An analysis of the Concession relation based on the discourse marker *aunque* in a Spanish-Chinese parallel corpus

La relación de Concesión analizada a través del marcador discursivo "aunque" en un corpus paralelo español-chino

Shuyuan Cao
Universitat Pompeu Fabra
C/ Roc Boronat, 138, 08018, Barcelona
shuyuan.cao@hotmail.com

Iria da Cunha
Universidad Nacional de Educación a Distancia
P/ Senda del Rey, 7, 28040, Madrid
iriad@flog.uned.es

Nuria Bel
Universitat Pompeu Fabra
C/ Roc Boronat, 138, 08018, Barcelona
nuria.bel@upf.edu

Abstract: The translation between Spanish and Chinese is particularly complicated because of the extensive grammatical, syntactic and discursive differences between the two languages. In this paper, based on the discourse marker in Spanish *aunque* ("although" in English), which usually signals the Concession relation, we will compare the discourse structure of Spanish and Chinese in the parallel corpus United Nations Corpus (UN). In order to perform the comparison, we will use the theoretical framework of Rhetorical Structure Theory (RST) by Mann and Thompson (1988).
Keywords: discourse analysis, translation, discourse marker, RST, parallel corpus

Resumen: La traducción español-chino es especialmente complicada debido a las grandes diferencias gramaticales, sintácticas y discursivas entre ambas lenguas. En este trabajo, comparamos las estructuras discursivas del español y el chino en el corpus paralelo *United Nations Corpus* (UN), partiendo del marcador discursivo en español *aunque*, que señala la relación de Concesión. Para realizar la comparación empleamos el marco teórico de la *Rhetorical Structure Theory* (RST) de Mann y Thompson (1988).
Palabras clave: análisis del discurso, traducción, marcador discursivo, RST, corpus paralelo

1 Introduction

The greater the linguistic distance is between a pair of languages, the greater the number of differences in their syntax and discourse structure. Therefore, the translation between two very different languages can be potentially more difficult. Comparative or contrastive studies of discourse structures offer clues to identify properly equivalent discourse elements in two languages. These clues can be useful for both human and machine translation (MT).

The emphasis on the idea that discourse information may be useful for Natural Language Processing (NLP) has become increasingly popular. Discourse analysis is an unsolved problem in this field, although discourse information is crucial for many NLP tasks (Zhou et al., 2014). In particular, the relation between MT and discourse analysis has only recently begun and works addressing this topic remain limited. A shortcoming of most of the existing systems is that discourse level is not considered in the translation, which therefore affects translation quality (Mayor et

al., 2009; Wilks, 2009). Notwithstanding, some recent researches indicate that discourse structure improves MT evaluation (Fomicheva, da Cunha and Sierra, 2012; Tu, Zhou and Zong, 2013; Guzmán et al., 2014).

Studies that use Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) contribute to discourse analysis research. RST is a theory that describes text discourse structure in terms of Elementary Discourse Units (EDUs) (Marcu, 2000) and cohrence relations. Relations are recursive in RST and are defined to hold between EDUs; these EDUs can be Nuclei or Satellites, denoted by N and S (Satellites offer additional information about Nuclei). The relations can be Nucleus-Satellite (e.g. Cause, Result, Concession, Antithesis) or Multinuclear (e.g. List, Contrast, Sequence).

RST is appropriate for our research because of the following reasons:

(1) RST is one of the most acknowledged and representative discourse structure theories. It shows a good applicability in a cross-cultural comparative study. As Mann and Thompson (1988: 247) declare:

RST is an abstract set of conventions. We can view the conventions as either independent or inclusive of particular relation definitions. The first view is more comprehensive, but the latter is more convenient - we use the latter. The first view would be essential for a cross-linguistic or cross-cultural comparative study in which relation definitions might differ.

Our research is also a comparative study of a parallel Spanish-Chinese corpus.

(2) The set of relations proposed in the framework of RST is representative for the analysis of the discourse of any field or text genre (Mann and Thompson, 1988). The investigation targets of our work are Spanish and Chinese written texts that contain many different subjects or themes.

