

John Benjamins Publishing Company



This is a contribution from *Languages in Contrast* 16:2
© 2016. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu. Please see our rights policy on <https://benjamins.com/content/customers/rights>

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Discourse connectives across languages

Factors influencing their explicit or implicit translation

Sandrine Zufferey

Université de Berne (Switzerland)

Coherence relations linking discourse segments can be communicated explicitly by the use of connectives but also implicitly through juxtaposition. Some discourse relations appear, however, to be more coherent than others when conveyed implicitly. This difference is explained in the literature by the existence of default expectations guiding discourse interpretation. In this paper, we assess the factors influencing impication by comparing the number of implicit and explicit translations of three polysemous French connectives in translated texts across three target languages: German, English and Spanish. Each connective can convey two discourse relations: one that can easily be conveyed implicitly and one that cannot be easily conveyed implicitly in monolingual data. Results indicate that relations that can easily be conveyed implicitly are also those that are most often left implicit in translation in all target languages. We discuss these results in view of the cognitive factors influencing the explicit or implicit communication of discourse relations.

Keywords: discourse connectives, discourse relations; translation, impication; English/French/German/Spanish

1. Introduction

Sentences forming a discourse are linked together by coherence relations such as *cause*, *concession*, and *contrast*. In some cases, these relations are marked explicitly by discourse connectives like *but*, *if* and *because* (Halliday & Hasan, 1976; Mann & Thomson, 1988; Sanders, Spooren & Noordman, 1992; Knott & Dale, 1994). In other cases, however, coherence relations are conveyed implicitly, in the absence of a connective. In fact, Taboada (2009) reports that up to 70% of discourse relations are not explicitly signaled by a lexicalized discourse connective in corpus data.

The reason for the rather low frequency of lexicalized connectives is that coherence relations can also be conveyed by a range of lexical, syntactic or graphic signals (Taboada, 2006). For example, the causal relation conveyed by *because* in (1) can also be conveyed by the use of a present participle as in (2).

- (1) Mary left the meeting early because she felt tired.
- (2) Feeling tired, Mary left the meeting early.

Similarly, the contrastive relation conveyed by *whereas* in (3) can easily be inferred in the absence of the connective as in (4), because the lexical opposition between the words *tall* and *small* and the parallel syntactic structure of the two clauses are strongly indicative of a contrast (Das & Taboada, 2013).

- (3) Eve is tall whereas Kate is small.
- (4) Eve is tall. Kate is small.

Importantly, however, in the absence of lexicalized connectives or alternative signals, some coherence relations seem to be intrinsically easier than others to reconstruct by inference. For example, the causal relation in (5) is easily inferred in the absence of a connective or an alternative signal. Indeed, using his world knowledge, the addressee of (5) concludes that being tired is a plausible reason for leaving the meeting early.

- (5) Mary was tired. She left the meeting early.

By contrast, the explicit concessive relation conveyed by *although* in (6) is lost in (7), and the implicit relation seems incoherent. In this case, using his world knowledge, the addressee cannot coherently relate the fact of losing one's job with a state of happiness.

- (6) Pete was happy, although he lost his job.
- (7) ?Pete was happy. He lost his job.

In this paper, we discuss the cognitive factors explaining why some coherence relations are easier to convey implicitly than others, and demonstrate that these factors affect the explicit or implicit translation of coherence relations that are conveyed by ambiguous lexicalized connectives.

Connectives are well known to be volatile items in translation, and translators often add, rephrase and remove them (Halverson, 2004; Zufferey & Cartoni, 2014). The study of translated texts thus offers a valuable case study for coherence relations, as analyzing the explicit and implicit translations of connectives across

several target languages provides an effective way to tease apart universal from language-specific aspects of text structuring.

