TRANSLATIONAL EXPLICITATION OF POLYSEMOUS ENGLISH DISCOURSE-CONNECTIVES. CORPUS-BASED CASE STUDIES

Sorina Postolea, PhD, "Al. Ioan Cuza" University of Iaşi

Abstract: Whether in the form of conjunctions, adverbials, prepositional phrases, or various other lexical combinations, discourse connectives play a crucial role in the way information is organised and conveyed in any form of communicative interaction. They function as complex signals of coherence relations, triggering various ways of interpreting the relationships established between the chunks of discourse they connect. Whereas some connectives are used to specify a single type of discourse relation (e.g. "thus" to indicate the result in a causal relation), others are polysemous, serving as procedural markers for the inference of several relations. Based on a parallel corpus, this paper examines five polysemous English discourse connectives ("so", "but", "while", "as", and "since") and their translations into Romanian. The analysis will show that the translation of these units is often accompanied by a process of explicitation whereby the different meanings of the source connectives are rendered unambiguous (or less ambiguous) in the target language.

Keywords: discourse connectives, translation, translation explicitation, coherence relations, parallel corpora, PDTB

Discourse connectives (DCs), also known as *discourse markers* (DMs), form the pragmatic (Fraser, 1999, p. 950) and functional (Zufferey & Degand, 2013, p. 1) category of lexical items whose main function is to signal the existence of a specific semantic relation between at least two chunks of discourse. From a grammatical-syntactic point of view, DCs may usually take the form of conjunctions, adverbials, or prepositional phrases, whereas from a semantic-pragmatic perspective, DCs are seen as procedural markers which triggervarious ways of interpreting the relationships established between the segments of discourse they connect. Taking into account the definition of *cohesion* as "an overt relationship holding between parts of the text, expressed by language specific markers" (Blum-Kulka, [1986]2004, p. 299) and of a *coherence relation* as "an aspect of meaning of two or more discourse segments that cannot be described in terms of the meaning of the segments in isolation" (Sanders, Spooren, & Noordman, 1992, p. 2), it could be said that DCs function as both *cohesion-* and *coherence-structuring devices*. However, whereas some connectives are used to signal a single type of discourse relation (e.g. "thus" to indicate the result in a causal relation), others are polysemous, serving as procedural markers for the inference of several relations. As Fraser puts it, DMs "have a core meaning which is procedural, not conceptual, and their more specific interpretation is 'negotiated' by the context, both linguistic and conceptual" (Fraser, 1999, p. 950). Or, in other words,

"depending on the context, the content of the arguments and possibly other factors, discourse connectives, just like verbs, can have more than one sense" (PDTB-Group, 2007, p. 26).

Starting from the observation that most languages possess a pre-defined set of DCs, but "they vary tremendously in the number of connectives they have to express relations and in the use they make of them" (Cartoni, Zufferey, & Meyer, 2013, p. 66), this paper focuses on five polysemous English discourse connectives ("so", "but", "while", "as", and "since") and their translations into Romanian. The merits of the translation-based study of discourse connectives within a multilingual framework have been highlighted by various authors (e.g. Noël, 2003; Aijmer, Foolen, & Simon-Vandenbergen, 2006; Degand, 2009; Zufferey, 2013), because "translators are language users whose linguistic choices are not only informative about the language they are producing, they are also highly indicative of their interpretation of the language they are receiving, and this interpretation is revelatory of the nature of the language that is received"(Noël, 2003). Among other things, translation-based studies of DCs may reveal important differences in the use and the available stock of DCs across different languages since, when seen from a contrastive perspective, "there is a general correspondence between the markers, but certainly not an exact mapping" (Fraser, 1999, p. 950). Translation-based analysis may also be a reliable way of investigating the polysemous nature of connectives both in the source and in the target language, as shown by Degand, 2009. For annotation and automation purposes, translation-based studies of DCs may serve as a path towards the disambiguation of various connective meanings in context (Cartoni, Zufferey, & Meyer, 2013). Moreover, in line with the approach adopted in this paper, translation-based analyses of DCs may bring important insight into the processes of text/discourse interpretation, processing, and (re-)ordering which emerge within and as a result of the process of translation itself.

