

Conjunctive cohesion in arabic united nations texts and their english translations: a corpus-based study

Chaalal Imen

A Doctorate Student in Applied Linguistics and Applied Language Studies
Department of Letters and the English Language
University of Mentouri Constantine 1
chaalalimen@yahoo.co.uk

Cette recherche qui est basée sur un corpus parallèle, étudie l'utilisation des conjonctions arabe et anglais dans certains textes des Nations Unies afin de découvrir la manière avec laquelle ils sont interprétés comme équivalents et de savoir comment les traducteurs les ajoutent ou les omettent dans leurs traductions. En utilisant deux logiciels de corpus, ces conjonctions seront comparées. Les résultats révèlent que la distribution des conjonctions en arabes et anglais est la même. La seule différence apparaît dans la fréquence de leur utilisation. Il semble que les textes arabes sont plus explicitement cohérents que leurs équivalents anglais par l'utilisation des conjonctions qui marquent l'addition, la cause et le temps.

This corpus-based study investigates the use of Arabic and English conjunctions in some United Nations texts in order to find out how they are interpreted as equivalents and to know how translators add or omit them in their translations. Using computer software, the use of conjunctions is compared. Findings reveal that the distribution of conjunctions in Arabic and English is the same. The only difference lies in the frequency of occurrence of those relations. It seems that Arabic texts are more explicitly cohesive than their English counterparts through the use of additives, causals and temporals.

Introduction

In the last few decades, the development of computer technology and the emergence of corpus linguistics tools and methods have brought many advantages to the disciplines of contrastive and translation studies. A great number of research works of textual cohesion have been suggested with respect to the applications of corpora in translation classrooms. The works of Bystrova-McIntyre (2012) and Giannossa (2012) are two examples of the genuine development in the study of cohesion, when the incorporation of corpus-based cross-linguistic research was considered as a mainstream methodology used within the translation framework.

It is important to note that although corpus-based studies have been applied in contrastive and translation studies dealing with specialised language/translation (e.g. Zhao et al.'s corpus-based study of cohesion in English medical texts and its Chinese translation (2009)), very little research has been undertaken in legal translation.

Cohesion, including conjunction, which is viewed as a genre-specific and text-type specific phenomenon has been slightly investigated for legal purposes. For example, the work of Trebits (2009) demonstrates the significance of conjunctive cohesion, as one linguistic aspect of text unity, in the monolingual framework (English). She describes the use of conjunctions in the EU documents with the help of corpus linguistic techniques and reveals some of the textual organization patterns the documents show. The study suggests also some implications for the teaching of conjunctions in EFL classrooms when teaching English for EU purposes (as one type of legal specific purposes). However, the attempt to study these devices in legal discourse across languages have been slightly neglected; only very few research works tackled this issue recently; e.g., Hanting's translation of conjunctive cohesion in legal documents (2013) and Ketabi and Jamalvand's corpus-based study of conjunction devices in English international law texts and its Farsi translation (2012).

In this paper, we attempt to undergo a corpus translation study of conjunctive devices across two languages, Arabic and English, in a specific genre, legal translation, particularly the United Nations texts. The aim is to uncover the differences and similarities between these devices across the two languages in terms of the categories used and their occurrence frequencies.

In this view, based on Halliday and Hasan's model of cohesion (1976), the study attempts to observe variation in the way each language makes use of conjunctions to signal cohesion relations and to identify the shift of these devices, at the textual level, in the English translations of some Arabic United Nations texts. The questions asked are whether the stylistics differences between Arabic and English exhibit different types of conjunctions in translation, and whether there are any significant differences in the frequency of occurrence of these devices across the two languages. Therefore, by adopting the notion of translation equivalence as the best basis for contrastive analysis, the study aims basically to examine the behaviour of conjunctions across the two languages and note to what extent they are interpreted as equivalents, and to raise the student's awareness of the vital role that these devices play in translation.