(3) In a contrastive study, the RST elements (segments, relations and discourse structure) can reveal how the discourse elements are being expressed formally in each language. This could reflect the similarities and differences of the rhetorical structure of the language pair. In turn, this may help us to elaborate some guidelines that include discourse similarities and differences for human and machine translation (MT) between the language pair Chinese-Spanish.

Discourse information is necessary for a good translation quality. Based on the Spanish discourse marker *aunque*, the following two examples show discourse similarities and differences between Spanish and Chinese.

(1):
(1.1) Sp: **Aunque** está enfermo, va a trabajar.

[**Aunque** está enfermo,]EDU $_S$ [va a trabajar.]EDU $_N$

(**Marker_1** is ill, goes to work.) [1]

(1.2) Ch: **虽然**他病了，**但是**他去上班了。
[**虽然**他病了，[2]]EDU $_S$ [**但是**他去上了。]EDU $_N$
(**Marker_1** he ill, **marker_2** he goes to work.)
(1.3) En: **Although** he is very ill, he goes to work. [3]

In example (1), Spanish and Chinese passages show the same rhetorical relation (Concession), and the order of the nucleus and the satellite is also similar. However, in Chinese, it is mandatory to include two discourse markers to show this relation: one marker "*suiran*" (虽然) at the beginning of the satellite and another marker "*danshi*" (但是) at the beginning of the nucleus. These two discourse markers are equivalent to the English discourse marker *although*. By contrast, in Spanish, to show the Concession relation, only one discourse marker is used at the beginning of the satellite (in this case, *aunque* 'although').

(2):
(2.1.1) Sp: Hace frío, aunque no llueve.

[Hace frío,]EDU $_N$ [**aunque** no llueve.]EDU $_S$

(Makes cold, **marker_1** no rain.)

(2.1.2) Sp: Aunque no llueve, hace frío.

[**Aunque** no llueve,]EDU $_S$ [hace frío.]EDU $_N$

(**marker_1** no rain, has cold.)

---

[1] In this work, we offer an English literal translation in brackets for the first two examples in order to understand each example better.
[2] All the Chinese characters and punctuations occupy two positions in a written text; therefore, the readers can see a blank space between the punctuation and bracket in the examples.
[3] In this work, for all the examples we give, all the English translations are translated from the Spanish sentences by authors.

(2.2) Ch: 很冷，虽然没有下雨。

[很冷，]EDU_N [**虽然**没有下雨。]EDU_S

(It′s cold, **marker_1** there is no rain.)

(2.3) En: It is cold, **although** there is no rain.

In example (2), the Chinese passage could have the same or the different rhetorical structure when comparing to the Spanish passages. In the Chinese passage, the discourse marker "*suiran*" (虽然) at the beginning of satellite, which is equivalent to the English discourse marker *although*, shows a Concession relation, and the order between nucleus and the satellite cannot be changed. In the Spanish passage, "aunque" is also at the beginning of satellite, which also corresponds to the English discourse marker *although*, and shows the same discourse relation, but the order between nucleus and satellite can be changed and this does not change the sense of the sentence.

From the two examples above, we can see that, in order to express a Concession relation in a written text, the Spanish discourse marker *aunque* can be translated into different Chinese discourse markers, but without relevant differences in the Chinese text coherence.

In this work, by using the Spanish discourse marker *aunque* ('although' in English), which shows the Concession relation; we will compare the discourse structure of Spanish and Chinese in the parallel corpus United Nations Corpus (UN) (Rafalovitch and Dale, 2009).

In Section 2, we will introduce some related works that use RST. In Section 3, we will give detailed information of the methodology. In Section 4, we will establish the results. In Section 5, we will conclude the research information and look ahead at future work.

2   Related Work

Thus far there have not been many studies addressing discourse analysis relating to our subject of study. Yet some comparative studies between Chinese and English by employing RST exist. Cui (1986) presents some aspects regarding discourse relations between Chinese and English; Kong (1998) compares Chinese and English business letters; Guy (2000, 2001) compares Chinese and English journalistic news texts.