Methodology

This study starts from two interconnected premises: on the one hand, 1) the translations of polysemous English DCs into Romanian may be used a "as a heuristics to uncover the meaning of Discourse Markers" (Degand, 2009), both in the source and target language and, on the other hand, 2) this type of disambiguation in context may be seen as a kind of explicitation prompted by the process translation itself. As shown by Zufferey & Cartoni (2014), the explicitation phenomena that accompany the translation of DCs may take various shapes, being dependent on multiple factors (e.g. the nature of both the source and the target language, the specific traits of the connective or of the discourse relation at hand). Moreover, these explicitation phenomena may be studied using various applied analyses. However, due to the limited amount of space available here and the lack of previous studies involving English and Romanian and of annotated resources, this study aims to be just a first, exploratory step into a more in-depth analysis of these phenomena.

The analyses carried out in this study draw on *parallel corpus* built by the author for her doctoral research. It comprises *275 parallel text pairs* (English source texts/ Romanian target texts), which amount to a total number of 548,591 words (268,342 for English and 280,249 for

Romanian). The texts were retrieved manually from the Internet and fully aligned at sentence-level by the author. They refer to ICT products and technologies and belong to *four textual genres* of general use in this field, i.e. ICT news articles, ICT press releases, ICT product descriptions, and ICT user manuals.

As mentioned before, this study focuses on five English discourse connectives. The analysis of their senses is based on the descriptions provided by the annotation manual of the Penn Discourse TreeBank (PDTB). PDTB is "a large-scale resource of annotated discourse relations and their arguments over the 1 million-word Wall Street Journal (WSJ) Corpus" (Prasad *et al.*, 2008). The five connectives were chosen due to 1) their polysemous nature (Prasad *et al.*, 2008; Degand, 2009; Zufferey & Degand, 2013), 2) their high frequency in the PDTB corpus, and 3) their relatively high frequency in the corpus at hand. Their translations were retrieved with *ParaConc*. The rough data provided by this analysis toolwere further processed by the author so as to discard the non-connective uses of the conjunctions at hand, as described in the literature (Halliday & Hasan, 1976; Degand, 2000; Zufferey & Degand, 2013). Moreover, the study draws on the *translation spotting* technique, i.e. "an annotation method that makes use of the translation of specific lexical items in order to disambiguate them" (Cartoni, Zufferery & Meyer, 2013, p. 68).

Findings

Corpus analysis revealed a wide array of possible translations for each of the five English connectives taken into account, ranging from 15 different items in the case of the connective *so*, to 7 in the case of *but*. For each DC, we also counted the instances in which the source DC was expressed by other means in the target language: e.g. omission (1), non-finite structures as in (2), punctuation (3). These instances are highlighted in red in each graph shown below.

(1) *That **because** there's WiFi N and a high-speed broadband option built in.*
Sunt incorporate tehnologia WiFi N şi o opţiune de bandă largă de mare viteză.

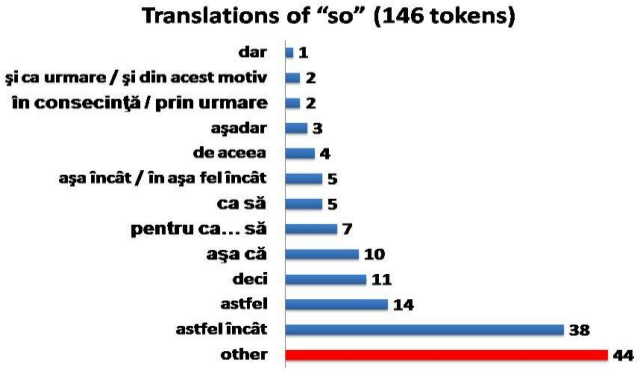
(2) *Use the tether to attach the stylus to the computer **so**you will not lose it.*
Ataşati stiloul de computer, folosind dispozitivul de fixare, pentru a nu-l pierde.

