, the idea that the book is written to 'help fill a niche' strongly suggests that few people are aware of this fact.

- 23 The supposedly presupposed quality of the content of causal SCs, as opposed to causal ACs, falls short of accounting for the subtle differences between the two markers, simply because taking into consideration a large number of examples proves that causal ACs may contain old just as well as new information, and that, in certain cases, it may simply be hard to decide which it is. As for causal SCs, it must be understood that their content is above all *presented* as old, actual information, likely to be already known to the addressee, which more often than not can prove strategic, as in the following examples:

(19) BNB460 Walpole was not deceived when, in 1738, Jacobite MPs argued in the House of Commons that **since their cause was clearly dead** the regular army could be reduced.

(20) CEK6760 There can be few places where you can indulge yourself so cheaply as the chateaux of France's Western Loire. Whether you choose a chateau hotel or stay in stately homes where families take in guests, splendour is the word and we found

it at the majestic Chateau de Noirieux in Briollay, Anjou. It was hardly necessary to open my bulging suitcase **since almost everything was provided, including bathrobe, slippers, hairdryer and toiletries.** (*Today* 1992)

- 24 In the above examples, the content of the causal SC is new rather than old information from the addressee's point of view; in (19), the reported speakers intend to deceive their addressee by presenting as actual fact something which is false. As for (20), the recourse to *since* is simply a journalistic device, which enables the writer to enumerate all the items provided at the chateau without being too heavy-handed about it, while the reader cannot possibly be expected to be already aware of all these details. This goes to show that characterising the content of an SC only as given information is missing the point of why *since* was selected by the speaker.
- 25 My hypothesis regarding the causal use of *since* is that this subordinator has a *backgrounding effect* on whatever content it introduces (*cf.* Guillaume 2014). In (20) for instance, the journalist does as if what he or she mentions is so obvious that the reader only needs to be gently reminded of it, instead of informed of it. The same strategy also accounts for the following example:
- (21) A8F640 The film, set on the beaches, in the bars and discos and in the shabby tourist rooms of downtown Acapulco, is distinguished by a central performance from Jackie Burroughs (also one of the directors) that is an astonishing tour de force of which it is impossible to speak too highly. That it is sustained almost to the point of embarrassment is exactly what is so good about it, **since discomfort is the name of this film's particular game.** (*The Guardian*, arts section, 1989)
- 26 In (21), the SC clearly contains the film critic's personal opinion, which can hardly be known to the reader prior to reading this column. Therefore, a subjective judgement is presented in the critique as a presupposed, unquestionable statement, by means of the recourse to *since*. Indeed, causal *since* primarily indicates that the clausal content is non negotiable,¹³ because it is presented as essentially non polemical (*cf.* De Cola-Sekali 1991: 69; Deléchelle 1993: 188-9). M. De Cola-Sekali goes as far as calling *since* a "linguistic immunity, an exemption".¹⁴ As for the backgrounding effect which is at play, it does not only apply to the content of the causal SC, but to the cause and effect relationship as a whole. The link between a cause in the SC and a consequence in the main clause is presented by the speaker as taken for granted, and, as such, it is not open for debate, a feature which is not found as such when resorting to ACs.
- 27 Syntactic criteria can also be taken into account when comparing the causal values of *as* and *since*. When comparing the placement of the SCs and the ACs with respect to the main clause in my corpus of examples, there is a striking similarity between the two conjunctions in that the postponing of the reason clause accounts for about 60% of the examples in each of the categories. Both ACs and SCs can appear before or after the main clause to which they are related,¹⁵ but with a bias towards the postponed position.¹⁶ The question is whether fronting or postponing the reason clause has the same effect in the two cases.
- 28 The concept of presupposition can again be resorted to in order to account for the placement of a subordinate, as it is commonly admitted that given information tends to appear in sentence initial position, where it can serve a cohesive function with what precedes and provide a frame for the rest of the sentence (see for instance Biber *et al.* 1999: 835ff). But in the case of ACs and SCs, it seems that a clear majority of such clauses provide what can be termed given information (this being even more true for *since* than for *as*), while in my corpora both ACs and SCs seem to favour final position.

- 29 Regarding the placement of causal SCs, I have tackled the question at length in Guillaume (2013). I came to the conclusion that, except for a few cases in which syntactic constraints, such as the presence of certain adverbs before *since* for instance, play a part, the difference between a fronted SC and a postponed SC is quite subtle, and above all a matter of emphasis. Whenever the SC is fronted, it is not the SC which is thrown into relief, but rather the cause and effect relationship between the two clauses. On the other hand, whenever the SC is postponed, it is the content of the main clause which is made prominent by the speaker.
- 30 On the other hand, subordinator *as*, contrary to *since*, does not have a backgrounding effect on the content of the subordinate clause which follows it. Whereas *since* presents the content of the subordinate as having been previously validated (thanks to its link with anteriority), *as* indicates simultaneity, and, as such, whenever an AC is fronted, the cause of what will follow is made prominent, while it is the consequence described by the main clause which attracts more attention whenever the AC is postponed.
- 31 Still, the presence of *so* at the beginning of the main clause in the next example, in which the AC is fronted, suggests that, in the same manner as with SCs, the fronting of an AC tends to highlight the cause and effect relationship even more than it does the content of the AC. Indeed, the phrases *as... so... or as... therefore... can only be resorted to if the clauses appear in that order, which is also the case of since... then...¹⁷*

(22) CCL164 Our Lord came to inaugurate the Kingdom. As he began his ministry it was present in his person. [...] His amazing words and miraculous deeds were signs of the coming Kingdom. **As his disciples were identified with him in his ministry and experienced the healing powers of the new age operating through them, so** they became partakers of the Kingdom.

(23) Accepting the view that man was prior in the creation, some Scriptural writers say that **as the woman was of the man, therefore**, her position should be one of subjection. Grant it, then **as the historical fact is reversed in our day, and the man is now of the woman**, shall his place be one of subjection? (Cady Stanton Elizabeth, *The Woman's Bible*, 1898, 20)

(24) ECH475 Speed on the climb will certainly avert problems later in the day. For an ascent of the Whymper Couloir of the Aiguille Verte for instance, a midnight start and a six hour target time are prerequisites to success. This means climbing the whole route in the dark, and **since an efficient descent on frozen snow should only take four hours, then** you should be back at the Couvercle Hut sipping beer by 11am.

- 32 Such phrases demonstrate once more that the causal value of both *as* and *since* undoubtedly stems from their temporal value, as *then* is also used to signal a succession of events, whereas *so* is a marker of identity, in the same way as *as*, and therefore highlights the concomitance between the actions described in the subordinate clause and the main clause respectively.

3. A theoretical approach to the differences between *as* and *since*

- 33 Trying to distinguish between the causal use of *since* and the causal use of *as* in terms of how much presupposed the content of the clause seems to be, does not yield very satisfactory results. How then can one account for the specificity of each of these markers, when they prove so close to each other in usage?