Other pairs of language within RST include Japanese and Spanish (Kumpf, 1986; Marcu et al., 2000), Arabic and English (Mohamed and Omer, 1999), French and English (Delin,

Hartley and Scott, 1996; Salkie and Oates, 1999), Dutch and English (Abelen, Gisla and Thompson, 1993), Finnish and English (Sarjala, 1994), Spanish and Basque (da Cunha and Iruskieta, 2010).

There are few contrastive works between Spanish and Chinese. None of them use RST. Yao (2008) uses film dialogues to elaborate an annotated corpus, and compares the Chinese and Spanish discourse markers in order to give some suggestions for teaching and learning Spanish and Chinese. In this work, Yao does not use any framework that based on discourse analysis; he just analyses and compares Spanish and Chinese discourse markers' characteristics and then makes conclusions. Taking different newspapers and books as the research corpus, Chien (2012) compares the Spanish and Chinese conditional discourse markers to give some conclusions on the conditional discourse marker for foreign language teaching between Spanish and Chinese. Wang (2013) uses a corpus of films to analyse how the subtitled Spanish discourse markers can be translated into Chinese, so as to make a guideline for human translations and audiovisual translation between the language pair.

The RST contrastive studies that use more than two languages are not common, for example, Portuguese-French-English (Scott, Delin and Hartley, 1998). In this work, a methodology has been presented for RST contrastive analysis while the empirical cross-lingual results have been published. Iruskieta, da Cunha and Taboada (2015) use RST as theoretical framework to compare Basque, Spanish and English, so as to create a new qualitative method for the comparison of rhetorical structures in different languages and to specify why the rhetorical structure may be different in translated texts.

3   Methodology

As the previous examples show, discourse similarities and differences exist between the Spanish sentences that contain the discourse marker *aunque* and their Chinese translated sentences. For this study, we have adopted the UN corpus as the research corpus. This corpus contains all 6 languages (English, Chinese, Spanish, Arabic, Russian and French) of the UN, consisting of around 300 million words for each language. Recorded in March of 2010, this corpus consists of 463,406 documents,

80,931,645 sentences in total. 326 million words have been calculated as the average number for five of the six official languages.

Table 1 shows the detail information of the UN corpus, and its subcorpus in Spanish and in Chinese.

Name	UN corpus	Spanish subcorpus	Chinese subcorpus
Nº of documents	463,406	70,509	65,022
Nº of sentences	80,931,645	13,052,875	10,839,473
Nº of words	326 million for each language	352,460,926	756,108,566
Nº of Sp-Ch parallel documents	/	62,738	

Table 1: Statistics of the Spanish and Chinese UN corpora

In this research, we have analysed 4 million Spanish words and its parallel Chinese texts, as corpus to study the marker *aunque*. We have extracted all the Spanish sentences (including repeated sentences) that contain the discourse marker *aunque* and all their Chinese parallel sentences manually. Then, we have carried out the RST analysis of these sentences manually by using RSTTool (O'Donnell, 1997). See for example Figure 1.
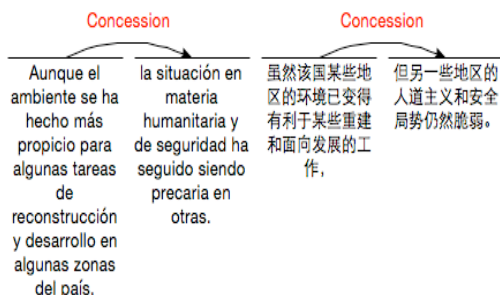

Figure 1: Example of RST analysis of a Spanish and Chinese sentence from the corpus with RSTTool[4]

---

[4] English translation of the sentence in Figure 1: Although the environment has become more conducive to some reconstruction and development in some areas of the country, the humanitarian and security situation has remained precarious in others.

Finally, we have compared the Spanish sentences and the parallel Chinese sentences through: discourse segmentation, RST relations, nuclearity order and discourse marker.

In this work, we followed the segmentation criteria proposed by Iruskieta, da Cunha and Taboada (2015) for both Spanish and Chinese. We use the relations established by Mann and Thompson (1988) and the relation of Same-unit mentioned by Carlson, Marcu and Okurowski (2001) to define the relations between the EDUs. We depart from discourse markers to detect the different discourse relations between EDUs and to decide the EDUs to be the nucleus or the satellites, and the nuclearity order.