1. Methodology

Based on the assumption that the contribution of corpus-based studies in contrastive studies may be used for practical purposes in translation, and based

on the significant role of parallel corpora in the examination of the differences and similarities between languages, this research makes use of a unidirectional parallel corpus of Arabic and English United Nations texts (PCUNTs).

The PCUNTs is a self-built Arabic/ English corpus consisting of 40 UN documents. It is a specialized one, since it focusses on one specific genre: legal translation. It consists of 9 GA Resolutions and 11 SC Resolutions published over a period of three years (2011-2013), and related to the most relevant events in the Middle East and North Africa, tackling mainly issues related to the crises in Syria, Iraq and Libya. The PCUNTs consists of two sub-corpora, 20 Arabic source texts (AUNTs) and their English translations (EUNTs). The whole corpus contains around 31,045 word tokens; the Arabic sub-corpus has fewer word tokens (13,753) in comparison to the English one, 17,292 word tokens. Because conjunctive cohesion tends to be more frequent in the corpus, it might be said that such a relatively small corpus is sufficient and seems to be adequate for the purpose of analysis.

It is worth noting that the texts are selected according to their reputation and availability. They are available online and downloadable via the official website of the UN Documentation¹(The ODS).

2. Tools of Research

For the purpose of analysis, two main software tools developed by Anthony (2011 and 2013) are used. First, the AntConc computer software (Anthony, 2011) is used; the word list function of the AntConc allows us to create the list of the most frequent conjunctive devices in the parallel corpus (PCUNTs), and to compare the two sub-corpora in terms of the types of conjunctions used. Second, the AntPconc computer software (Anthony, 2013) is used in order to examine the differences and similarities between Arabic and English devices in real context. Then, the Software Package for Social Sciences (SPSS 22) is used for the statistical analysis and the comparison of the two sub-corpora; it is employed to examine whether or not significant differences exist between the two sub-corpora in terms of the use of conjunctive devices.

3. Results and Discussion

The results of conjunctive cohesion analysis are classified into two types. The first one deals with the results of the semantic analysis of cohesion, in which the frequency of occurrence and percentages of these devices are recorded and

¹ <http://documents.un.org/>

presented. The second one reports the results of the statistical analysis of the corpus. By using SPSS, the differences in the averages are subjected to a paired t-test in order to find out whether or not there are statistically significant in the use of conjunctions across the two sub-corpora.

Results of the Semantic Analysis of Conjunctions in the PCUNTs

The results of conjunctions in the UN texts are presented in Table 1. The types of conjunctive relations, occurring at both intra-sentential and inter-sentential levels across the two languages, are additives, adversatives, causals, temporals and continuatives. The overall frequency and distribution of the identified conjunctions across the two sub-corpora show a total of **895** of the five types of conjunctive devices in AUNTs, and a total of **596** in EUNTS.

| Conjunction | AUNTs | | EUNTs | |
|-------------|-------------------|------------|-------------------|------------|
| | N° of Occurrences | Percentage | N° of Occurrences | Percentage |
| Text 1 | 33 | 6.35 | 21 | 3.61 |
| Text 2 | 26 | 6.34 | 14 | 2.95 |
| Text 3 | 27 | 6.80 | 17 | 3.35 |
| Text 4 | 13 | 8.17 | 11 | 4.36 |
| Text 5 | 52 | 6.22 | 33 | 3.36 |
| Text 6 | 53 | 6.68 | 39 | 4.19 |
| Text 7 | 25 | 7.26 | 24 | 5.21 |
| Text 8 | 66 | 6.58 | 46 | 4.04 |
| Text 9 | 28 | 5.27 | 18 | 3.13 |
| Text 10 | 53 | 5.57 | 49 | 3.92 |
| Text 11 | 26 | 6.35 | 16 | 3.36 |
| Text 12 | 23 | 4.44 | 14 | 2.44 |
| Text 13 | 31 | 8.15 | 16 | 3.52 |
| Text 14 | 91 | 5.59 | 60 | 2.99 |
| Text 15 | 25 | 3.74 | 21 | 2.70 |
| Text 16 | 109 | 5.86 | 69 | 3.17 |
| Text 17 | 23 | 4.97 | 16 | 3.10 |
| Text 18 | 98 | 7.72 | 41 | 3.28 |
| Text 19 | 48 | 5.11 | 34 | 3.21 |
| Text 20 | 45 | 6.40 | 37 | 4.22 |