- 34 Regarding causal relationships as a whole, A. Leroux (2009; 2012) explains that, more often than not, there is no actual need to have recourse to an explicit subordinator, as the mere juxtaposition of the events described is self-explanatory. Among others, this explains why, in the case of *as*, temporal concomitance can pave the way for a causal relationship (see examples 9 and 10). But what is even more interesting in this point of view is that one may suppose that, in the cases when an explicit conjunction such as *since* or *as* is used, it may be so in order to express something more than a simple causal relationship.
- 35 Following in the footsteps of H. Adamczewski, G. Deléchelle explains that connectors such as *since* or *as* work within a complex sentence in the same way, to some extent, as modals do in a simple sentence, in that they express above all the speaker's point of view on the relationship between the main clause and the subordinate clause (1993: 189). However, as was suggested by A. Leroux (2009: 97), I shall look at it slightly differently, and regard such conjunctions as indicators of the way in which the speaker anticipates the addressee's reaction, among others their reaction to the connection established by the speaker between the subordinate clause and the main clause.
- 36 For that purpose, I shall now resort to some of the linguistic tools developed by A. Culioli in his utterer-centred approach, in order to clarify the position of the speaker and that of the addressee, with regard to each other, and also in respect of the causal relationship between the main clause and the subordinate clause evidenced by the recourse to *as* or to *since*.
- 37 The concept of location is a fundamental operation in A. Culioli's *Theory of Enunciative and Predicative Operations*¹⁸ (TOE for short): "a located term is one which has been situated, specified, or determined. [...] The basic idea is that a linguistic object only acquires a determined value by means of a system of location." (Bouscaren *et al.* 1992: 150). The operation itself is symbolised by epsilon: $\underline{\epsilon}$, hence $\langle x \underline{\epsilon} y \rangle$ means that x is located relative to y . Location may correspond to several different values:¹⁹
- **identification** $\langle x = y \rangle$: in this case, the located term is identifiable with the locator, for instance *At that time, the King of France was Louis XIV.*
 - **differentiation** $\langle x \neq y \rangle$: here, x cannot be identified with y , the located term is different from the locator, for example *Meg has a doll.*
 - **disconnection** $\langle x \omega y \rangle$: in this special case, the located term cannot be located in respect of the locator, and therefore it is neither identified with it, nor different from it. Each of the items is on a different plane, which is why they are 'disconnected' from each other. Modal uses of the English preterite are examples of hypothetical disconnection, in that they describe states of affairs which are disconnected from the actual situation in which the speaker is located, for instance *if pigs had wings, they would fly.*
 - **hypothetic plane** $\langle x * y \rangle$: this complex value, very often linked to hypothesis and the use of *if*, combines all the other values, in that it may correspond to neither equal nor different, or again either equal or different, or even both equal and different. Complex markers such as the pronoun *one* may be characterised thanks to the hypothetic plane, as it may designate no one in particular (*One drinks a lot of tea in England*) or someone specific, but in an indirect manner, who may be different from the speaker and possibly identified with the addressee (*I can tell that one has had a lot of fun*), or again identified with the speaker (*One needs a little comfort after all*).
- 38 If one tries to apply such concepts to the use of *as* or *since* as causal subordinators, it seems relevant to consider to what extent the speaker (symbolised by S_0 in the TOE)²⁰

endorses the content of the causal clause and, as a result, the cause and effect relationship which brings about the consequence described in the main clause. In both cases, it seems fairly obvious that the speaker who resorts to causal *as* or causal *since* takes the content of the clause, as well as the cause and effect relationship that follows from it, for granted.

- 39 A subtle difference can however be posited between the use of the two conjunctions in the way in which the speaker construes the position of the addressee (symbolised by S'_0 in the TOE) in respect of these same phenomena (content of the subordinate clause and validity of the cause and effect relationship between subordinate clause and main clause). My hypothesis is that the position of the speaker can be said to be equated with that of the addressee ($S_0 = S'_0$) in the case of *since*; the content of the subordinate clause as well as the cause and effect relationship are *presented* as unproblematic, the addressee is therefore *expected to agree*, hence $S_0 = S'_0$. Of course, this does not mean in actual fact that everyone will agree, but the speaker does not leave room for disagreement.
- 40 With causal *as* on the other hand, the addressee is not taken into account when establishing a cause and effect relationship between an AC and a main clause. Whereas *since* takes the addressee's opinion into consideration, but only insofar as they will agree, the recourse to *as* is disconnected from any representation of the addressee. The content of the subordinate clause, as well as the cause and effect relationship, are simply *felt* by the speaker as unproblematic, and, at best, the addressee is *not expected to disagree*, hence $S_0 \omega S'_0$. Their reaction is neither anticipated, nor even relevant. The speaker and the addressee are no longer on the same plane.
- 41 This hypothesis is in contradiction with G. Deléché's analysis, who, in his thesis (1989), argues that *as* is congruent whereas *since* is polemical. This is an interesting point, even if it gives too conspicuous a role to the content of a causal SC, whereas causal *since* has rather a backgrounding effect in my opinion. G. Deléché also says that the use of *since* materialises the presence of the addressee, and more specifically of someone who is expected to disagree with the speaker.²¹ Whereas I strongly agree that *since* takes the addressee into account, I do not believe that the latter is construed as uncooperative. On the contrary, whenever *since* is resorted to, what is presupposed by the speaker is not only the content of the SC and the cause and effect relationship which it entertains with the main clause, it is also the fact that the addressee supposedly agrees with these unconditionally. In the same manner, causal *as* cannot be used in a polemical context, which is in keeping with the way in which it is used to express identity.
- 42 The difference between the respective positions of speaker and addressee may also explain why the content of the SCs may be felt as more presupposed than that of the ACs, even if it was argued that the degree of presupposition cannot account by itself for the differences between causal *as* and causal *since*. In a theoretical approach different from the TOE, O. Ducrot's definition of presupposition also stresses the existence of different possible stances of a given speaker towards the content of his speech, in which SCs could correspond to *we*, while ACs would be more comparable to the use of *I*:

En introduisant une idée sous forme de présupposé, je fais comme si mon interlocuteur et moi-même nous ne pouvions faire autrement que de l'accepter. Si le posé est ce que j'affirme en tant que locuteur, si le sous-entendu est ce que je laisse conclure à mon auditeur, le présupposé est ce que je présente comme

commun aux deux personnages du dialogue, comme l'objet d'une complicité fondamentale qui lie entre eux les participants à l'acte de communication. Par référence au système des pronoms, on pourrait dire que le présupposé est présenté comme appartenant au "nous", alors que le posé est revendiqué par le "je" et que le sous-entendu est laissé au "tu".²² (1984 : 20)

- 43 Even though the difference between the two conjunctions is very tenuous indeed, traces of it may be found in the examples of the corpus, for instance in fiction, in the way in which the narrator is located in respect of the content of a novel. Here are examples from fictional first-person narratives; but whereas one is a classic thriller in which the hero narrator unfolds his tale without pretending that it is a diary, the other one more clearly takes the form of a diary:

(25) Grand larceny was nothing to be laughed at, **especially since I was guilty**. The maximum was ten years in prison. I would worry about that later. (John Grisham, *The Street Lawyer*, Random House 1998: 313)

(26) **Since the firm had been my life for the past seven years**, it wasn't difficult to predict its actions. There was too much pride and arrogance to suffer the indignities being exposed. (John Grisham, *The Street Lawyer*, Random House 1998: 385)

(27) New Year's Resolutions

I WILL NOT [...]