(3) *Do not remove the outer cover, **as** this may result in electric shock.*
*Nu îndepărtaţi stratul izolator**;** pericol de electrocutare.*

So

The English connective *so* produced the greatest number of occurrences in the corpus, i.e. 146 tokens. According to the PDTB manual, this DC is basically associated with just one type of discourse relation, i.e. "CONTINGENCY: Cause: result". However, its Romanian translations in the corpus show a slightly different picture. Although the compound *so that* was excluded from this analysis, its closest Romanian translation, *astfel încât* (Fr: *de sorte que*), was actually the most frequent translation of *so* in our corpus, with 38 tokens. Thus, there seems to be a functional subdivision within the "result" sense subtype described for *so* in the PDTB manual: 1) *aşa încât/ în aşa fel încât/ ca să/ pentru ca (să)/ astfel încât* seem to correspond to a purpose-

oriented meaning of *so*, as in (4) and (5) below, whereas 2) the other Romanian connectives shown in the graph seem to signal a conclusive/result meaning of *so* (Fr. *donc*/*par conséquent*), as in (6) and (7).



(4) *...you can simultaneously charge your mobile device **so** you never run out of power halfway*

*...puteţi să vă încărcaţi simultan dispozitivul mobil **astfel încât** să nu vă lase fără alimentare la mijlocul drumului*

(5) *...and has a line-in jack and headphone jack **so** you can use it with an mp3 using and for private listening*

*...şi mai au o mufă line-in li una pentru casti, **ca să** le puteţi folosi cu un mp3 player şi pentru auditii private*

(6) *The software default is to use the fastest setting **so** no user-intervention is required.*

*Software-ul are ca opţiune implicită utilizarea celei mai rapide setări, **deci** nu este necesară intervenţia utilizatorului*.

(7) *Flash memory card specifications constantly change **so** compatibility may change without warning.*

*Specificaţiile cardului de memorie flash se schimbă în mod constant, **prin urmare** compatibilitatea se poate schimba fără avertizare prealabilă.*

In terms of distribution, 55 translations of *so* signal a relation of *purpose* whereas 44 other instances convey a conclusive meaning, similar to that described by *therefore* in English.

Another interesting aspect which seems to be worth noting in the case of *so* is the great number of cases in which it was translated by other means in Romanian. In fact, in 36 out of the 44 instances accounted for in the graph above, the source DC was actually translated by a Romanian non-finite clause, i.e. the structure *pentru+infinitive*:

(8) *Set up your online accounts **so** you can send and receive email, and more.*

Configuraţi-vă conturile online pentrua putea trimite şi primi e-mailuri şi altele.

On the one hand, seeing that this structure is mainly used to mark a *purpose* relation, this shows that, in the corpus at hand, *so* is mainly used to convey causal relations referring to *purpose* (in 62,3% of its occurrences). This is not surprising, since a large part of the textual genres included in the corpus have a strong directive component (e.g. user manuals). On the other hand, this also shows that, in the case of this DC, there is a tendency to translate English finite clauses by non-finite structures in Romanian. In turn, this raises the issue of the minimal units that should be taken into account as significant in a future project to annotate Romanian texts, i.e. should Romanian non-finite clauses be seen as arguments for connectives or not?

But

The polysemous nature of the English connective *but* becomes immediately apparent in its translations into Romanian. In the PDTB manual, *but* is held to mark a wide array of discourse relations, which may be roughly divided into three main categories: the class of COMPARISON, with its "Contrast" and "Concession" types, and the class of EXPANSION, with its "conjunction" subtype. All these possible meanings of *but* are present in its translations in our corpus. The category of contrast relations is represented by the Romanian conjunctions: *dar*, *însă*, and *iar*, with 95 tokens or ~70% of all occurrences (9). The concessive meaning encoded by *but* justified its being translated by *deşi* and *cu toate acestea* (10), whereas its function as a marker of *conjunction* (as described in the PDTB) seems to be rendered by the Romanian conjunctions *şi* and *ci* (11).