The paper is organized as follows. In Section 2, we introduce the notion of temporal and causal discontinuity as factors explaining the use of connectives to explicitly mark discourse structure, and argue that the definition of discontinuity must be broadened to encompass all cases of coherence relations involving a perceptive shift between two mental spaces. In Section 3, we describe the meanings of three French causal connectives that can be used to communicate two distinct coherence relations (*dans la mesure où*, *en effet* and *or*). In all cases, the two coherence relations are unequal in terms of their potential for implicitation: one of them is predicted to be easily conveyed implicitly because of the cognitive factors discussed in Section 2, while the other is predicted to produce a loss of coherence when conveyed implicitly. In Section 4, we present a multilingual corpus study in which we assess these predictions by comparing the explicit and implicit translations of the three French connectives across three target languages: English, German and Spanish, using parallel directional corpora. The results presented in Section 5 demonstrate that discontinuous coherence relations give rise to more explicit translations compared to continuous relations. This effect is in addition coherent across the three target languages, and independent of the percentage of implicit relations produced by each connective, and the range of translation equivalents provided by each target language system. We discuss these results from the perspective of universal cognitive factors influencing discourse structuring, and provide further avenues of enquiry for the study of coherence relations across languages in Section 6.

2. Cognitive factors explaining the explicit communication of discourse relations

The difference between coherence relations that can and cannot be conveyed implicitly, that is in the absence of connectives or alternative signals, is accounted for in the literature by cognitive factors influencing readers' default expectations about discourse. More specifically, Murray (1995; 1997) formulated the 'continuity hypothesis' stating that readers or listeners expect events in a discourse to be presented in a temporally linear manner and/or to be causally related to each other. This principle has implications for the explicit or implicit communication of coherence relations. According to Murray (1997:228) "continuity can be conveyed easily via additive or causal relations. When a reader encodes a text event that is discontinuous in the absence of a marker or indication of the discontinuity, reading is more difficult". Following this principle, relations that do not introduce events in the order in which they occurred in the world, as for example temporal

relations with the connective *after* as in (8), are more difficult to process compared to temporal relations in which the world order is respected, such as temporal relations with the connective *before* as in (9).

(8) Mary went to bed after watching a movie.

(9) Mary watched a movie before she went to bed.

Sanders (2005) underlined more specifically the special status of causal relations as elements creating a high degree of coherence in a discourse. He formulated the ‘causality-by-default hypothesis’ stating that readers start out by assuming that the relation between discourse segments is causal, unless the semantic content of the segments prevents a causal interpretation. This principle explains why concessive relations are difficult to convey implicitly, as they involve the negation of a potential cause (Sanders, Spooren & Noordman, 1992).

Results from experimental studies confirm the existence of temporal and causal continuity principles as well as a default preference for causal interpretations, influencing discourse processing. Indeed, when readers have to read discontinuous concessive relations in the absence of a connective, reading is perturbed, but no similar effect occurs with continuous causal relations (Murray, 1997). The special status of causal relations is also corroborated by the fact that readers read faster and recall better a segment preceded by a causal connective than by an additive connective (Sanders & Noordman, 2000). Moreover, even in the absence of a connective, causal inferences influence the processing of upcoming words in a sentence (Kuperberg, Paczynski, & Ditman, 2011). Finally, when subjects are asked to continue a sentence ending with a period, their answers are often causally related to the first segment (Murray, 1997).

The validity of these cognitive factors has also been demonstrated to affect language use on the basis of corpus data. In the Penn Discourse Tree Bank corpus (Prasad et al., 2008), continuous relations such as causality, temporal sequence, and instantiation are conveyed implicitly much more often than discontinuous relations such as concession and temporal precedence (Asr & Demberg, 2012). This corpus study also revealed that other coherence relations like conditional relations were almost never conveyed implicitly. This phenomenon cannot, however, be explained in terms of causal or temporal discontinuity, as conditional relations are neither non-causal nor non-temporal.

In order to account for additional cases of discourse relations that are seldom conveyed implicitly, Zufferey and Gyax (2015) suggested broadening the notion of discontinuity to include all relations involving a form of perspective shift between the two related segments. They focused more specifically on the case of confirmation relations, as in (10).

- (10) John thought that the teacher was ill. And indeed, he did not come to school this morning.

In (10), the first segment is presented from the perspective of John, while the second segment is presented from the speaker's own perspective. According to Sanders (1994), the various narrative perspectives represented in a text can be thought of in terms of distinct mental spaces (Fauconnier, 1985). Shifting mental spaces in the absence of a connective has been demonstrated to disturb sentence processing in confirmation relations (Zufferey & Gygax, 2015). In addition, the use of lexical markers of attributed thoughts or utterances is correlated with the explicit communication of discourse relations in corpus data (Patterson & Kehler, 2013). These results thus tend to confirm that shifting narrative points of view require explicit linguistic marking.