Waste money on: pasta-makers, ice-cream machines or other culinary devices which will never use; books by unreadable literary authors to put impressively on shelves; exotic underwear, **since pointless as have no boyfriend**.

[...] Obsess about Daniel Cleaver **as pathetic to have a crush on boss** in manner of Miss Money Penny or similar. (Helen Fielding, *Bridget Jones's Diary*, 1996)

(28) 9st 3 (but post-Christmas), alcohol units 14 (but effectively covers 2 days **as 4 hours of party was on New Year's Day**), cigarettes 22, calories 5424. (Helen Fielding, *Bridget Jones's Diary*, 1996)

(29) 2 Bloody Marys (count as food **as contain Worcester sauce and tomatoes**) (Helen Fielding, *Bridget Jones's Diary*, 1996)

- 44 In the two Grisham examples, the narrator merely recaps facts that are already known to the reader, whereas Bridget Jones supposedly writes her diary for her benefit only, and is therefore not expected to take a possible reader into consideration. The latter case is a possible illustration of the disconnection between speaker and addressee, which has been described regarding *as*. In (27), the narrator resorts to causal *since* followed by causal *as*, but throughout the book the examples of ACs are quite common. In reverse, no example of ACs were found in the Grisham novel, but of course one may argue that it is related to the style of the author, as it is nowadays well-known that the authors' use of words and phrases is a matter of subjectivity, and differs from one person to the next.
- 45 Although the heterogeneity of my corpus regarding *as* and the fact that I was not the only one collecting examples makes it almost impossible to provide reliable statistics, it must be said that, during the gathering of the examples, both myself and some of my students have felt that it proved on the whole easier to find examples of causal ACs in newspaper articles rather than in literary fiction, a genre in which examples of causal SCs are on the other hand very easily come by. Although I wish to remain very cautious regarding such statements, this view could be supported to some extent by the different stances of the speaker regarding the addressee which I have tried to characterise in respect of causal ACs and causal SCs. Whereas a fictional narrator, whether they be omniscient or identified with a character in particular, often tends to

recap and capitalise upon events that have already been narrated to the reader, a journalist has to present connections between facts that may be already known to certain readers, but not to others, and therefore must ignore this factor by simply not taking the addressee's possible previous knowledge into account:

(30) The second point was important **as the backdrop to the end of empire was, as it is now, a revanchist Russia** (then the Soviet Union), threatening eastern Europe and thus Britain's strategic interest of maintaining the balance of power on the continent.

(*The Economist*, 9/02/2017, The Art of Leaving. What the break-up of the British Empire can tell us about Brexit)

(31) Much has been written about "post-truth" politics in the context of the recent presidential election, and rightly so, **as Trump's relationship with the truth is not the same old conservative legerdemain**. (*New York Times*, politics sections, 2017)

- 46 In examples (20) and (21), the recourse to *since* was justified by the reporter's characterisation of the information contained in the SC as undebatable, even though it could hardly already have been known to the addressee. Here the stance is somewhat different, in that an educated reader may already be aware of such views, but a less well-informed reader can read about these assumptions for the first time; in either case, the journalist strongly asserts their opinion and does not expect the reader to disagree with it.
- 47 The idea that the distribution of causal clauses in general could be genre-related has already been put forward by linguists working on former states of the English language. For instance, Rissanen (1998) studies the distribution of causal conjunctions in Early Modern English and brings out striking differences in their frequency according to genres, especially as far as *for* and *because* are concerned. Claridge and Walker (2001: 38-9) find that *as* and *since* are comparatively more rare than other causal conjunctions in Early Modern English, as they are today, and that they are almost restricted to formal written English such as sermons, with *since* being somewhat more common than *as*.

4. Searching examples in a parallel corpus and exploring the issue of genres

- 48 If the stance of the speaker and their expectations towards the addressee really are different whenever they choose *as* over *since* (or the opposite) to express causality, certain genres might indeed be more typical of one subordinator or the other: whereas newspaper articles and diaries could favour *as*, the narrator in classic fiction could be partial to *since*. Although I have noticed such tendencies during the gathering of examples, I cannot vindicate them, as my corpus is both too small and much too heterogeneous to allow me to prove such claims.
- 49 It seemed however possible to undertake a small-scale experiment which could either invalidate these impressions altogether, or support them to some extent. For that purpose, I needed to be able to compare samples of causal ACs and SCs gleaned from the same sources. Given the difficulty in finding relevant causal examples of *as* (see part one), the only way to gather examples of the two sorts efficiently was to resort to a parallel corpus, in which the possible French translations of causal *as* and *since* could help discriminate between the varied uses of *as* (and even of *since*, for that matter), and

thus enable me to obtain a sample whose origin could be easily looked up in order to test the hypothesis connected to genres.