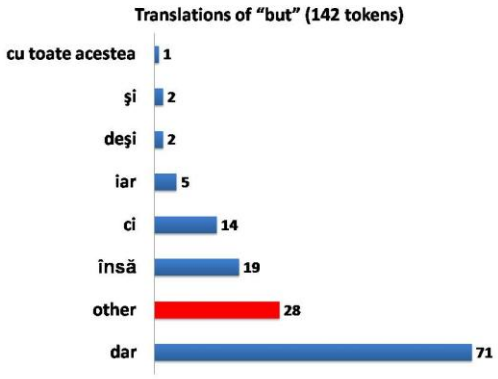
(9) *The machine prints, **but** the text is wrong, garbled, or incomplete.*
*Maşina imprimă, **dar** textul este greşit, deformat sau incomplet.*

(10) *No operators have yet announced plans to launch WiMax 2 networks, **but** the demonstration was an impressive glance into the future of mobile data.*
*Niciun operator nu a anunţat până în prezent planuri pentru lansarea de reţele WiMax 2, **deşi** demonstraţia a fost o vedere în viitorul reţelelor mobile de date wireless.*

(11) *...will not only get access to our technology for free, **but** will be champions for better broadband across Europe*
*...vor beneficia nu numai de acces gratuit la tehnologia noastră, **ci** vor fi **şi** susţinători ai îmbunătăţirii serviciilor în banda largă din întreaga Europă*

It seems that just as in the case of *but* in English, the two Romanian connectives usually described as strictly *adversative* in traditional grammars (or as markers of a relation of *contrast* according to the PDTB framework), i.e. *dar* and *însă*, may also render a *concessive* meaning, as in the following example from our corpus:

(12) *Enterprises increasingly need to achieve organization-wide compliance, **but** the end-goal often requires an unrealistic amount of time and resources.*
*Companiile au tot mai multă nevoie să obţină o conformitate cu reglementările la nivelul întregii organizaţii, **însă** scopul final necesită adesea o cantitate nerealistă de timp şi de resurse.*



In this particular example, in which the two segments seem to be linked by a concessive relation, *însă* could be just as well replaced by the Romanian *dar* or *deşi*, the latter being traditionally associated only with *concessive* relations. This shows that there is a strong link between contrast and concession relations which, while being documented for other languages, such as English (see, for instance, Izutsu, 2008), is still poorly investigated as far as the Romanian language is concerned.

It should also be noted that the Romanian conjunction *ci* could also be seen as polysemous, at least according to the PDTB framework. As a translation of the English structure *not only... but also*, i.e. in the Romanian structure *nu numai/doar... ci şi*, it seems to be the marker of a *conjunction* relation (13), whereas when used on its own it is usually a signal of *contrast* (14):

(13) *The built-in memory card reader is **not only** convenient, **but also** faster than most other forms*
*Cititorul de carduri de memorie încorporat **nu** este **doar** uşor de utilizat, **ci** este **şi** mai rapid decât majoritatea altor forme...*

(14) *To resume, do not press the direct selection keys, **but** press any other key like Ctrl.*
*Pentru a relua funcţionarea, nu apăsaţi direct tastele de selectare, **ci** apăsaţi o altă tastă, de exemplu Ctrl.*

While

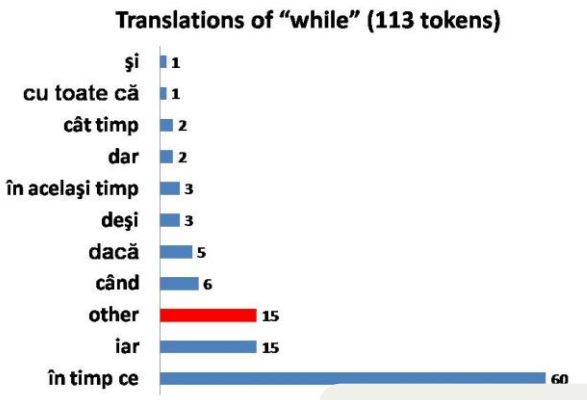
In the corpus at hand, the results for the English *while* are similar to those reported by Cartoni, Zufferey & Meyer (2013) for its French translations. The 10 lexical items used in Romanian for the source connective may be divided into three categories: 1) items that are usually used with both a temporal (15) and non-temporal meaning (16), i.e. *în timp ce*, *în acelaşi timp*, *cât timp*, 2) items with only a temporal meaning, i.e. *când* (17), and 3) items that usually render only a non-temporal meaning, i.e. *iar*, *dacă*, *deşi*, *dar*, *cu toate că*, *şi* (18).