In addition to confirmation relations, shifts in mental spaces have also been evidenced in some concessive and corrective relations (Pander Maat, 1998) and in conditional relations (Dancygier, 1999). We therefore submit that all coherence relations that involve a discontinuity due to shifts between mental spaces cannot easily be conveyed implicitly. We assess this hypothesis by comparing the implicitation of French connectives that convey continuous causal and temporal relations and discontinuous concessive, confirmative and conditional relations. From a theoretical perspective, this study will test the influence of a broader range of cognitive factors than temporal and causal discontinuity.

3. A case study involving three polysemous French connectives

In order to compare the explicit and implicit communication of coherence relations across languages, we selected three French connectives that are ambiguous between two distinct coherence relations. The study of ambiguous lexicalized connectives enables us to compare the communication of distinct discourse relations while keeping the connective constant, thus avoiding biases related to intrinsic differences between several connectives in terms of frequency, register, etc. This is an important factor given the fact that individual connectives display important differences in the number of explicit relations they convey in corpus data (Patterson & Kehler, 2013).

The three connectives that we selected can all be used to convey two coherence relations that are not equal in terms of implicitation. While one of them is continuous and should therefore be easily conveyed implicitly, the other is discontinuous and should not be easily conveyed implicitly. We illustrate the relations conveyed by these connectives in this section, using real examples from our corpus.

First, according to the French lexicon of connective LEXCONN (Roze, Danlos & Muller, 2012), the French connective *dans la mesure où* can either convey a causal relation as in (11) or a conditional relation as in (12).

- (11) Cette évolution est importante et doit être maintenue *dans la mesure où* c'est l'avenir de l'union qui est en jeu.
 “This is an important development which must be supported as it is the future of the EU which is at stake.”
- (12) Cette association des peuples ne sera possible que *dans la mesure où* un bon climat général existera dans l'opinion à l'égard du travail de l'Union et de ses institutions.
 “It will only be possible to involve people in this way if public opinion is generally well-disposed towards the work of the European Union and its institutions.”

The relation of causality is continuous and therefore expected to be easily conveyed implicitly, while the relation of condition is expected to be more difficult to convey implicitly, because it involves a shift between two mental spaces (see Section 2).

Second, studies focusing on the French connective *en effet* (Charolles & Fagard, 2012; Danlos, 2012; Iordanskaja & Mel'čuk, 1999; Rossari, 2002; Zufferey & Gygax, 2015) all indicate that this connective is ambiguous between a relation of causality as in (13) and a relation of confirmation as in (14).

- (13) Monsieur le président, nous ne pouvons pas voter en faveur de ce rapport.
En effet, l'Union Européenne devrait mettre le social au coeur de sa construction.
 “Mister President, we cannot vote in favour of this report. The European Union should actually place social issues at the heart of the building of Europe.”
- (14) La commission propose le rejet cette initiative. C'est *en effet* assez compliqué.
 “The commission suggests rejecting this initiative. It is indeed quite complicated.”

Again, the relation of causality is hypothesized to be easily conveyed implicitly while the relation of confirmation is not, because it involves a perspective shift between the mental space of an external speaker in the first segment and the speaker's own mental space in the second segment. In the case of (14), the external perspective is that of the committee in the first segment, while the speaker's own perspective is presented in the second segment.

Finally, according to LEXCONN, the French connective *or* is ambiguous between a relation of background introducing either a temporal or an argumentative continuation as in (15), and a relation of concession as in (16).

- (15) Le progrès technologique est un slogan trop connu. *Or*, la mutation à laquelle nous assistons est analogue à celle de la naissance de l'imprimerie.
“Technological progress is a familiar slogan. In fact, the transformation we are witnessing is reminiscent of what happened with the invention of printing.”
- (16) La liberté ne se décrète pas. *Or*, la construction communautaire ne semble pas consciente de cette réalité.
“Liberty cannot be decreed. But those who are building Europe seem unaware of that fact.”

The relation of background is either temporally or causally continuous while the relation of concession is discontinuous. It is therefore expected that the latter will be conveyed implicitly less often than the former.

In sum, each connective can convey either a continuous or a discontinuous relation, and we expect this difference to affect the number of implicit translations for each of them.