Among the 4 million Spanish words, we have obtained 99 Spanish sentences that contain the discourse marker *aunque*. However, not all the 99 sentences are different: we find 13 repeated sentences. Therefore, 86 sentences are included in the results. Table 2 includes an example of our database.

	Spanish	Chinese
Sentence with discourse marker in bold	Observando con satisfacción que, **aunque** queda pendiente una labor considerable, las partes interesadas han hecho avances reales en el logro del objetivo de la ordenación sostenible de la pesca.	满意地注意到**虽然**仍有大量工作要做，**但**有关各方已朝可持续渔业管理方向取得了实际进展。
Discourse relation	Concession	Concession
Nucleus-Satellite order	S-N	S-N

Table 2: Example of the parallel sentences included in the analysis[5]

4 Results

As previously mentioned, the discourse marker *aunque* in Spanish shows the Concession relation. In its parallel Chinese corpus, *aunque* presents the following translation cases. There are 47 independent Chinese sentences that hold the Concession relation. For showing the

---

[5] English translation of the sentence in Table 2: Realizing satisfied that, although considerable work remains pended, the interesting parts have made real progress towards the goal of sustainable fisheries.

Concession in its parallel Chinese corpus, *aunque* has been translated into 6 different Chinese discourse markers, which are formed by two words, as "*suiran... danshi*" (虽然… 但), "*suiran... dan*" (虽然… 但), "*jinguan... danshi*" (尽管... 但是), "*jishi... reng*" (即使... 仍), "*suiran... que*" (虽然... 却), "*sui... er*" (虽... 而); and other 2 different Chinese discourse markers with just one word, which are "*jishi*" (即使) and "*jinguan*" (尽管).

For the case of the two discourse markers in Chinese, they have the same meaning, and are equivalent to the English discourse marker 'although…but'. In this case, it is mandatory to include two discourse markers to show the Concession relation: one marker at the beginning of the satellite and another marker at the beginning of the nucleus. In the case of two discourse markers, the rhetorical structure of Chinese sentences is S-N. For the case of one discourse marker, these discourse markers are equivalent to English discourse 'although', and in the Concession relation, it is necessary to put them at the beginning of the satellite, the rhetorical structure is N-S.

In Spanish, for expressing a Concession relation, the order of nucleus and satellite can change and it does make sense syntactically. Because of the changeable order of nucleus and satellite in Spanish, in UN corpus, for the Concession relation case, the rhetorical structure between Spanish and Chinese is similar; it depends on how many Chinese discourse markers have been used in the translation.

Apart from indicating the Concession relation, there are some other special translations to Chinese of the sentences including *aunque* in Spanish:

1) There are 22 Chinese sentences that comprise the Antithesis relation. In Antithesis relation, *aunque* has been translated to "*dan*" (但), "*sui*" (虽), and "*er*" (而). All these Chinese discourse markers carry the same meaning and are equivalent to 'but' in English. In this special case, for each pair of the parallel sentences, the number of discourse markers between Spanish and Chinese is the same. The rhetorical structure of these 22 Chinese sentences is N-S.

Though *aunque* has been translated into different Chinese discourse markers and these Chinese discourse markers show an Antithesis relation, which is different from the Concession

relation. In RST, an Antithesis relation means the situations presented in nucleus and satellite are in contrast, while in a Concession relation the situations presented in both EDUs are compatible (Mann and Thompson, 1988). We consider that the changed relation during the translation does not affect readers to understand the information of context. Here we give an example in the UN corpus to explain the situation.

(3):
Sp: [El objetivo es alentar o servir de inspiración a los ciudadanos para que presten servicios voluntarios,]EDU_N [**aunque** la decisión queda en manos de la persona o la organización.]EDU_S

Ch: [鼓励或激励公民志愿服务，]EDU_N [但让个人或组织自己做出选择。]EDU_S

En: The goal is to encourage or inspire citizens to volunteer, although the decision is in the hands of the person or organization.