- 50 The Czech linguist O. Nadvornikova describes the making up of the parallel corpus Intercorp and discusses the pros and cons of working on such a multi-lingual corpus (2017). Intercorp is easily accessible on the internet (<https://www.korpus.cz/>) after a free registration. The KonText interface allows simple as well as more sophisticated searches in one or more of the languages available in Intercorp version 10. To this day, the French version contains 93 644 000 words and the English version 123 172 000 words.
- 51 Regarding the possible translations of *since* into French, “*puisque*” appears to be the most obvious, followed by “*étant donné que*”, a phrase in which, just in the same way as *since*, causation is linked to the temporal anteriority of an event. Other translations, such as “*parce que*”, “*vu que*”, for example, may be encountered, but they are less typical of *since*.²³ As for *as*, both “*comme*” and “*puisque*” are common translations, with also an obvious etymological likeness between *as* and “*comme*”, which in French can also express simultaneity as well as comparison.
- 52 To keep things simple and in order to work on a small, manageable sample, I have chosen to use “*puisque*” as a touchstone in the KonText searches in the Intercorp French and English versions, as it stands good chances of being translated by (or being the translation of) either causal *as* or causal *since*. Indeed, it appears to be fairly neutral between the two, whereas, for instance, “*étant donné que*” may turn out to be biased towards *since* and “*comme*” towards *as*.
- 53 Another important aspect was to decide which sub-corpora to include in the experiment. As O. Nadvornikova explains (2017: 4), the sub-corpora found in the various versions of Intercorp are very different from one another in both their nature and their size. The so-called Core part of the corpus consists in literary or scientific texts and their commercial translations (the so-called scientific texts being in fact present to a lesser degree, as we shall see), but only accounts for 17% of Intercorp as a whole.²⁴
- 54 The largest part of Intercorp comes from a website which provides free subtitles in various languages for films and series, namely www.opensubtitles.org. Another significant part (in terms of size) are translations of European Commission documents and transcriptions of debates in the European Parliament (2007 to 2011). A much smaller proportion of the corpus is made up of translations of the Bible. Finally, journalistic texts, labelled Syndicate or EuroPress, are present, but only to a small degree (only 3% of Intercorp as a whole).
- 55 For my purpose, which is mostly to compare the frequency of causal *as* and causal *since* in various genres, it seemed fairly obvious that both the Core corpus and the press corpus had to be included. The Bible, on the other hand, is too specific a text and contains archaisms, not to mention the heterogeneity of its content in terms of authors, genres, and when the various texts were written. As for the European institutions’ documents or transcriptions, they also are hard to classify according to genre, because they contain documents of very varied nature, from oral transcriptions of debates to legal texts. Also, the sheer size of this sub-corpus made it difficult to include it in the experiment, as it was bound to return an overwhelming number of examples. Finally, the subtitles also constitute a specific genre which is neither literature nor press; above all, a quick look at certain examples shows that the quality

of the translations is extremely poor – these freely downloadable subtitles are not devised by professional translators, but mostly by fans, who resort to automatic translators, with barely any proofreading or none at all; as a result, the content of this sub-corpus is not at all reliable for linguistic research.

56 The search on *puisque/as* in the Core and press subcorpora of the French and English versions of Intercorp 10 yields 395 hits, while 381 results are found for the *puisque/since* request. Such numbers are comparable and can be dealt with by just one person. I have therefore attempted to assess from what subcorpora the examples came in each case.

57 Nevertheless, as is always the case with such researches, some of the results turned out to be irrelevant. By default, KonText provides a lot of context for each of the examples, which is very helpful for linguistic analysis, but means that, for instance, the presence of *as* in the English text in the near context of a phrase translating “*puisque*” will be included, even though it is not the translation of “*puisque*”; given the frequency of *as* in English and the diversity of its uses (see part one), this paves the way for a large number of errors in the *puisque/as* search, as in the following examples:

(32) **Puisque** les décideurs politiques opèrent dans des environnements de deuxième ordre, le cheminement optimal des réformes (même dans des cas apparemment simples tels que les réformes des prix) ne peut être conçu sans considération nécessaire des conditions courantes et sans mesurer les conséquences des multiples marges à échelle différente.

→ **Because** policymakers operate in second-best environments, optimal reform trajectories—even in apparently straightforward cases such **as** price reform cannot be designed without due regard to prevailing conditions and without weighing the consequences for multiple distorted margins. (Syndicate)

(33) [...] alors j’ai demandé que les mitrailleurs me donnent aussi une arme, **puisque** je savais me servir d’un fusil [...]

→ (...) and I wanted the sub-machine-gunners to let me have a Kalashnikov **as well**, **because** I could handle a gun [...] (Core)

58 I have also termed non-relevant the few examples in which “*puisque*” was translated by (or was the translation of) phrases such as *seeing as* and *insofar as*, which, although they are related to the causal value of *as*, are more complex and therefore different from the use of causal *as* by itself.

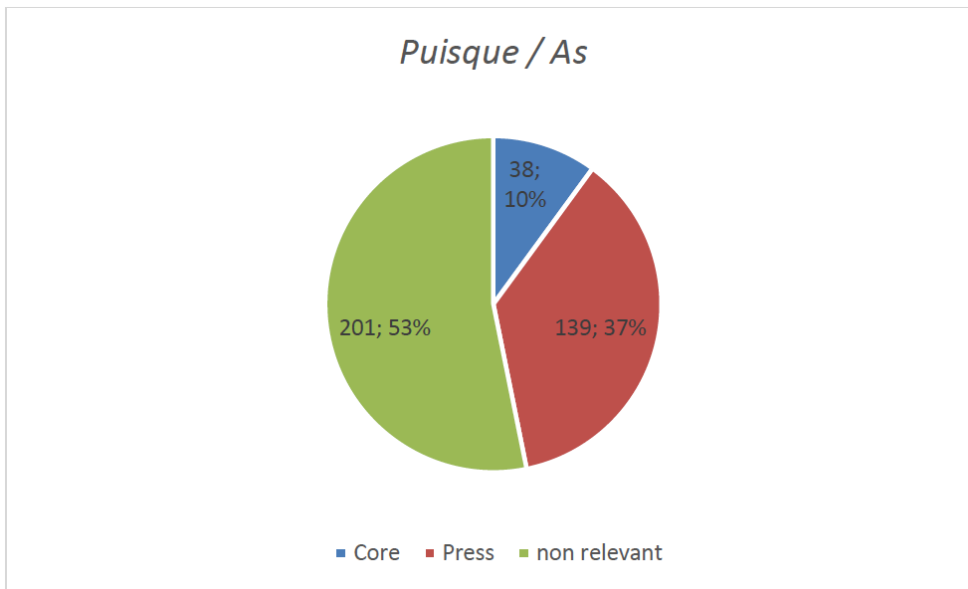
59 As far as *puisque/since* is concerned, the presence of temporal *since* in the vicinity of a phrase translating “*puisque*” could also yield non-relevant examples, but the confusion was much less frequent than with *puisque/as*, as we shall see:

(34) L’objectif plus ambitieux du Sommet mondial de l’alimentation, de réduire de moitié le nombre total de personnes souffrant de la faim dans le monde, est encore plus lointain, **puisque** le nombre de personnes sous-alimentées n’a baissé que de 17 pour cent depuis 1990-1992.