4. A multilingual corpus study

We conducted a multilingual corpus study to compare the explicit and implicit translations of discourse relations conveyed by the three French connectives discussed in Section 3.

In this study, we used parallel directional corpora extracted from the Europarl corpus (Koehn, 2005), forming the Europarl direct corpus (Cartoni, Zufferey & Meyer, 2013). The Europarl corpus is freely available, and automatically sentence-aligned for research purposes. It contains debates from the European Parliament. During the parliamentary sessions, deputies speak in their mother tongue and statements are later transcribed and translated into all the other official languages of the European Union. In addition to representing a large variety of languages, one of the main advantages of Europarl is that statements are produced by a variety of different speakers and translated by a large team of expert translators, thus avoiding individual biases in language use. In addition, because the context of production is constant for all the languages, the various sub-parts of the corpus are very homogeneous (Cartoni & Meyer, 2012), enabling sound cross-linguistic comparisons.

For the experiments reported in this paper, we used the sub-parts of Europarl encompassing original French texts and their translations in English, German and Spanish. In the French part of the corpus, we randomly extracted 200 occurrences of the three ambiguous connectives discussed in Section 3: *dans la mesure où*, *en effet*, and *or* with their aligned translations, using the bilingual concordance tool Paraconc (Barlow, 2008). We then proceeded to annotate all the occurrences of French connectives with the appropriate coherence relation. All occurrences of *dans la mesure où* were tagged with either a *cause* or a *condition* tag. All occurrences of *en effet* were tagged as either *cause* or *confirmation* tag. Finally, all relations conveyed by *or* received either a *background* or a *concession* tag. In order to ensure the reliability of the annotation, a random sample of 20% of the relations was independently annotated by a second annotator. Results indicate that the two annotators agree in 90% for the annotation of *dans la mesure où*, in 88% of the cases for *en effet* and in 90% of the cases for *or*.

In addition to the annotation of discourse relations in French, we manually spotted the translation equivalents for all occurrences of the three French connectives in English, German and Spanish translated texts.

In this experiment, our goal was to compare the number of explicit and implicit translations of each connective depending on which relation they conveyed in context. We included in the category of implicit translations, noted as 'zero' in Section 5, only cases where the connective was replaced by nothing, a full stop or a comma. Alternative syntactic constructions and reformulations using non-lexicalized connectives were treated as explicit translations, as the meaning of the relation did not have to be inferred but was given by linguistic signals.

5. Results: comparing implicitation across languages and relations

In this Section, we report the explicit and implicit translations of three French causal connectives and give their translation equivalents in English, German and Spanish. We include in the list of explicit translation equivalents only lexical items representing at least 10% of the total number of occurrences, in order to ensure a comparable representation of main translation equivalents across languages. All infrequent translations are grouped in the category called 'other'.

We compared implicit relations, marked as zero, and explicit relations, corresponding to all other translation choices including equivalent connectives, phrases and syntax. In order to analyze the data, we used 2x2 tests with two factors: explicit or implicit translation and the two possible discourse relations. We used either χ^2 tests of independence, or when the expected value of one of the cells was inferior to 5, Fisher's exact test. Given that we have clear hypotheses about the

direction of the effects, we systematically used one-tailed tests. The translations of *dans la mesure où* are reported for each target language separately, in Tables 1 to 3.

Table 1. Translations of *dans la mesure où* in English per relation.

	<i>insofar as</i>	<i>as</i>	<i>because</i>	<i>zero</i>	<i>in that</i>	<i>inasmuch as</i>	<i>to the extent that</i>	<i>other</i>	Total
Cause	39	15	13	13	11	7	11	44	153
Condition	23	0	1	0	1	5	1	16	47
Total	62	15	14	13	12	12	12	60	200

Table 1 indicates that *dans la mesure où* has seven frequent translations in English, among which are ‘zero’ translations, in other words implicitations. We compared the number of implicit (zero) and explicit translations of *dans la mesure où* (all the other translation choices), and found that the relation of cause lead to significantly more zero translations compared to the relation of condition (Fisher’s exact test $p < 0.05$). It is noticeable that zero translations were never found in cases of conditional relations.

Table 2. Translations of *dans la mesure où* in German per relation.