Table 1: Conjunctive Cohesive Devices in the PCUNTs

The word frequency analysis reveals that conjunctions are comparatively higher in use in the Arabic STs than in the English TTs. Thus, it can be said that there is a major difference in the use of this category between the two sub-corpora. Conjunctions are relatively more frequent in Arabic than in English.

Subcategories of Conjunctive Devices in the PCUNTs

What follows in Table 2 is the result of the analysis of the subcategories of conjunctive devices in the PCUNT. These subcategories are counted and then

calculated as a percentage in each pair of texts. The average percentage is also calculated in order to reveal to what extent the corpus depends on these particular types.

| Corpus Conjunctive Devices | AUN Texts | | EUN Texts | |
|----------------------------------|----------------------|------------|----------------------|------------|
| | N° of Occurrences | Percentage | N° of Occurrences | Percentage |
| Additive | 585 | 65.36 | 303 | 50.83 |
| Adversative | 27 | 03.01 | 21 | 03.52 |
| Causal | 136 | 15.19 | 102 | 17.11 |
| Temporal | 96 | 10.72 | 83 | 13.92 |
| Others | 52 | 05.81 | 89 | 14.93 |
| Total | 895 | 100% | 596 | 100% |

Table 2: Frequencies and Percentages of Subcategories of Conjunction

A glance at the table reveals that additives (65.36% in AUNTs and 50.83% in EUNTs) are the most frequent devices of conjunctions, followed by causals (15.19% in AUNTs and 17.11% in EUNTs) and temporals (10.72% in AUNTs and 13.92% in EUNTs). Adversatives, however, are found to be the least frequently used in the PCUNTs (03.01% in AUNTs and 03.52% in EUNTs). Other types of conjunctions, i.e. continuative conjunctions are also of low frequencies (05.81% in AUNTs and 14.93% in EUNTs); they rank fourth and third, respectively, in the two sub-corpora.

a) Additive Conjunctions in the PCUNTs

The overall percentages of additive conjunctions in the PCUNTs constitute half the overall occurrences of conjunctive devices (65.36% in AUNTs and 50.83% in EUNTs). In Table 3 below, the figures show that all the Arabic STs exhibit higher use of additives than English TTs do. Therefore, it can be said that additives seem to be more frequent in Arabic than in English.

| Additive Conjunctions | AUNTs | EUNTs |
|-----------------------|-------|-------|
| Total of Occurrences | 585 | 303 |
| Average Percentage | 3.90 | 1.66 |

Table 3: Additive Conjunctions Frequency and Average Percentage in the PCUNTs

b) Causal Conjunctions in the PCUNTs

Causal conjunctions are the second most widely used conjunctive devices in the parallel corpus. The comparison between the frequencies of occurrence of causals in the parallel corpus shows that the two sub-corpora display slight differences in terms of the frequency of occurrence of causals. The results in Table 4 reveal that the use of causals in Arabic STs is higher than in the English

TTs. In addition, the figures show that the average use of causal conjunctions in the Arabic sub-corpus is slightly higher than the one in the English sub-corpus.

| Causal Conjunctions | AUNTs | EUNTs |
|----------------------------|--------------|--------------|
| Total of Occurrences | 136 | 102 |
| Average Percentage | 0.94 | 0.65 |