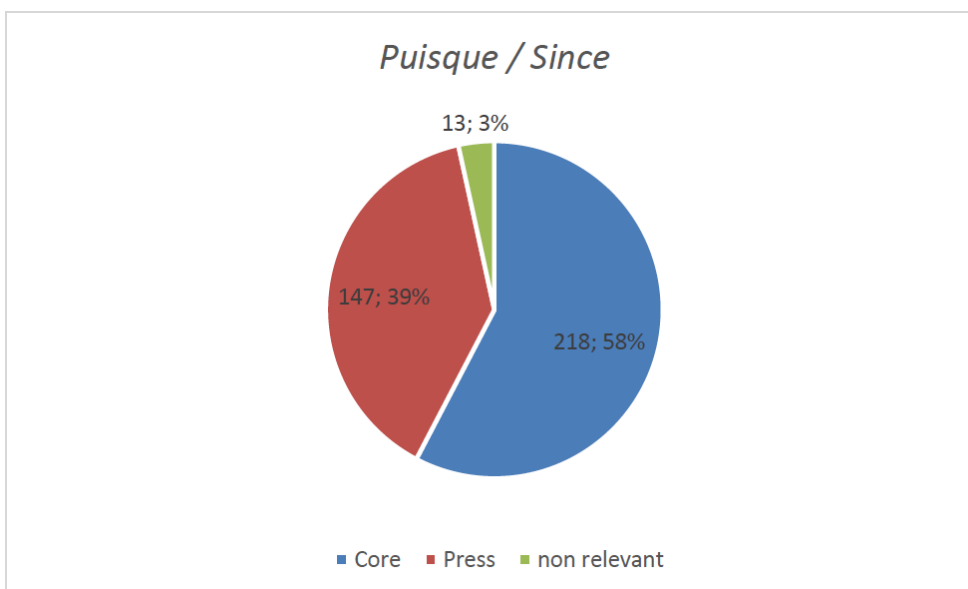
→ The more ambitious WFS target of halving the total number of hungry people worldwide is even more distant, **with** the number of undernourished people having fallen by only 17 % **since** 1990-1992. (Syndicate)

60 In the end, I was able to devise statistics on 378 examples for each search (including non-relevant examples), instead of the 395 and 381 examples which had been announced respectively.²⁵ The results of this small-scale experiment turn out to show undeniable differences as far as the distribution of *since* and *as* as a translation for “*puisque*” are concerned:

B. Distribution according to the Intercorp sub-corpora of the examples found when looking for translations of “*puisque*” using *as*



B. Distribution according to the Intercorp sub-corpora of the examples found when looking for translations of “*puisque*” using *since*



- 61 Even leaving aside the very high frequency of non-relevant hits (more than half of the examples returned), causal *as* has more than three times as many chances of appearing in the press corpus than in the Core corpus. This is all the more remarkable because, in both the English and the French versions of Intercorp 10, the Core corpus is much bigger than the sum of the Syndicate and PressEurop examples: roughly speaking, the French Core corpus is twice as big as the press corpus, while the English Core corpus is three times as big as its journalistic counterpart.²⁶
- 62 Regarding *puisque/since*, the risk of being confronted to non-relevant examples is very low, contrary to *puisque/as*. *Since* can be encountered in both literary and scientific texts, as well as in press examples. The relative superiority of the representation of the

former group (58% *as* as opposed to 39%) must be corrected by the fact that, as stated above, the Core corpus is much bigger than the press sub corpus. One may therefore deduce from these findings that *since* is very common in both subcorpora; unlike *as*, it does not seem to be over represented in either of the two groups.

- 63 But how far are these differences between the findings connected to *as* and those connected to *since* really significant? The fact that the number of non-relevant examples is much higher with *as* than with *since* is not unexpected, as the difficulties of finding causal (and only causal) examples of *as* were already described in the first part; a search in the BNC, even when eliminating *as though* and *as if*, only returned 10% causal hits, while assigning advanced students the task of finding examples still made it necessary to do away with 40% of the examples. Therefore, if the goal of this experiment had been simply to gather examples of the causal value of *as*, the method used, with roughly 47% chances of finding relevant examples, could have been said to be relatively effective. Nevertheless, it certainly tends to interfere with the comparison with *since*, which only has a very low rate of irrelevant examples.
- 64 Above all, even in a corpus whose parts are clearly identified as belonging to a specific genre, the putting together of heterogeneous texts under the same heading makes it difficult to claim that one subordinator is more typical of a genre than of another. In Intercorp, it is mostly the Core corpus whose unity seems problematic. Most of the examples gathered during the *puisque/since* and *puisque/as* searches are taken from only a few books (by today's standards, the subcorpora in Intercorp are rather small for each given language). Some of them undeniably belong to the literary genre; one finds, among others, texts by A. de Saint-Exupéry (*Le Petit Prince*), L.-F. Céline (*Voyage au bout de la nuit*), A. Camus (*La Peste*), J. Verne (*Le Tour du monde en 80 jours*), A. Nothomb (*Stupeur et tremblements*), J.R.R. Tolkien (the *Lord of the Rings* series) J. K. Rowling (the *Harry Potter* series), G. Orwell (1984) as well as French and English translations of novels and plays by various Czech authors, such as M. Kundera, V. Havel and F. Kafka, to quote but the best-known. All these texts are extremely different in terms of style, and they were written and published over a very long period of time. To make things even more difficult, the Core corpus also comprises so-called scientific texts, which, nevertheless, turn out to be mostly scientific essays, hence perhaps the fact that they were put in the same category as the literary texts. Most 'scientific' examples in my Intercorp searches came from S. Hawking (*A Brief History of Time*), W. Heisenberg (*Physique et Philosophie*, in French), M. Foucault (*Les Mots et les choses*).
- 65 Therefore, apart from J. K. Rowling, J.R.R. Tolkien, G. Orwell, S. Hawking and a few others, most of the uses of *as* and *since* in the Core corpus are translations from French, German, Czech, and possibly other languages. Still, these translations are made by professionals, and were sanctioned by a publishing house, which means that they are supposed to abide by certain standards in terms of quality, unlike the subtitles part of the corpus, which was eliminated in my searches because it clearly lacked reliability.
- 66 As far as the Syndicate and PressEurop subcorpora are concerned, it is possible to find out when and where the article was published, as well as its original title. As a result, one may know which was the language of the original text and which is the translation. Most of the articles used were originally published on a website, namely Project Syndicate (<https://www.project-syndicate.org>), which aims at quality journalism and provides free translations of their articles in numerous languages.²⁷ As for VoxEurop (PressEurop), it has a team of volunteer contributors and translators, and also takes up

articles from well-known European and international newspapers and magazines. It must also be noted that, unlike a significant proportion of the texts in the Core corpus, all the media articles are fairly recent –at the very least, they were published in the 21st century, the earliest ones dating back to 2003, but most of them being more recent than this.