	<i>da</i>	<i>(in)sofern</i>	<i>in dem Maße</i>	<i>wenn</i>	<i>zero</i>	<i>other</i>	Total
Cause	75	22	5	1	12	40	155
Condition	4	8	8	11	0	14	45
Total	79	30	13	12	12	54	200

When German is the target language, *dans la mesure où* has five frequent translations, including zero translations. We compared the number of zero translations of *dans la mesure où* between the two discourse relations, and found that the relation of cause lead to more zero translations in German compared to the relation of condition (Fisher’s exact test $p < 0.05$). Again, relations of condition do not lead to any implicit translation.

Table 3. Translations of *dans la mesure où* in Spanish per relation.

	<i>en la medida en que</i>	<i>zero</i>	<i>other</i>	Total
Cause	123	5	30	158
Condition	31	0	11	42
Total	154	5	41	200

In Spanish, *dans la mesure où* only has one frequent translation, and implicitations represent only 2.5% of the data. In Spanish, the relation of condition does not lead

to any implicitation, as was also the case in English and German. However, given the very low number of implicitations (only five in total), the difference between the two relations is not statistically significant (Fisher's exact test, $p = 0.28$).

Table 4. Translations of *en effet* in English per relation.

	zero	indeed	in fact	other	Total
Cause	81	20	30	36	167
Confirmation	7	14	2	10	33
Total	88	34	32	46	200

The translations of *en effet* are reported separately for each target language in Tables 4 to 6. In English, the connective *en effet* has three frequent translations, among which zero translations are the most frequent. When compared across cause and confirmation relations, the number of implicitations is significantly more numerous in cases of causal relations compared to confirmation relations ($\chi^2 = 8.33$, $df = 1$, $p < 0.01$).

Table 5. Translations of *en effet* in German per relation.

	zero	denn	nämlich	in der Tat	so	other	Total
Cause	61	30	23	13	17	24	168
Confirmation	4	2	2	11	0	13	32
Total	65	32	25	24	17	37	200

In German, *en effet* has five frequent translations. Like in English, zero is the most frequent translation choice. When implicitations are compared across coherence relations, they are found to be more numerous in cases of causal relations compared to confirmation relations ($\chi^2 = 6.95$, $df = 1$, $p < 0.01$).

Table 6. Translations of *en effet* in Spanish per relation.

	en efecto	zero	efectivamente	other	Total
Cause	92	36	23	4	156
Confirmation	25	4	11	3	44
Total	107	40	34	7	200

In Spanish, *en effet* has three frequent translations, and contrary to German and English, zero translations represent only the third translation choice. In spite of this difference, when implicitations are compared across relations, they are also significantly more numerous in cases of causal relations compared to confirmation relations ($\chi^2 = 4.2$, $df = 1$, $p < 0.05$).

Results for the connective *or* are presented separately for each target language in Tables 7 to 9.

Table 7. Translations of *or* in English per relation.

	zero	<i>however</i>	<i>but</i>	<i>well</i>	<i>yet</i>	other	Total
Continuation	59	6	7	9	2	33	116
Concession	11	26	17	5	10	5	84
Total	70	32	24	14	12	48	200

In English, the connective *or* has five frequent translations, and zero translations represent the most frequent choice. When compared across relations, implications are significantly more numerous when *or* conveys a relation of continuation compared to a relation of concession ($\chi^2 = 30.54$, $df = 1$, $p < 0.001$).

Table 8. Implication of *or* in German per relation.

	<i>jedoch</i>	<i>nun</i>	zero	<i>aber</i>	<i>doch</i>	other	Total
Continuation	21	25	24	9	12	23	114
Concession	34	8	7	14	9	14	86
Total	55	33	31	23	21	37	200

In German, *or* has five frequent translations. Contrary to English, zero translations represent only the third most frequent option. When compared across languages, the number of implicit translations is significantly more numerous in cases of continuation than concession relations ($\chi^2 = 5.75$, $df = 1$, $p < 0.05$).

Table 9. Implication of *or* in Spanish per relation.

	<i>ahora bien</i>	<i>sin embargo</i>	zero	other	Total
Continuation	75	10	10	13	108
Concession	69	14	2	7	92
Total	144	24	12	20	200

Finally, in Spanish, *or* has only three frequent translations, and zero translations represent the less frequent choice. When compared across relations, the implications produced by *or* in Spanish are also significantly related to the relation of continuation ($\chi^2 = 4.42$, $df = 1$, $p < 0.05$).

