



The use of a corpus management tool for the preparation of interpreting assignments: a case study

Pablo S. Pérez-Pérez
Universidad de Málaga, Spain
psperez@uma.es

DOI: 10.12807/ti.110201.2018.a08

Abstract: This paper is about a series of experiments conducted at the University of Malaga as part of a PhD thesis with students from the fourth and final year of the Degree in Translation and Interpreting. The experiments were done in order to analyse whether the use of a corpus management programme in the preparation of the vocabulary for an interpreting assignment had an impact on their interpreting outputs. Our hypothesis was that the use of the programme would have a positive impact on interpreting assignments. To carry out a case study, for some experiments the students prepared the vocabulary in advance without the use of the programme, and for other experiments the vocabulary was prepared using the programme. The results from the different groups of students were compared to see which of the groups performed better, the students who prepared the vocabulary with the use of the corpus management programme, or those who did not use the programme.

Keywords: interpreting, corpus, word list, concordances

1. Introduction

New technologies have proven to be very useful and necessary for the work of translators. There is a large selection of computer-aided translation tools available in today's market (Bowker and Corpas Pastor, 2015). Corpus management programmes would be an example of one of these tools used in translation, not only in the profession, but also for research in Translation Studies for decades, as there are many areas of research where studies have been and can be carried out in translation with the use of these programmes. As Bendazzoli and Sandrelli point out (2009): “the development of CTS (corpus-based studies on written translation) has been more advanced than the development of CIS (corpus-based interpreting studies) since the very beginning of this scholarly venture”, but although research with corpus and interpreting is not so developed, Corpas Pastor (2008, p. 98) observes that there is a growing interest in the use of corpus for research, training and the professional practice of interpreting.

One of the uses of a corpus management programme for interpreting can be that of helping interpreters when they prepare for a conference interpreting

assignment. This paper will show what features can be of use to interpreters when a corpus management programme is used to help in the preparation process of the vocabulary.

When a conference interpreting assignment comes up, the client is expected to provide the interpreter with all the information available in advance about the conference assignment, so as to allow the interpreter a good and thorough preparation of all the vocabulary that will or couldⁱ be used during the conference. The interpreter may receive from very little information (e.g. only the date, venue and the name or the topic of the conference) to very detailed information, including the full conference schedule, and all the relevant documents from all the speakers (their papers, computer presentations, and CVs). This is not the norm: as Torres Díaz confirms, conference organisers do not usually have much knowledge about the profession and do not take care in providing the interpreters with the necessary documents to prepare the interpreting assignment in advance (Torres Díaz, 1998, p. 44).

2. Description of functions of corpus management programmes for the preparation of the vocabulary for an interpreting assignment

Corpus management programmes can help interpreters obtain the vocabulary that could appear in a particular conference, and information necessary for the preparation for the assignment. There are many different corpus management programmes available; for example, AntConc (Anthony, 2014) or WordSmith Tools (Scott, 2012). The interpreter needs to upload documents about the topic of the assignment to the programme. The documents are used as input for the corpus management programme. It is very important to make sure that the documents that are going to be used for the preparation of the vocabulary come from trustworthy sources. These would use the proper vocabulary and terminology relevant to the conference topic.

2.1. Word List

By uploading the documents to the corpus management programme, it will be able to create a ‘Word List’ⁱⁱ. The programme will automatically put together all the words from all the documents that have been uploaded to it, and it will provide a list with the words that appear the most often in the documents that have been used as a documentation source related to the topic of the interpreting assignment and, in this list, the terms will be ordered by frequency. If the words at the top of the list are those which are used the most in relation to a specific topic, it will very likely mean that the words with such high frequency of appearance will be amongst those which will be used the most during the conference. The digitised programme provides, in a faster manner, a list of the most frequently used words from a given topic. This is

ⁱ *Could* is used in this sentence because, as Jimenez Serrano affirms, there is a difference between a ‘finite’ and an ‘infinite’ text. This means that in translation the work is carried out on an already defined and ‘finite’ text, whereas in interpreting the text (or discourse) is provided during the conference when the speaker is taking part in it, and thus, it tends to be unknown or ‘infinite’ (Jiménez Serrano, 1998).