- 67 Therefore, the Core corpus is very heterogeneous, while the journalistic examples seem to make up a more coherent sub-group, but are taken from two very specific websites and may therefore lack the diversity of what is usually labelled the press, which can normally be local, national or international, or encompass broadsheet newspapers or tabloids... In reverse, Project Syndicate tend to normalise the contributions which they publish and are aimed only at an international audience. VoxEurop offers more variety in this respect, as it quotes an amazing list of sources from all types of press (<http://www.voxeurop.eu/en/sources>), but it is even less present in Intercorp as a whole than Project Syndicate. In either subcorpus, lots of the English examples are professional translations from other languages, but it is probably even more true in the Core corpus, as many articles from the two websites (more particularly Project Syndicate) were written in English originally.
- 68 And yet, in spite of all these shortcomings, of which one needs to be aware, the result yielded by the experiment on *puisque/as* and *puisque/since* do point in the direction of my first hypothesis: it seems to be the case that the distribution of causal *as* in journalistic articles is significantly higher than that of causal *since*, which stands more chances of being encountered in literary texts or scientific essays. The simple fact that so many examples of journalistic uses are evidenced for both conjunctions, and more particularly for *as*, attracts attention in itself, because in both the English and the French Intercorp 10 subcorpora, the Core corpus is much bigger than the sum of the Syndicate and PressEurop examples. Therefore, leaving aside all the reservations previously made, the presence of so many examples of causal *as* which can be labelled as journalistic is, after all, very significant, even if it is not decisive.
- 69 What the experiment helps most set into relief is the lesser frequency of *as* in literary and scientific texts, however limited and heterogeneous the Core corpus may be, and its high frequency in recent journalistic texts, although these were much less represented. As for *since*, it is frequent in both subcorpora, even though it is comparatively more frequent in literary and scientific texts (but this can simply be due to the size of the Core corpus as compared to the press corpus).
- 70 Such findings can be related to my hypothesis that the speaker (or rather the writer, in this case) does not take into account whether the addressee is already aware of the content of the AC, and whether they will sanction the cause and effect relationship posited by the speaker. The detached, objective stance of journalist writing, whose main goal is to impart information, even if it is never completely devoid of judgement, is compatible with such an analysis of causal *as*. On the other hand, in a work of fiction such as a novel, or even, for that matter, in a scientific essay, the author tends to rely in the course of their writing on events that have already been described, facts that have already been mentioned or proved, and, in such a case, the backgrounding effect of causal *since*, which anticipates nothing but agreement on the part of the addressee, is particularly well suited to such contexts.

Conclusion

- 71 I have attempted in this paper to argue that the opposition between *as* and *since* whenever they are used as causal subordinators cannot be explained only by taking into account the degree of presupposition of the content of the subordinate, that is whether it is given or new information. Not only does such a distinction prove too subjective to be reliable in most cases, but the analysis must also take into account the expectation of the speaker towards the addressee's reaction. While the causal relationship materialised by an SC is *presented* as unproblematic and is addressee-oriented (in that they are expected to endorse it), the recourse to a causal AC corresponds to a causal relationship which is *felt* as unproblematic, and is speaker-oriented insofar as the addressee's reaction is neither anticipated nor relevant.
- 72 Such theoretical considerations can help explain why it is very often difficult to account for the differences between these two causal markers, which easily commute with each other in a vast majority of examples. Indeed, the two different strategies boil down to a very similar effect in the end: a *since* causal relationship is presented as undebatable and non-polemical because the speaker expects the addressee to agree with them unconditionally, while *as* as a causal subordinator does not leave any more room for debate, because what the addressee might think is not deemed relevant.
- 73 Nevertheless, the constitution of a small corpus of several hundred examples in order to study the causal values of conjunctions *as* and *since* first suggested to me and to certain students, involved to a small degree in this project, that it was easier to find examples of causal *since* than of causal *as* in literature and essays, while causal *as* was perhaps more frequent than, or at least as frequent as causal *since* in newspaper articles (in the broad sense, that is including their internet versions). As this impression did not contradict, but rather supported, the hypothesis that the speaker's representation of the addressee is not the same when they resort to causal *as* or *since*, I decided to take a look at an online corpus which was both parallel (hence making it easier to find relevant examples of causal *as*), and also comprised subcorpora with rather distinctive features.
- 74 Despite shortcomings evidenced in the making-up of the subcorpora and the fact that the search for actual examples of causal ACs once more proved much more arduous than that of causal SCs, thus disrupting to some extent the possible comparison between the two, it is possible to evidence a tendency for ACs to be more present in the press, while SCs are very common in literature, essays and the press as well.
- 75 Starting with hardly more than an impression, the recourse to Intercorp has enabled me to substantiate to some extent my hypotheses regarding both the frequency and the significance of ACs and SCs. But this is clearly not telling enough, and only a large-scale experiment in a corpus such as the COCA could be significant. Just like Intercorp, the COCA is divided into sub-sections which represent different genres (see for instance Shahrokny-Prehn and Höche 2011). But it is much larger than Intercorp, especially as far as English is concerned, and also more homogeneous, as it is mostly made up of recent texts and transcriptions. Nevertheless, making a list of all the relevant examples of causal (and causal only) ACs and SCs in so big a corpus, while keeping track of the subcorpora in which they appear, would require either an unlimited amount of time and manpower, or a sophisticated algorithm, which, to my present knowledge, remains to be invented.²⁸

BIBLIOGRAPHY

- Adamczewski H. (2000). "The Architecture of English Grammar: from the Theory of Phases to the Principle of Cyclicity", Plenary conference at the International Symposium of English Linguistics at the Université de Toulouse – Le Mirail. 7th and 8th July 2000. <http://www.linguistique.org/>
- Altenberg B. (1984). "Causal linking in spoken and written English", *Studia Linguistica* 38: 20-69.
- Biber D. et al. (1999). *Grammar of Spoken and Written English*. Harlow (UK): Longman.
- Bouscaren J., Chuquet J. and Danon-Boileau L. (1992). *Introduction to a Linguistic Grammar of English. An Utterer-Centered Approach*. Translated and adapted by J. Bouscaren and R. Flinham. Gap: Ophrys, 1992.
- Claridge C. and Walker T. (2001). "Causal clauses in written and speech-related genres in Early Modern English", *ICAME Journal* 25: 31-63. <http://clu.uni.no/icame/ij25/ccicame.pdf>
- Culioli A. (1995). *Cognition and Representation in Linguistic Theory*. Texts selected, edited and introduced by M. Liddle. Translated with the assistance of J. T. Stonham. Amsterdam: John Benjamins.
- . (1999). *Pour une linguistique de l'énonciation. Domaine notionnel*. Tome III. Paris: Ophrys.
- Dancygier B. and Sweetser E. (2000). "Constructions with *If*, *Since*, and *Because*: Causality, Epistemic Stance, and Clause Order", in E. Couper-Kuhlen and B. Kortmann (eds) *Cause-Condition-Concession-Contrast: Cognitive and Discourse Perspectives*. Berlin: Mouton de Gruyter, 111-42.
- De Cola-Sekali M. (1991). "Connexion inter-énoncés et relations intersubjectives: l'exemple de *because*, *since* et *for* en anglais", *Langages. Numéro spécial: Intégration syntaxique et cohérence discursive*. Mary-Annick Morel (ed.) 62-78.
- Delechelle G. (1980). "La valeur causale de *as* et *with*", *Études anglaises* 90 : 249-258.
- . (1989). "L'expression de la cause en anglais contemporain. Étude de quelques connecteurs et opérations" Thèse d'État. Paris: Université de Paris III – Sorbonne Nouvelle.
- . (1993). "Connecteurs et relations inter-énoncés" *Séminaire pratique de linguistique anglaise*. J.-R. Lapaire and W. Rotgé (eds.) 173-94. Toulouse: PU du Mirail.
- Dubois J. et al. (1999). *Dictionnaire de linguistique et des sciences du langage*. Nouvelle édition. Paris: Larousse-Bordas.
- Ducrot O. (1983). "Puisque: Essai de description polyphonique", *Revue Romane* 24: 166-85.
- (1984). *Le Dire et le Dit*. Paris : Les Editions de Minuit.
- Groussier M.-L. and Riviere C. (1996). *Les mots de la linguistique. Lexique de linguistique énonciative*. Paris: Ophrys.
- Guillaume B. (2012). "From ambiguity to deceptiveness: the case of hybrid *since*-subordinates in English" *E-rea* 9.2.
- . (2013). "Clause order in sentences containing a *since*-subordinate" *E-rea*. erea.revues.org/3476
- . (2014). *A Corpus-Based Study of Since-Clauses in Contemporary English*. Toulouse: PU du Mirail.
- Guimier C. (1997). "AS, connecteur inter-propositionnel: ses valeurs en discours", *Anglophonia-Sigma* 2: 5-36.