(15) *Press and hold the key **while** clicking the trackpad.*
*Tineţi apăsată tasta **în timp ce** faceţi clic trackpad.*

(16) *Highly colored areas consist of a large number of dots, **while** lighter areas consist of a smaller number of dots.*
*Suprafeţele colorate intens sunt constituite dintr-un număr mare de puncte, **în timp ce** suprafeţele mai deschise sunt constituite dintrun număr mai mic de puncte.*

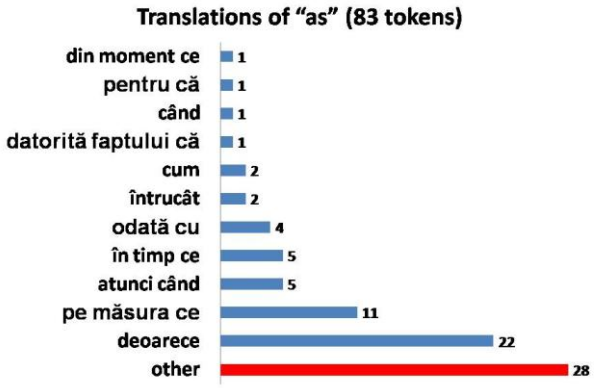
(17) *To show the menu **while** you are in an app, swipe down from the top frame onto the screen.*
*Pentru a afişa meniul **când** sunteţi într-o aplicaţie, treceţi rapid cu degetul în jos de la rama superioară până pe ecran.*

(18) ***While** the interface retains the same 'look and feel' across devices, it's tailored to the individual characteristics of each kind of device.*
***Deşi** interfaţa are "acelaşi look şi creează aceeaşi senzatie" pentru toate sistemele, este adaptată caracteristicilor individuale ale dispozitivelor.*



These three categories of Romanian connectives also correspond to the three main senses attributed to *while* in the PDTB, i.e. "TEMPORAL: Synchrony", "COMPARISON: Contrast", and

"COMPARISON: Concession", although they are not the only ones represented in the corpus. The translation-based analysis used in this study, reveals, once more, the polysemous nature of some other Romanian DCs. As shown in (15) and (16) above, within the first category, *în timp ce* may have a contrastive and/or temporal-synchrony meaning. This explains why it is the most frequent translation of *while* in the corpus, since it covers two of its basic meanings. Although it may be used to link synchronous events, *în acelaşi timp* is mainly used to render *conjunction* or *contrast* relations (19), whereas *cât timp* expresses the same combination of duration and condition as the English *as long as* or the French *tant que* (20).



(19) *...to also record programs directly to an external hard disk, **while** other content that is already saved onto USB devices can likewise be viewed...*
*...permite utilizatorilor să-şi înregistreze emisiunile preferate direct pe HDD-ul portabil, **în acelaşi timp** conţinutul deja înregistrat poate fi oricând redat...*

(20) *...you'll see the Internet Sharing icon **while** you're sharing your cellular data connection.*
*...va fi vizibilă pictograma de partajare a conexiunii la Internet **cât timp** partajaţi conexiunea celulară de date...*

Within the third category of Romanian translations for the connective *while*, *iar* and *dar* have a primarily contrastive meaning whereas *deşi* and *cu toate că* are usually used only for concessive relations (18). *Dacă* may render both a contrastive and a condition meaning of *while*, as in (21):

(21) ***While** many-core is more of a design perspective, […] it's reinventing chip design based on the assumption that high core counts is the new norm.*
***Dacă** „many-core" este mai degrabă o îmbunătăţire de design, […] multi-core reinventează designul cipurilor pornind de la ideea că noua regulă constă în importanţa numărului mare de nuclee.*

As

Just like in the case of *while*, the temporal/non-temporal dimensions of the English connective *as* are easy to distinguish in its Romanian translations. On the one hand, *deoarece*, *întrucât*, *cum*, *datorită faptului că*, *pentru că*, and *din moment ce* have no temporal meaning in Romanian, being mainly markers of the "CONTINGENCY:Cause:reason" category of meaning (22). On the other hand, *pe măsură ce*, *atunci când*, *odată cu*, and *când* are only used for temporal relations, with the first three items rendering a notion of synchrony (23). As discussed above, *în timp ce* has a twofold contrastive/temporal semantic breadth.