ⁱⁱ For the purpose of this study, the freeware AntConc programme by Laurence Anthony from Waseda University in Tokyo, Japan was used: <http://www.laurenceanthony.net/software.html>

one of the many benefits of using the programme, in comparison to preparing a glossary manually, which would take a long time, as all the documents would have to be read one by one by the interpreter, in order to find and extract the words and phrases for a glossary. As Costa et al. (2014, p. 31) observe:

corpora and corpus management tools have proven most beneficial for interpreters as a device to speed up the preparation phase and to improve the quality of the input. A corpus can provide vast amounts of domain expert knowledge and accurate terminological and phraseological information in an efficient, effortless and inexpensive way.

The Word List provided is a full list with all types of words from the documents, i.e. not only the vocabulary that an interpreter would need to include in a glossary for an interpreting assignment, but also basic and too unspecific words such as articles, prepositions, easy nouns and verbs, and so on. In order to avoid the inclusion of easy words, a ‘Stoplist’ can be used as a ‘filter’ list of words. A ‘Stoplist’ is a list of words which we do not want to have in the Word List provided by the corpus management programme.

Once the stoplist is uploaded to the programme, if we then ‘ask’ the programme to provide us with the Word List, we will see how the words included in our stoplist (articles, pronouns, etc.) are not shown in the new Word List.

2.2. Clusters

The programme can also provide terms which are used and formed by more than one word, known as *n-grams* (Yannakoudakis and Angelidakis, 1988; Oakes, 1998), as we can ask the programme to give us a list of words which tend to be used together with another word (or another two words, three words, etc.). These are called clusters (for example a noun which tends to be used with a certain adjective, also known as collocations). This is another useful piece of information for the interpreter to prepare the vocabulary for a conference, as sometimes in a topic a term with a high frequency of use can consist of more than one word, i.e. a *multiword unit* (Barreiro et al., 2013, p. 27). Carrying out this task manually would be very time-consuming.

In order to create a glossary for a conference with the use of the ‘Word List’ and the ‘Clusters’, the steps described above have to be followed for the two languages used during the conference. The programme allows saving the information provided in a .txt format, which interpreters will be able to open on their computers and then work with and save in another format which they might prefer to create their glossary in any relevant language combination.

2.3. Concordances

When interpreters are preparing for an assignment, it is not sufficient to prepare a list of words in the working languages. They need to know and understand the meaning of the words which will potentially be part of the assignment, as these words will be used within a context during the conference, and not individually. The more knowledge interpreters have about the topic before the conference, the betterⁱⁱⁱ.

ⁱⁱⁱ Ding (2017) conducted an experiment on knowledge of propositional content where participants in two groups received different levels of background information before an interpreting task. The control group received a list of terms related to the speech topic, whereas the experimental group received the same list of terms plus bilingual *Translation & Interpreting* Vol 10 No 1(2018)

For this reason, there is another feature that can be of use for interpreters when preparing the vocabulary for a conference. If interpreters need to have further information about any of the words provided by the programme in the Word List, they would only need to click on any of the words on which more information is needed, and the programme will automatically move to the 'concordances' screen. This feature will show interpreters all the lines of text where any given term is included in all the documents that have been uploaded to the programme.

Finally, to read the word in context, another click on any of the lines will show the full text including the term on which the interpreter needs more specific information.

3. Potential advantages of corpus management programmes for vocabulary preparation for conference interpreting

As has been described above, a corpus management programme can be used by interpreters as a tool for preparation of a conference assignment, providing a great deal of information that otherwise would have to be obtained in a traditional manner, i.e. by having to extract the words one by one or having to look for a specific word in a text to read it in context. The following is a summary of the potential advantages of the use of a corpus management programme for the preparation of related vocabulary that the interpreter might use in a conference:

- a) the programme provides the interpreter with a Word List ordered by frequency;
- b) a frequency list of terms which are formed by more than one word (multiword units) can also be obtained, referred to in the programme as 'clusters';
- c) the results of the word lists can be saved in a .txt format, and then copied and pasted in a format that the interpreter might normally work with, in order to prepare the glossary for the interpreting assignment;
- d) the words from the Word List can also be seen in a context, by using the concordances feature;
- e) if more information about a given word is needed, the interpreter can very easily access the full text in which the word occurs by clicking on the concordance line from the list, which will provide the context for a thorough understanding of the meaning;
- f) by using the programme, time is saved in the preparation for the conference interpreting assignment, as the interpreter does not have to read and analyse each individual document that is going to be used for the preparation of the assignment, because these documents can be uploaded to the corpus management programme, which will provide the interpreter with the vocabulary in less time in comparison with having to do it in a traditional manner.

background articles. Participants in the experimental group performed better than participants in the control group.