In this example we can see that, the Spanish passage holds a N-S rhetorical structure of the Concession relation while the Chinese passage holds the same rhetorical structure in the Antithesis relation. Merely, the main idea of these two passages is the same, which is to offer services voluntarily and let the person or the organization to choose by their own.

2) The translation of *aunque* represents another RST relation in 15 Chinese sentences, which is a multinuclear relation (N-N) known as List.

There are few occasions in the UN corpus where *aunque* has been translated into "*tongshi*" (同时), which in Chinese means *at the same time*. The selected example of the translated Chinese discourse marker "*tongshi*" (同时) in the UN corpus is the following:

(4):
Sp: [Acoge complacida el progreso logrado en la rehabilitación de escuelas, el suministro y la distribución de material didáctico y la capacitación de maestros,]EDU_N [**aunque** subraya la necesidad de fomentar la capacidad.]EDU_S

Ch: [欢迎学校在修复、教材供应和分配以及教师培训方面取得进展，]EDU_N [**同时**强调需要进行能力建设。]EDU_N

En: Welcomes the progress made in the rehabilitation of schools, provision and distribution of educational materials and teacher training, **while** emphasising the need for capacity building.

In example (4) we can see that the Spanish passage uses *aunque* to show a Concession relation but the Chinese passage uses "*tongshi*" (同时) ('meanwhile' in English), a multinuclear relation (List) to deliver the information. In the Spanish passage, the highlighting part is the first EDU; the second EDU is the additional information of the first EDU. In the Chinese passage, both EDUs are same important. Though the rhetorical structures (discourse relations and the nuclearity order) between the two passages are different, they all show the same basic information. This shows that, though there are improvements in schools, still the need for capacity building should be emphasised.

3) There are 2 translated Chinese sentences that do not contain the translation of *aunque*. Example (5) shows one of these cases.

(5):

Sp: [Reconoce que, **aunque** las medidas adoptadas para aplicar los resultados de las grandes cumbres y conferencias y los períodos extraordinarios de sesiones de las Naciones Unidas en las esferas económica y social y esferas conexas celebrados durante los últimos diez años servirán para promover el desarrollo social]EDU_S [**también** será necesario contar con una cooperación y una asistencia para el desarrollo más intensas y eficaces en los planos internacional y regional y avanzar hacia una mayor participación, justicia social y equidad en las sociedades]EDU_N

Ch: [确认为执行过去十年间在经济、社会和有关领域举行的联合国各次主要的首脑会议、会议和特别会议的成果而采取的行动将进一步促进社会发展]EDU_N [但也必须加强和有效开展国际和区域合作与发展援助,逐步扩大参与,加强社会正义和增进社会公平。]EDU_S

En: Recognizes that, although the measures taken to implement the outcomes of the major summits and conferences and special sessions of the United Nations in the economic, social and related fields held during the past ten years will further promote social development, also it is necessary to depend on the cooperation and assistance for more strengthened and effective development in the international level and regional, and move towards to a greater participation, social justice and equality in societies.

In the Spanish passage, the discourse marker *aunque* shows a Concession relation. The marker *también* ('also' in English) is included in the sentence too. The Chinese passage just translates the discourse marker "*también*" as "*dan*" (但). Although the Spanish passage and the Chinese passage both hold a nucleus-satellite (N-S) relation, the rhetorical relation is different. A Condition relation (S-N) is held between two Spanish sentences while the Chinese parallel sentences have an Antithesis relation. This means that in Spanish the emphasised part (nuclear span of relation) is the second EDU, whereas in Chinese the opposite occurs.

Table 3 includes the discourse structures in Chinese detected in our corpus equivalent to the sentences in Spanish including the discourse marker *aunque* (that is, showing a Concession relation). This table could be used by Spanish-Chinese human translators and could be useful for MT researchers. When translating the Spanish discourse marker *aunque* to Chinese, for showing a Concession relation, they could follow the rules included in Table 3.