Table 4: Causal Conjunctions Frequency and Average Percentage in the PCUNTs

c) *Temporal Conjunctions in the PCUNTs*

The frequency and percentage of temporal in the PCUNTs demonstrate that the two sub-corpora tend to use temporal conjunctions quite similarly. The figures in Table 5 reveal that the average use of temporal conjunctions in the Arabic sub-corpus is almost the same compared to the one in the English sub-corpus.

| Temporal Conjunctions | AUNTs | EUNTs |
|------------------------------|--------------|--------------|
| Total of Occurrences | 96 | 83 |
| Average Percentage | 0.70 | 0.56 |

Table 5: Temporal Conjunctions Frequency and Average Percentage in the PCUNTs

d) *Continuative Conjunctions in the PCUNTs*

Table 6 reveals that continuatives are used much less frequently in the parallel corpus. The comparison between the occurrences of continuatives shows that the two sub-corpora display slight differences. The frequency of use of this category in the English TTs is higher than the Arabic STs. Similarly, the average use of continuative conjunctions in the English sub-corpus is slightly higher than the one in the Arabic sub-corpus. However, because this difference is not quite obvious, it cannot be said that there are significant differences between Arabic texts and their English translations.

| Continuative Conjunctions | AUNTs | EUNTs |
|----------------------------------|--------------|--------------|
| Total of Occurrences | 52 | 89 |
| Average Percentage | 0.33 | 0.44 |

Table 6: Continuative Conjunctions Frequency and Average Percentage in the PCUNTs

e) *Adversative Conjunctions in the PCUNTs*

This subcategory is the least conjunctive device used in the parallel corpus. The frequency of occurrence and average percentage of this semantic relation in Table 7 show that adversative conjunctions are quite scarce in the two sub-corpora. That is to say, occurrences of this category are very few in both AUNTs and EUNTs. The average use of adversatives in the two sub-corpora is scarce and almost identical.

| Adversative Conjunctions | AUNTs | EUNTs |
|--------------------------|-------|-------|
| Total of Occurrences | 27 | 21 |
| Average Percentage | 0.17 | 0.12 |

Table 7: Adversative Conjunctions Frequency and Average Percentage in the PCUNTs

Results of the Statistical Analysis of Conjunctions in the PCUNTs

By using SPSS, the results analysed are subjected to a paired t-test, in order to find out whether or not there are significant differences between the two sub-corpora. They are presented as follows:

| Semantic Category | AUNTs Sub-corpus | | EUNTs Sub-corpus | | Paired Samples Test | |
|------------------------------|------------------|------|------------------|------|---------------------|-----------------|
| | M | SD | M | SD | t | Sig. (2-tailed) |
| Subcategories of Conjunction | | | | | | |
| Additives | 3.90 | 0.72 | 1.71 | 0.50 | 12, 11 | 0.000 |
| Adversatives | 0.17 | 0.16 | 0.12 | 0.13 | 2, 220 | 0.039 |
| Causals | 0.94 | 0.65 | 0.64 | 0.47 | 2, 651 | 0.016 |
| Temporals | 0.70 | 0.50 | 0.56 | 0.35 | 1, 762 | 0.094 |
| Continuatives | 0.33 | 0.20 | 0.44 | 0.22 | - 1, 963 | 0.065 |

Table 8: T-test Results for the Subcategories of Conjunction in the PCUNTs

The results show that statistically significant differences between the two sub-corpora are identified in two subcategories of conjunction: additives ($t = 12, 11, p = 0.000 < 0.05$), and causals ($t = 2, 651, p = 0.016 < 0.05$). Thus, the results reveal that these two types are considered as statistically significant relations. However, in the case of adversative, temporal and continuative devices, no statistical significant differences exist between AUNTs and EUNTs. The proportion of p is higher than 0.05 in these three sub-categories: adversatives ($t = 2, 220, p = 0.039 > 0.05$), temporals ($t = 1, 762, p = 0.094 > 0.05$) and continuatives ($t = -1, 963, p = 0.065 > 0.05$).