(22) *Do not dispose of batteries in a fire* ***as*** *they may explode.*
Nu aruncaţi bateriile în foc ***deoarece*** *pot exploda.*

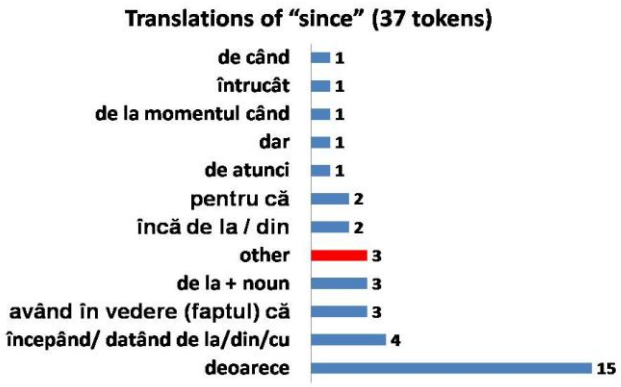
(23) *...word suggestions are displayed* ***as*** *you type.*
... ***pe măsura ce*** *tastaţi vi se afişează sugestii de cuvinte.*

It is worth noting that the instances in which *as* was translated by Romanian temporal connectives (26 cases, including *în timp ce*), non-temporal DCs (29 cases) or by other means (28 cases) are relatively evenly distributed in the corpus. Like in the case of *so*, there seems to be an important tendency to translate mainly the non/temporal, *causal* coherence relations signalled by *as* using other lexical or syntactic means (24) or simply by omission (25):

(24) *Don't miss this distinguished model in the keyboard world,* ***as*** *everything is right at your fingertips.*
Nu rataţi acest model excepţional de tastatură care vă pune toate funcţiile la îndemână.

(25) *Hold the AC power plug by the head when removing it from the wall socket,* ***as*** *pulling the lead can damage internal wires.*
*Pentru a scoate cablul din priză, apucaţi de ştecher****;*** *trăgând de cablu se pot distruge firele interioare.*

Since



The phenomenon of *disambiguationthrough translation*, which seems to be apparent in most of the examples discussed so far, is also visible in the case of the English DC *since*. The distinction between the temporal and non/temporal meanings of the source unit becomes quite clear when considering its translations from the corpus, with *deoarece*, *având în vedere (faptul că*), *pentru că*, *dar*, and *întrucât* having only (mainly causative-reason) non-temporal senses (26) and all the other items having only a temporal semantic span (27).

(26) *Do not use benzene, thinner, or rubbing alcohol* ***since*** *it may adversely affect the surface causing discoloration, etc.*
Nu folosiţi benzen, solvenţi sau alcool pentru frecare ***deoarece*** *aceste substanţe pot afecta suprafaţa producând decolorarea acesteia, etc.*

(27) ***Since*** *1992 the TCO certification program has had a significant influence on improved…*
Începând cu anul 1992 programul de certificare TCO a avut o influenţă semnificativă…

An interesting fact is that Romanian has no connective per se to express the temporal relation of succession rendered by *since* in English (or Fr. *depuis*). This explains the various alternative structures used in the corpus to translate it, mainly through reformulations: *începând/ datând de la/din/cu* (starting/ dating from/with), *de la + noun* (from + noun), *de atunci* (from

then), etc. Moreover, this raises some interesting questions about the status of *since* as a temporal connective, at least within the framework put forth by the PDTB, seeing that in many cases it takes just one chunk of discourse as its argument, thus seemingly flaunting the basic definition of DCs and acting as a preposition. Additional criteria seem to be needed in order to better differentiate between the connective and non-connective uses of *since*. However, these issues are not within the scope of this paper.

Conclusions

The small study conducted in this paper seems to confirm, at least provisionally, the two hypotheses set down initially. The translations of polysemous English DCs into Romanian may be used a "as a heuristics to uncover the meaning of Discourse Markers" (Degand, 2009), both in the source and target language. As far as the source language is concerned, the *translation spotting technique* seems to reveal straight away the polysemous or non-polysemous nature of the source DCs, since different connectives, which are usually non-interchangeable, are used to translate the various coherence relations at work in the source text. In turn, the translation-based study of DCs may also reveal interesting data about the target language, since some target language DCs are used to translate relations with which they have not been associated in traditional approaches (as in the case of *dar* and *însă* used to translate the *concession* meanings of *but*).