Nuclearity order (N-S/ S-N / N-N)	Disc. markers	Position of disc. marker (N/S)
S-N	suiran..danshi (虽然... 但是)	N&S
S-N	suiran..dan (虽然... 但)	N&S
S-N	jinguan….danshi (尽管... 但是)	N&S
S-N	jishi..reng (即使... 仍)	N&S
S-N	suiran..que (虽然... 却)	N&S
S-N	sui…er (虽... 而)	N&S
N-S	jishi (即使)	S
N-S	jinguan (尽管)	S

Table 3: Chinese discourse structures equivalent to Spanish discourse structures including the discourse marker *aunque* and Concession relation

5 Conclusion and Future Work

In this work, we have explored the sentences that contain the Spanish discourse marker *aunque* and their Chinese parallel sentences in the UN subcorpus. In the Spanish subcorpus, *aunque* shows the Concession relation. However, in the Chinese subcorpus, this marker has many different Chinese discourse markers,

and these Chinese discourse markers hold different RST relations. Besides, in some parallel sentences, there is no translation of *aunque*.

The original language of the official documents in the UN corpus is English. The parallel corpus is translated from English, so the Spanish-Chinese parallel corpus is actually made up of two parts. One is the translation between English and Spanish, and the other is the translation between English and Chinese. These translated Spanish and Chinese documents make up the UN Spanish-Chinese parallel corpus. The UN parallel Spanish-Chinese corpus is not a direct translation corpus. Therefore, due to the linguistic realization (normally known as translation strategy) a Spanish discourse marker could be translated to different discourse markers in its parallel Chinese corpus. Also for a same sentence, nuclearity order and the number of discourse markers between these two languages could be different. In the 86 analysed sentences, the rhetorical structure between Spanish and Chinese is quite similar. This means that the rhetorical structure has been changed when doing the translation work. We think this explains why the Spanish discourse marker *aunque* has been translated to different Chinese discourse markers and why it has not been translated in only a few instances.

In this work we have only analysed the structure of independent sentences, only intra-sentence discourse elements have been analysed, and the analysis does not bring us many discourse differences between Spanish and Chinese. However, we expect to find more discourse differences when analysing a whole text.

This research is a corpus-based preliminary study. For our future work, we will use a larger Spanish-Chinese parallel corpus and compare their nucleus-satellite order to find more discourse similarities and differences in order to provide discourse information for the translation between this language pair.

6 Acknowledgements

References

Abelen, E., G. Redeker, and S. A. Thompson. 1993. The rhetorical structure of US-American and Dutch fund-raising letters. *Text* 13(3): 323-350.

Carlson, L., D. Marcu, and M. E. Okurowski. 2001. Building a discourse-tagged corpus in the framework of Rhetorical Structure Theory. In *Proceedings of the Second SIGdial Workshop on Discourse and Dialogue*, pp, 1-10. Aalborg (Denmark), 1-2 September.

Cui, S. R. 1985. *Comparing Structures of Essays in Chinese and English*. Master thesis. Los Angeles: University of California.

Chien, Y. S. 2012. *Análisis contrastivo de los marcadores condicionales del español y del chino*. PhD thesis. Salamanca: Universidad de Salamanca.

da Cunha, I., and M. Iruskieta. 2010. Comparing rhetorical structures of different languages: The influence of translation strategies. *Discourse Studies* 12(5): 563-598.

Delin, J., A. Hartley, and D. R. Scott. 1996. Towards a contrastive pragmatics: Syntatic choice in English and French instructions. *Language Sciences* 18(3-4): 897-931.

Fomicheva, M., I. da Cunha, and G. Sierra. 2012. La estructura discursiva como criterio de evaluación de traducciones automáticas: una primera aproximación. *Empiricism and analytical tools for 21 century applied linguistics: selected papers from the XXIX International Conference of the Spanish Association of Applied Linguistics (AESLA)*: 973-986.

Guy, R. 2000. Linearity in Rhetorical Organisation: A Comparative Cross-cultural Analysis of Newstext from the People's Republic of China and Australia. *International Journal of Applied Linguistics* 10(2): 241-58.

Guy, R. 2001. What Are They Getting At? Placement of Important Ideas in Chinese Newstext: A Contrastive Analysis with Australian Newstext. *Australian Review of Applied Linguistics* 24(2): 17-34.