Discussion of the Results

Based on the findings, it can be said that conjunctions are not employed in AUNTs as frequently as in EUNTs. Significant differences are found in two out of five subcategories, namely, additives and causals in terms of their frequency of occurrence, but no significant differences are identified in the statistical result of adversatives, temporals and continuatives. In fact, because the UN texts are consistent in logic and well structured, conjunctions are predominantly used to display the logical relations between propositions. The high or low frequency of subcategories of conjunctions in the two sub-corpora is attributed to the following aspects, which characterise the organisation of the UN documents:

The nature of the UN documents, which is typically descriptive and rigid, explains the remarkable predominance of additives in the PCUNTs. In fact, the

prevalence of additives is ascribed to the writers' intention to add information, clarify ideas or opinions of their governments, and support the facts or events that took place.

In addition, the importance given to causals and temporals in the second and third ranking respectively is due to, first, the argumentative characteristics of the UN texts that aim to persuade readers, and, second, to their narrative function of successions of facts, events or precise reports that have taken place.

Moreover, the use of continuatives justifies the importance of logical relations in the PCUNTs. The two languages highlight the explicit means to show semantic relations at the inter-sentential level, that is why continuatives are clearly employed. Continuatives like *in this regard* and *with reference to* are maintained in English in order to show clearly the semantic relations of sequence. The same applies to their Arabic counterparts.

However, the category of adversatives is the least frequently employed conjunction in the parallel corpus. Adversatives are employed essentially with the aim of highlighting contrasts, drawing conclusions or providing details. For the aim of precision and accurateness, this category is avoided in the two languages, suggesting that adversativity is not significant for legal texts, and the use of contrasting conditions is not typical for this type of texts.

Cross Language Interpretation

As far as cross language differences are concerned, the distribution of conjunctions in the Arabic and English sub-corpora is the same. The only difference lies in the frequency of occurrence of those relations. It seems that Arabic texts are more explicitly cohesive than their English counterparts through the use of additives, causals and temporals.

The prevalence of these three subcategories of conjunction in the Arabic sub-corpus is attributed to the narrative function of successions of facts, events or precise reports that have taken place in some of the UN texts. Williams (1989) argued that Arabic is close to the human lifeworld; it uses the narrative form in texts of any type, more than English does. This explains why all argumentations in Arabic are expressed through narrative structures.

Regarding additives, Al-Jabr (1987:147), who examined narrative texts, maintained that this category "incorporates 'right branching' and highlights the seriation of events in this type of discourse". The frequent use of causals is ascribed to the argumentative characteristics of the UN texts, which aim to convince readers. Smith and Farwley (1983, in Al-Jabr, 1987) explained that causality is predominantly employed in the English texts, as it serves a useful function in narrative texts. Moreover, temporals are found to be of high frequency of occurrence as Arabic uses narration more than English does. Williams (1989) suggested that showing a higher employment of temporals and a smaller one of causals is typical in Arabic texts than is the case in English.

It is evident that the additive relation has a higher percentage than the other subcategories in each of the twenty AUNTs. While Arabic cohesion relies mostly on additives, English cohesion relies more on causal, temporal and continuative relations. The in-depth analysis of data reveals that the mostly used additive devices in the parallel corpus are conjunctions **و** and its English equivalent *and*, expressing the semantic relation of addition. The comparison between the Arabic STs and their English translations reveals also that Arabic texts use a much larger number of additive **و** than their English counterparts *and*.

Moreover, the findings reveal that Arabic uses a greater proportion of causals than English. This is attributed to the fact that Arabic tends to make relationships between sentences more explicit, and finds it important to clarify this relation more than English. Regarding temporal conjunctions, the UN texts employ these items in the two sub-corpora, quite similarly. This is ascribed to the tendency of writers and translators to make the sequential/chronological relationships between sentences explicit, as they reflect the narrative function of successions of facts, events or precise reports that have taken place. The significant proportion of temporal conjunctions in the UN texts is, therefore, used mainly to express the external time form rather than how the writers/translators organised their thoughts.