- Halliday M.A.K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Kennedy G. D. (1998). *An Introduction to Corpus Linguistics*. Harlow (UK): Longman.
- Leroux A. (2009). "Relations de cause en anglais et en français dans la presse: la place du lecteur", *Synergies Pologne* 6 : 93-102.
- . (2012). "La relation inter-énonciative et le marquage syntaxique des relations de cause: étude contrastive anglais-français", *CORELA - Paramétrer le sens? Etudes de cas*. <http://corela.edel.univ-poitiers.fr/index.php?id=2429>
- Meier E. (2002). "Causal subordination in English and Norwegian", *Nordic Journal of English Studies* 1.1: 33-64.
- Molencki R. (2007). "The Evolution of Since in Medieval English" *Connectives in the History of English*. U. Lenker and A. Meurman-Solin (eds.). Amsterdam (Netherlands): John Benjamins, 97-113.
- Nadvornikova O. (2017). "Pièges méthodologiques des corpus parallèles et comment les éviter", *CORELA*. HS 21. <https://journals.openedition.org/corela/4810>
- Poudat C. and Landragin F. (2017). *Explorer un corpus textuel: Méthodes - pratiques - outils*. Paris: De Boeck Supérieur.
- Ranger G. (2018). *Discourse Markers: An Enunciative Approach*. London: Palgrave Macmillan.
- Rissanen M. (1998). "Towards an integrated view of the development of English: Notes on causal linking" in J. Fisiak and M. Krygier (eds.) *Advances in English Historical Linguistics*. Berlin: de Gruyter, 389-406.
- Shahrokny-Prehn and Höche (2011). "Rising through the registers - A corpus-based account of the stylistic constraints on Light Verb Constructions" *Corpus* 10. <https://journals.openedition.org/corpus/2110>

NOTES

1. The examples which are preceded by a reference consisting of a few letters and numbers are all BCN examples. I shall come back at length to the corpora used to collect examples.
2. It can also be noted that both examples are taken from scientific textbooks, a genre which provides numerous examples of such causal relationships, and is extremely well represented in the BNC.
3. <https://www.korpus.cz/>
4. For an in-depth analysis of examples (3) and (4), see Guillaume 2014: 181-3. As for example (5), it was not part of my original corpus of SCs but it is one which I came across later, as I continue to collect examples of SCs (see also example 14).
5. My definition of an SC is a clause in which *since* introduces a finite verb clause. When using the BNC, I had to do away with a little more than 10% of the examples obtained (see chart above), in which *since* introduced a noun phrase or was used as an adverb. These, however, may not solely be due to parsing errors, but also to a different acceptance of what is a subordinating conjunction; see Guillaume 2014: 13-4 for more details.
6. The transcriptions of oral discourse only represent 10% of the BNC (Kennedy 1998: 50).
7. Indeed, the presence of an explicit subordinator to signal a cause and effect relationship is not necessary most of the time, as this type of relationship is often self-evident (*cf.* Leroux 2009;

2012). It may therefore be frequently omitted in oral speech. The same remark can apply to *as* as a causal subordinator (cf. Altenberg 1984: 41-3).

8. “la relation d’identité marquée par *as* se ramène à un lien de continuité logique entre *q* et *p* (on parle couramment de lien naturel, de prolongement, d’équivalence entre la cause et la conséquence, etc)” (Guimier 1997: 32)

“the relationship of identity signalled by *as* boils down to a link of logical continuity between *q* and *p* (it is often described in terms of innate compatibility, continuation, equivalence between cause and consequence, etc)” (my translation).

9. <http://ruscorpora.ru/>

10. However, they also say the same of the *because*-clauses, which is more questionable given the complexity of *because* as a subordinating conjunction. Indeed, *because* is one of the most frequently used conjunctions in order to express a cause and effect relationship nowadays, and cannot therefore be characterised in a few words.

11. In other related theories, such an opposition can be labelled in a different manner (thematic/rhematic for instance, in H. Adamczewski’s approach).

12. Halliday uses an initial capital letter for *Given* and *New*, because he considers them to be functions (1985: 31; 38).

13. I wish to thank J.-M. Merle for suggesting this phrase, but remain responsible for the use made of it here.

14. “une ‘franchise’, un laissez-passer linguistique” (1991: 69; my translation).

15. Altenberg (1984) compares data regarding the placement of reason clauses in the spoken material of the *London-Lund Corpus of Spoken English* and the written material of the *Lancaster-Oslo/Bergen Corpus of British English*. Among others, he deduces from his observations that the fronting of reason clauses is much more common in written texts, as this is the logical thing to do, but it may require some hindsight, whereas spontaneous speech more often treats reason clauses as afterthoughts, no matter how presupposed they may be, because there is usually not enough time to rationally organise one’s thoughts in the course of an oral discourse. My own material does not contain enough oral transcriptions, however, to allow me to discuss this view.

16. Regarding the SCs in my original corpus (the one used in Guillaume 2014), 199 examples are placed before the main clause (59%) while 134 appear after it (41%), and a couple of examples are interpolated with the main clause. As for the ACs in the corpus gathered with the help of my Master’s students, 105 examples are fronted (59%), against 73 examples appearing after the main clause (41%), and again one finds an odd example interpolated with the main clause.

17. See Deléchelle 1985: 531 for an analysis of these phrases.

18. The large number of books, PhD dissertations or again research articles concerning various languages – but mostly English or French – which have been inspired by the TOE in the past four decades are very often written in French, but there exist some translations or original works which endeavour to extensively present the TOE to an English-speaking readership. See for instance Bouscaren *et al.* (1992), Culioli (1995) or again Ranger (2018).

19. For more detail, see, among others, Groussier and Rivière 1996: 71-2 – or, in English: Bouscaren *et al.* 1992: 150-2.