Guzmán, F., S. Joty, Ll. Màrquez, and P. Nakov. 2014. Using Discourse Structure Improves Machine Translation Evaluation. In

Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, pp, 687-698. Baltimore (USA), 22-27 June.

Iruskieta, M., I. da Cunha, and M. Taboada. 2015. A Qualitative Comparison Method for Rhetorical Structures: Identifying different discourse structures in multilingual corpora. *Language resources and evaluation* 49(2): 263-309.

Kong, K. C. C. 1998. Are simple business request letters really simple? A comparison of Chinese and English business request letters. *Text* 18(1): 103-141.

Kumpf, L. 1975. *Structuring Narratives in a Second Language: A description of Rhetoric and Grammar*. PhD thesis. Los Angeles: University of California.

Mann, W. C., and S. A. Thompson. 1988. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text* 8(3): 243-281.

Marcu, D. 2000. The rhetorical parsing of unrestricted texts: A surface-based approach. *Computational Linguistics* 26(3): 395-448.

Marcu, D., L. Carlson, and M. Watanabe. 2000. The automatic translation of discourse structures. In *1st North American chapter of the Association for Computational Linguistics conferences*, pp, 9-17. Seattle (USA), 29-4 April to May.

Mayor, A., I. Alegria, A. Díaz de Ilarraza, G. Labaka, M. Lersundi, and K. Sarasola. 2009. Evaluación de un sistema de traducción automática basado en reglas o porqué BLEU sólo sirve para lo que sirve. *Procesamiento del Lenguaje Natural* 43: 197-205.

Mohamed, A. H., and M. R. Omer. 1999. Syntax as a marker of rhetorical organization in written texts: Arabic and English. *International Review of Applied Linguistics in Language Teaching (IRAL)* 37(4): 291-305.

O'Donnell, M. 1997. RST-tool: An RST analysis tool. In *Proceedings of the 6th European Work-shop on Natural Language Generation*, pp, 92-96. Duisberg (Germany), 24-26 March.

Rafalovitch, A., and R. Dale. 2009. United Nations general assembly resolutions: A six-languages parallel corpus. In *Proceedings of MT Summit XII*, pp, 292-299. Ottawa (Canada), 26-30 August.

Salkie, R., and S. L. Oates. 1999. Contrast and concession in French and English. *Languages in Contrast* 2(1): 27-56.

Sarjala, M. 1994. Signalling of reason and cause relations in academic discourse. *Anglicana Turkuensia* 13: 89-98.

Scott, R., J. Delin, and A. F. Hartley. 1998. Identifying congruent pragmatic relations in procedural texts. *Languages in contrast* 1(1): 45-82.

Tu, M., Y. Zhou, and C. Q. Zong. 2013. A Novel Translation Framework Based on Rhetorical Structure Theory. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pp, 370-374. Sofia (Bulgaria), 4-9 August.

Wilks, Y. 2009. *Machine Translation: Its scope and limits*. 3ª ed. New York: Springer.

Wang, Y. C. 2013. *Los marcadores conversacionales en el subtitulado del español al chino: análisis de La mala educación y Volver de Pedro Almodóvar*. PhD thesis. Barcelona: Universitat Autònoma de Barcelona.

Yao, J. M. 2008. *Estudio comparativo de los marcadores del discurso en español y en chino a través de diálogos cinematográficos*. PhD thesis. Valladolid: Universidad de Valladolid.

Yue, M. 2006. *Hanyu caijingpinglun de xiucijiegou biaozhu ji pianzhang yanjiu* (汉语财经评论的修辞结构标注及篇章研究,[Annotation and Analysis of Chinese Financial News Commentaries in terms of Rhetorical Structure]). PhD thesis. Beijing: Communication University of China.

Zhou, L. J., B. Y. Li, Z. Y. Wei, and K. F. Wong. 2014. The CUHK Discourse TreeBank for Chinese: Annotating Explicit Discourse Connectives for the Chinese TreeBank. In *Proceedings of the International Conference on Language Resources and Evaluation*, pp, 942-949. Reykjavik (Iceland), 26-31 May.