6. Discussion

We compared the number of explicit and implicit translations of three French connectives that can all be used to convey two discourse relations. We hypothesized that in each case, one discourse relation would lead to more implicit translations than the other, due to cognitive biases in discourse interpretation. Our results clearly confirmed this prediction. For all connectives and relations, the discontinuous discourse relations of condition, concession and confirmation lead to a higher number of explicit translations compared to continuous causal and temporal relations. The difference was statistically significant in all but one case, the translation of *dans la mesure où* in Spanish that lead to a very low number of implicitations.

From a theoretical perspective, our results tend to confirm that considering all cases of relations involving a perspective shift between two mental spaces as discontinuous is on the right track. A noticeable difference emerged however between on the one hand relations of concession and confirmation that consistently yielded a small number of implicit translations, and on the other hand conditional relations that lead to no implicit translation at all. The reason for this difference may be that removing the connective produces a different effect in both cases. More specifically, contrary to concessive and confirmation relations, removing the connective in conditional relations does not produce a loss of coherence, but the interpretation of the relation becomes causal rather than conditional, due to the causality-by-default bias. For this reason, removing the connective seems to be harmful for conditional relations, as the intended relation is lost. Further work targeting more specifically the processing of conditional relations as well as the alternative lexicalizations that can be used to express them should shed further light on this observation.

Even though the number of implicitations was variable across target languages and connectives, some regularities emerged from our translation data. First, for all connectives, English produced the highest number of implicit translations, followed by German and lastly Spanish, a language in which the number of implicit translations was consistently low. In addition, for all connectives, Spanish also had a lower range of translation equivalents compared to English and German. These differences can be explained by the fact that Spanish and French are both Romance languages that in many cases possess very close translation pairs. Some of them like *en effet* and *en efecto* are even cognate words that evolved from a common source in Latin (Bertin, 2002; Fagard, 2011). In non-corpus based theories about translation, Romance languages are sometimes considered to be more cohesive than Germanic languages (Vinay & Darbelnet, 1958; Hervey & Higgins, 1992). Our data seem to corroborate this intuition, but this result should be replicated

through larger-scale analyses between the two language families before firm conclusions can be reached in this respect. A cross-linguistic study of cohesion should in addition encompass the use of cohesive devices at large in order to assess the possibility that these languages resort to alternative means than connectives to express discourse cohesion.

Similarities in the number of implicitations across languages were also found when comparing the translation equivalents on the three connectives. In all target languages, *en effet* produced the highest number of implicitations, followed by *or* and lastly *dans la mesure où*. This difference suggests that some intrinsic properties of connectives make them more or less likely to be translated implicitly. This result corroborates Patterson and Kehler's (2013) finding that individual connectives are strong predictors for the implicit or explicit realization of discourse relations. Striking differences are also found between connectives conveying the same coherence relation. In their comparison of English and French causal connectives, Zufferey and Cartoni (2012) reported that the number of implicit translations of French causal connectives ranged from 4.5% for *parce que* to 7.5% for *car*. By contrast, in our study, *en effet* produced 44% of implicit translations in English. A possible explanation for this important difference is that the causal uses of *en effet* appear mostly in sentence initial position, with the result that the two related segments are distinct sentences. This separation implies that juxtaposition becomes a salient translation choice (Danlos, 2012). Similar features of connectives that could explain these wide differences should be more systematically analyzed in future work.

In our study, we categorized coherence relations in terms of continuity or discontinuity only. Future work should also include features characterizing the related arguments, for example the use of pronouns or full referents, as well as the dependencies between relations, such as embedded versus non-embedded relations. Indeed, using statistically trained classifiers, Patterson & Kehler (2013) found that these features were useful to automatically predict the explicit or implicit nature of a relation. In addition, the role of connectives versus alternative signals for the explicit or implicit translation of discourse relations should also be systematically compared in future work.