20. See for instance Groussier and Rivière 1996: 71-2 for more detail on these symbols.

21. “*Since* s’appuie sur la thématité de *q* pour imposer la conclusion *p*. Une fois *q* admis, on ne peut plus ne pas admettre aussi *p*. [...]”

C’est cette double négation qui donne à l’énoncé son tour polémique, dans la mesure où l’on crée une image virtuelle d’un co-énonciateur qui refuserait d’admettre *p*.” (1989 : 560-1)

“*Since* relies on the presupposed quality of *q* in order to render *p* unavoidable. Once *q* has been validated, it becomes impossible not to admit that *p* is true as well. [...]”

It is such a double negation which gives a polemical flavour to the phrase, inasmuch as it

conjuges up the picture of an addressee who would not want to accept that *p* is true.” (my translation)

22. “When presenting an idea as presupposed, I pretend that the addressee and myself cannot do anything but take it for granted. If supposition is what I assert as a speaker, if implication is what I let the addressee conclude of their own accord, presupposition is what I present as shared opinion between both the speaker and the addressee, as some fundamental common ground which bonds the participants in the dialogue. With reference to the system of the personal pronouns, it can be said that presupposition equates to *we*, while supposition is best represented by *I* and implication by *you*.” (my translation)

23. For instance, “*parce que*” is more often regarded as a counterpart of *because*, “*vu que*” of *given that*, etc.

24. The proportion of each sub-corpus varies for each language. I shall come back to this point in more detail as far as French and English are concerned.

25. The difference between the number of hits announced and the one appearing in my statistics is probably imputable to human error (my own, then) rather than to the software.

26. The figures concerning the various versions (namely, the different languages) of Intercorp 10 are available at this address: <http://wiki.korpus.cz/doku.php/en:cnk:intercorp:verze10>.

27. “*Project Syndicate* produces and delivers original, high-quality commentaries to a global audience. Featuring exclusive contributions by prominent political leaders, policymakers, scholars, business leaders, and civic activists from around the world, we provide news media and their readers cutting-edge analysis and insight, *regardless of ability to pay*.” (<https://www.project-syndicate.org/about>)

28. “D’une manière générale, les annotations qui relèvent de la morphosyntaxe et de la syntaxe sont automatisables avec un taux de correction satisfaisant, alors que la plupart des annotations relevant de la sémantique et de la pragmatique sont soit non automatisables dans l’état actuel du TAL, soit automatisables mais avec un **taux d’erreur** non négligeable.” (Poudat and Landragin 2017: 42)

‘Generally speaking, annotations that pertain to morphosyntax and syntax can be automated with a satisfying rate of correction, whereas most semantic or pragmatic annotations are either not possible to automate in the present state of natural language processing, or may be automated but with a significant **error rate**.’ (my translation)

ABSTRACTS

This article deals with English causal subordinate clauses introduced by *since* or *as*. Both these markers may convey different meanings according to contextual variations, and can express temporal relations, from which their causal value is derived. The semantic closeness between *as* and *since* whenever they express a causal relation makes it necessary to harvest a large number of attested examples in order to compare and contrast them. The recourse to large on-line corpora such as the *British National Corpus* gives rise to specific practical difficulties, however, especially as far as *as* is concerned; because of the high frequency of the subordinators in question, one is confronted to thousands of examples, few of which turn out to have a clear causal value in the case of *as*, hence the recourse to other means in order to make up a large enough corpus of examples.

The concept of *presupposition* is examined, as it has often been argued that causal *since* subordinates, unlike *as* subordinates, introduce a presupposed content. But when confronted to a large number of examples, this criterion falls short of accounting for the subtle differences between the two conjunctions. A more theoretical approach is required; thanks to the tools provided by A. Culioli's *Theory of Enunciative and Predicative Operations*, it becomes possible to formalise the hypothesis that a *since* causal relationship is *presented* as unproblematic and is addressee-oriented, whereas an *as* causal relationship is *felt* as unproblematic, and is speaker-oriented.

This difference can be felt in the fact that it seems to be easier to find examples of causal *as* clauses in the press, while the causal use of *since* is more widespread in general, and does not seem to be specific to a genre in particular. In order to put this impression to the test, I have resorted to the multi-lingual parallel corpus *Intercorp*.

Nous étudions dans ce travail les propositions causales introduites par les subordonnants *since* ou *as* en anglais. Ces marqueurs peuvent avoir des significations différentes en fonction du contexte et leur valeur causale dérive de leur valeur temporelle. La proximité sémantique entre eux lorsqu'ils expriment la cause rend indispensable de travailler sur de nombreux exemples authentiques, afin de pouvoir les comparer et mettre au jour leurs différences. Néanmoins, dans ce cas précis, et plus particulièrement en ce qui concerne *as*, le fait d'utiliser de grands corpus disponibles en ligne, tel le *British National Corpus*, pose des problèmes pratiques bien spécifiques. En effet, en raison de la grande fréquence de ces conjunctions, on se trouve confronté à des milliers d'exemples dont un très petit nombre seulement illustre la valeur causale, surtout en ce qui concerne *as*, d'où la nécessité d'avoir recours à d'autres moyens afin de constituer un corpus d'exemples suffisant.

Nous examinons le concept de « présupposition », dans la mesure où il a souvent été écrit que le contenu des subordonnées causales en *since* serait préconstruit, contrairement à ce qui se passe avec *as*. Néanmoins, ce critère est loin d'être suffisant pour expliquer les subtiles nuances entre l'emploi de *as* ou bien de *since* avec une valeur causale dans un grand nombre d'exemples. En ayant recours au cadre théorique et aux outils de la *Théorie des Opérations Enonciatives* d'A. Culioli, nous proposons d'envisager *since* comme le marqueur d'une relation causale présentée comme non problématique et prenant en compte le co-énonciateur, tandis que *as* exprime pour sa part une relation causale ressentie par l'énonciateur comme n'étant pas problématique, sans pour autant qu'il ne se préoccupe véritablement de la position du co-énonciateur.

La différence d'emploi entre ces deux conjunctions se ressent lors de la recherche d'exemples, dans la mesure où il semble plus facile de trouver des exemples de *as* avec une valeur causale dans la presse, tandis que cette valeur de *since* est plus largement répandue d'une manière générale, et ne semble pas devoir correspondre à un genre en particulier. Nous proposons une expérience afin de tester cette impression, et avons recours pour cela au corpus parallèle multi-langues *Intercorp*.

INDEX

Keywords: *as*, *since*, causal subordinate clauses, presupposition, corpus linguistics, utterer-centred approach, *Intercorp*

Mots-clés: *as*, *since*, propositions subordonnées causales, présupposition, linguistique de corpus, linguistique de l'énonciation, *Intercorp*

AUTHOR

BÉNÉDICTE GUILLAUME

MCF-HDR, Université Côte d'Azur, CNRS, BCL, France