Briefly said, it is obvious that the use of conjunctions of various categories plays an important role in the organisation of legal documents, particularly the UN texts, as it guarantees the configuration of the logical flow of Arabic source texts and their English translations. Findings reveal that the distribution of conjunctions in the Arabic and English sub-corpora is the same. The similarities are maintained for the purpose of accurateness, transparency and logicity that characterize the language of the UN texts. The only difference, however, lies in the frequency of occurrence of those relations. It seems that Arabic texts are more explicitly cohesive than their English counterparts through the use of additives, causals and temporals.

Conclusion

To sum up, using the corpus-based approach to contrastive and translation studies the study shows that the distribution of conjunctive devices in Arabic and English displays many similarities in the choice of the types of conjunction. Nevertheless, significant differences appear in the frequency of occurrence of these types across the two languages. The differences in textual cohesion are said to be due to language particular preferences and legal genre specificity. That is to say, they are due to the stylistic features of legal texts, and because of the two languages internal rules and stylistic preferences in employing these

devices, which, as a result, represent tendencies typical of each language. This paper may help both students of translation and professionals working in law improve their discourse skills and produce more accurate translations.

References

- Al-Jabr, A.M. (1987). *Cohesion in text differentiation: A study of English and Arabic*. (Doctoral thesis, University of Aston, Birmingham). Retrieved from <http://eprints.aston.ac.uk/10288/>
- Anthony, L. (2011). AntConc Build 3.2.4. [Free computer Software Tool]. Centre for English language education in science and engineering. Waseda University: Tokyo. Retrieved February 25, 2014 from <http://www.antlab.sci.waseda.ac.jp/software.html>
- Anthony, L. (2013). AntConc Build 3.4.1. [Free Computer Software Tool]. Centre for English language education in science and engineering; Waseda University: Tokyo. Retrieved February 25, 2014 from <http://www.antlab.sci.waseda.ac.jp/software.html>
- Bystrova-McIntyre, T. (2012). *Cohesion in translation: A corpus study of human-translated, machine-translated, and non-translated texts (Russian into English)*. (Doctoral Dissertation, Kent State University). Retrieved from https://etd.ohiolink.edu/rws_etd/document/get/kent1353451112/inline
- Giannossa, L. (2012). *A corpus-based investigation of lexical cohesion in English and Italian non-translated texts and in Italian translated texts*. (Doctoral thesis, Kent State University). Retrieved from https://etd.ohiolink.edu/rws_etd/document/get/.../inline
- Halliday, M.A.K., & R, Hasan. (1976). *Cohesion in English*. London: Longman.
- Hanting, P. (2013). Translating conjunctive cohesion in legal documents. *Perspectives: Studies in Translatology*, 22 (1), 1-20. doi: 10.1080/0907676X.2013.777463.
- Ketabi, S. & Jamalvand, A.A. (2012). A Corpus-based Study of Conjunction Devices in English international law texts and its Farsi translation. *International Journal of Linguistics*, (4), 4, 362-371. doi:10.5296/ijl.v4i4.2578.
- ODS. Document System of the United Nations. Retrieved from <http://documents.un.org/>
- SPSS. (2013). IBM Corp. Released. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.
- Trebits, A. (2009). Conjunctive cohesion in English language EU documents – A corpus-based analysis and its implications. *English for Specific Purposes*, 28 (2009) 199–210. doi:10.1016/j.esp.2009.04.004
- Williams, M.P. (1989). *A Comparison of the textual structures of Arabic and English written texts: A study in the comparative orality of Arabic*. (Doctoral thesis, University of Leeds).
- Zhao, J., Yan, W., & Zhou, Y. (2009). A Corpus-based study of cohesion in English medical texts and its Chinese translation. *International Journal of Biomedical Science*, 5 (3), 313-320. Retrieved from www.ijbs.org