Our study focused on the implicitation of discourse relations in translation. Another important direction for future work will be the comparison between cases of implicitation and explicitation in translation. Indeed, translators tend to add connectives more often than to remove them in translation, corresponding to the supposed translation universal of explicitation (Blum-Kulka, 1986). From the perspective of discourse connectives, it is expected that connectives conveying continuous relations will be added more often by translators than connectives conveying discontinuous relations. As discontinuous relations are more frequently used

to mark relations in source texts (Asr & Demberg, 2012), they provide less room for explicitation in translation than continuous relations. Zufferey and Cartoni (2014) confirmed that the connective *en effet* produced significantly more explicitations when it conveyed a causal relation compared to a confirmative relation in translated French texts. Similar assessments of explicitation should be conducted with a broader range of connectives to strengthen the validity of this observation.

In sum, our study provided some confirmation that the cognitive factors affecting the explicit or implicit communication of discourse relations in monolingual data also influence the explicit or implicit translation of connectives conveying them. Future work will benefit from the integration of a broader range of languages pertaining to diverse language families, in order to deepen our understanding of cross-linguistic differences and similarities in discourse structuring.

Acknowledgement

This work was financed by an Ambizione grant from the Swiss National Science Foundation for the MULDIS project (PZ00P1_148024).

References

- Asr, F. and Demberg, V. 2012. Implicitness of Discourse Relations. Proceedings of the 24th International Conference on Computational Linguistics, 2669–2684. Mumbai, India.
- Barlow, M. 2008. Parallel Texts and Corpus-Based Contrastive Analysis. In *Current Trends in Contrastive Linguistics. Functional and Cognitive Perspectives*, M. de los Ángeles Gómez González et al (eds), 101–121. Amsterdam: John Benjamins. doi:10.1075/sfsl.60.o8bar
- Bertin, A. 2002. L'Emergence du Connecteur *En Effet* en Moyen Français. *Linx* 46:37 – 50. doi:10.4000/linx.90
- Blum-Kulka, S. 1986. Shifts of Cohesion and Coherence in Translation. In *Interlingual and Intercultural Communication*, J. House and S. Blum-Kulka, 17–35. Tübingen: Narr.
- Cartoni, B. and Meyer, T. 2012. Extracting Directional and Comparable Corpora from a Multilingual Corpus for Translation Studies. Proceedings of *LREC 2012*, Istanbul, Turkey, May 23–25 2012.
- Cartoni, B., Zufferey, S. and Meyer, T. 2013. Using the Europarl Corpus for Linguistic Research. *Belgian Journal of Linguistics* 27:23–42. doi:10.1075/bjl.27.o2car
- Charolles, M. and Fagard, B. 2012. *En Effet* en Français Contemporain: de la Confirmation à la Justification/Explication." *Le Français Moderne* 80: 171–197.
- Dancygier, B. 1999. *Conditionals and Prediction. Time, Knowledge and Causation in Conditional Constructions*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511486463