On the other hand, it seems that the process of translation itself results in a type of explicitation in the target language, understood here as *disambiguation throughtranslation*. In this process of explicitation, after having interpreted the meaning of the source DC in its original context, the translator chooses the target DC that seems to best convey the source coherence relation in the target language. When the target DC is clearly monosemous in the target language (i.e. *deşi* is only used to mark concessive relations in Romanian), explicitation seems to take the shape an overt, obligatory choice, which narrows down the larger range of possible interpretations of the source coherence relation to just one possible interpretation in the target text (e.g. *so* translated as *astfel încât*). However, when the target language disposes of an equally polysemous equivalent (e.g. the couple *while – în timp ce*), this kind of explicitation seems to become only optional.

Further, fine-grained analyses are needed to shed more light on these phenomena. This study represents only a first and tentative step towards a more in-depth and systematic analysis into the translation phenomena that accompany the transfer of discourse relations across languages.

Acknowledgement

This research was carried out within the COST Action IS1312, *TextLink: Structuring Discourse in Multilingual Europe*, which benefits from the joint support of the European Cooperation in Science and Technology (COST) programme and the European Science Foundation.

BIBLIOGRAPHY:

Aijmer, K., Foolen, A., & Simon-Vandenbergen, A.-M. (2006). Pragmatic markers in translation: a methodological proposal. In K. Fischer (Ed.), *Approaches to discourse particles* (pp. 101-114). Amsterdam: Elsevier.

Blum-Kulka, S. ([1986]2004). Shifts of cohesion and coherence in translation. In L. Venuti (Ed.), *The Translation Studies Reader* (S. Kitron, Trans., Taylor & Francis e-Library ed., pp. 298-313). London and New York: Routledge.

Cartoni, B., Zufferey, S., & Meyer, T. (2013). Annotating discourse connectives by looking at their translation: The translation-spotting technique. *Dialogue and Discourse*, 4 (2), 65-86.

Cuenca, M. J. (2013). The fuzzy boundaries between discourse marking and modal marking. In L. Degand, B. Cornillie, & P. Pietran (Eds.), *Discourse Markers and Modal Particles. Categorization and description* (pp. 191-216). Amsterdam / Philadelphia: John Benjamins.

Degand, L. (2009). Describing polysemous discourse markers: What does translation add to the picture? In S. Slembrouck, M. Taverniers, & M. Van Herr (Eds.), *From will to well. Studies in Linguistics offered to Anne-Marie Simon-Vandenbergen* (pp. 173-183). Gent: Academia Press.

Degand, L. (2000). Prepositional causatives or causal connectives? Discursive constraints. *Journal of Pragmatics*, 32 (6), 687-707.

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics* (31), 931-952.

Halliday, M., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Izutsu, M. N. (2008). Contrast, concessive, and corrective: Toward a comprehensive study of opposition relations. *Journal of Pragmatics* (40), 646–675.

Noël, D. (2003). Translations as evidence for semantics: an illustration. *Linguistics*, 41 (4), 757-785.

Pavel, A. N. (2013). Discoursive markers in Romanian. Terminological and conceptual aspects. In I. Boldea (Ed.), *Studies on literature, discourse and multicultural dialogue* (pp. 214-224). Tîrgu Mureş: Arhipelag XXI.

PDTB-Group. (2007). The Penn Discourse TreeBank 2.0 Annotation Manual. http://www.seas.upenn.edu/~pdtb/PDTBAPI/pdtb-annotation-manual.pdf.

Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., et al. (2008). The Penn Discourse TreeBank 2.0. *Proceedings of the 6th InternationalConference on Language Resources and Evaluation*, (pp. 2961–2968). PLACE.

Sanders, T., Spooren, W., & Noordman, L. (1992). Towards a taxonomy of coherence relations. *Discourse Processes* (15), 1-36.

Zufferey, S., & Cartoni, B. (2014). A multifactorial analysis of explicitation in translation. *Target*, 26 (3), 361-384.

Zufferey, S., & Degand, L. (2013). Annotating the meaning of discourse connectives in multilingual corpora. *Corpus Linguistics and Linguistics Theory*, 1-24.