- Danlos, L. 2012. Formalisation des Conditions d'Emploi des Connecteurs 'En Réalité' et '(Et) En Effet'. Proceedings of the *Third Congrès Mondial de Linguistique Française*, Lyon, France, 493 – 508.
- Das, D. and Taboada, M. 2013. Explicit and Implicit Discourse Relations. A Corpus Study. Proceedings of the *2013 Annual Conference of the Canadian Linguistic Association*, Alberta, Canada.
- Fagard, B. 2011. La Construction 'En Effet' dans les Langues Médiévales Romanes et la Question de l'Emprunt." *Oslo Studies in Language* 3:26–69.
- Fauconnier, G. 1985. *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge, MA: Bradford.
- Halliday, M. and Hasan, R. 1976. *Cohesion in English*. London: Longman.
- Halverson, S. 2004. Connectives as a Translation Problem. In *An International Encyclopedia of Translation Studies*, H. Kittel et al (eds), 562–572. Berlin/New York: Walter de Gruyter.
- Iordanskaja, L. and Mel'čuk, I. 1999. Textual Connectors across Languages: French *En Effet* vs. Russian *V Samon Dele*. *RASK* 9/10:305–347.
- Koehn, P. 2005. Europarl: A Parallel Corpus for Statistical Machine Translation. Proceedings of the *10th Machine Translation Summit*, Phuket, Thailand, 79–86, September 13-15.
- Knott, A. and Dale, R. 1994. Using Linguistic Phenomena to Motivate a Set of Coherence Relations. *Discourse Processes* 18(1):35–62. doi:10.1080/01638539409544883
- Kuperberg, G. Paczynski, M. & Ditman, T. 2011. Establishing Causal Coherence across Sentences: An ERP Study. *Journal of Cognitive Neuroscience* 23:1230 – 1246. doi:10.1162/jocn.2010.21452
- Hervey, S. and Higgins, I. 1992. *Thinking Translation. A Course in Translation Method, French-English*. London: Routledge.
- Mann, W. and Thomson, S. 1992. Relational Discourse Structure: A Comparison of Approaches to Structuring Text by 'Contrast'. In *Language in Context: Essays for Robert E. Longacre*, S. Hwang and W. Merrifield (eds), 19–45. Dallas: SIL.
- Murray, J. 1995. Logical connectives and local coherence. In *Sources of Cohesion in Text Comprehension*, R. Lorch and E. O'Brien (eds.), 107 – 125. Hillsdale, NJ: Erlbaum.
- Murray, J. 1997. Connectives and Narrative Text: The role of Continuity. *Memory and Cognition* 25: 227–236. doi:10.3758/BF03201114
- Pander-Maat, H. 1998. Classifying negative coherence relations on the basis of linguistic evidence. *Journal of Pragmatics* 30: 177–204. doi:10.1016/S0378-2166(98)00024-1
- Patterson, G. and Kehler, A. 2013. Predicting the presence of discourse connectives. Proceedings of the *2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, 914–923.
- Prasad, R. Dinesh, N. Lee, A. Miltsakaki, E. Robaldo, L. Joshi, A. and Webber B. 2008. The Penn Discourse Treebank 2.0. Proceedings of the *6th International Conference on Language Resources and Evaluation*, Marrakesh, Morocco, 2961–2968.
- Rossari, C. 2002. Les Adverbes Connecteurs: Vers une Identification de la Classe et des Sous-Classes. *Cahiers de Linguistique Française* 24:11–43.
- Roze, C. Danlos, L. and Muller, P. 2012. *LEXCONN: A French Lexicon of Discourse Connectives*. *Discours* 10, published online.
- Sanders, J. 1994. *Perspective in Narrative Discourse*. Ph.D. dissertation, Tilburg University, The Netherlands.

- Sanders, T. 2005. Coherence, Causality and Cognitive Complexity in Discourse. *Proceedings of the First International Symposium on the Exploration and Modelling of Meaning*, Biarritz, France, 105–114.
- Sanders, T. and Noordman, L. 2000. The Role of Coherence Relations and their Linguistic Markers in Text Processing. *Discourse Processes* 29:37 – 60. doi:10.1207/S15326950dp2901_3
- Sanders, T. Spooren, W. and Noordman, L. 1992. Towards a Taxonomy of Coherence Relations. *Discourse Processes* 15(1):1–36. doi:10.1080/01638539209544800
- Taboada, M. 2006. Discourse Markers as Signals (or not) of Rhetorical Relations.” *Journal of Pragmatics* 38:567–592. doi:10.1016/j.pragma.2005.09.010
- Taboada, M. 2009. Implicit and explicit coherence relations. In *Discourse, of Course*, J. Renkema (ed.), 125 – 138. Amsterdam, The Netherlands: John Benjamins. doi:10.1075/z.148.13tab
- Vinay, J.-P. and Darbelnet, J. 1958. *Comparative Stylistics of French and English. A Methodology for Translation*. Amsterdam and Philadelphia: John Benjamins. doi:10.1075/btl.11
- Zufferey, S. and Cartoni, B. 2012. English and French Causal Connectives in Contrast. *Languages in Contrast* 12(2):232–250. doi:10.1075/lic.12.2.06zuf
- Zufferey, S. and Cartoni, B. 2014. A Multifactorial Analysis of Explicitation in Translation. *Target* 26:361–384. doi:10.1075/target.26.3.02zuf
- Zufferey, S. and Gygax, P. 2015. The Role of Perspective Shifts for Processing and Translating Discourse Relations. *Discourse Processes*. published online ahead of print. doi: 10.1080/0163853X.2015.1062839

Author's address

Sandrine Zuffrey
Institut de Langue et de Littérature françaises
Université de Berne
Länggassstrasse 49
CH-3000 Berne 9 - Suisse
sandrine.zufferey@rom.unibe.ch