4. Case study

A case study was carried out in order to assess the impact of the use of a corpus management programme on the preparation of the vocabulary that will likely be used in an interpreting assignment. This was part of a PhD thesis using corpus methodology in an empirical study (Pérez Pérez, 2013), where we had originally speculated that the case study could result in any of these three potential findings, to wit: a) the use of the programme would have a positive impact on the outcome of the interpreting assignment; b) it would have no positive or negative impact whatsoever; c) it would have a negative impact on the outcome of the interpreting assignment. Our hypothesis was that the use of the corpus management programme for the preparation of the vocabulary for an interpreting assignment would have a positive impact on the outcome of the interpreting assignment.

To analyse the impact of the use of a corpus management programme for the preparation of an interpreting assignment, a series of experiments were conducted at the University of Malaga with students of the fourth (and final) year of the Degree in Translation and Interpreting undertaking the subject of Simultaneous Interpreting (English ↔ Spanish). The students took part in the experiments voluntarily and they were randomly divided into two groups:

Table 1. Students included in Group 1 and students included in Group 2

Group 1	These students prepared the vocabulary for the interpreting assignments with no indications whatsoever as to how to carry out such preparation (i.e. the preparation was done in a 'traditional' way, extracting the information and the vocabulary manually from the documents they looked for and used for the preparation of the vocabulary);
Group 2	These students prepared the vocabulary for the interpreting assignments with the use of the corpus management tool.

4.1. Participants

After we spoke to their lecturer, and in order to recruit volunteers for the experiments, the students were advised that they were going to do simultaneous interpreting exercises in which they would have to interpret from English to Spanish about a topic that they were going to be told about in advance. The experiments were then carried out outside of their normal class schedule, which allowed for the possibility of working with both groups separately and also meant that the students turning up were willing to participate. A total of 27 students participated voluntarily in the experiments with three different topics. Some of them only took part in experiments where they did not have to use the computer programme to prepare the vocabulary; other students took part first in experiments without using the programme, and then in experiments using the programme to interpret a different topic; and some students only took part in experiments where they had to use the computer programme to prepare the vocabulary. Based on the vocabulary preparation method, all participants were then included in either group 1 or group 2, as is shown in Table 1. A total of 48 simultaneous interpreting performances were carried out with three different topics.

4.2. Description of the experiments

The experiments were about the following topics: 1) 'Spas: health benefits'; 2) 'Solar energy: solar panels'; and 3) 'Iridology'. For each topic a fictional

presentation was created in English^{iv} with a duration of approximately 8 minutes. The following table shows the number of students in each group for each topic:

Table 2. Number of students in each group for each topic

Topic	Distribution	Distribution
Spas	8 students / Group 1	8 students / Group 2
Solar Panels	8 students / Group 1	8 students / Group 2
Iridology	8 students / Group 1	8 students / Group 2

The information about the topics to be interpreted by each group of students was sent exactly at the same time in advance (7 days before the interpreting session). By providing the students with the information the same number of days in advance, the only difference between each group was the use of the corpus management programme for the preparation of the vocabulary with Group 2, in comparison with Group 1 students preparing the vocabulary in a traditional way. Group 2 students who had to use the programme received training on how to work with it. The training was offered outside of their normal class schedule, as was the case for all interpreting experiments, and prior to receiving information about the topic they were going to interpret.

Below is the English translation of the Spanish emails sent to the students before the experiments:

Example 1 is the email sent to the students who had to prepare the vocabulary for the interpreting assignment without the use of the corpus management programme:

These are the instructions for the simultaneous interpreting experiment that will take place next Wednesday 25th March at 1.10 p.m. in the Simultaneous Interpreting Room.

The exercise will be:

Simultaneous interpreting from English into Spanish
Topic: 'Spas: health benefits'.

You will be in the booths (one person per booth) and when you are ready, I will start reading the text in English which you will have to interpret into Spanish. The duration of the exercise is approximately 8 minutes.

In order to prepare for this exercise, you will have to look for documents related to the topic of the interpreting assignment. With documents about the topic, both in English and in Spanish, you will be able to prepare the vocabulary that might be used during the exercise.

I'll see you on Wednesday 25th.

Thank you all very much in advance.

Pablo Pérez

^{iv} Although the subject the students were undertaking involved English < > Spanish interpreting, all experiments were carried out from English into Spanish, so that the results could be compared in the same interpreting direction.

Example 2 is the email sent to the students who had to use the corpus management programme to prepare the vocabulary and had been taught in advance how to use the programme:

These are the instructions for the simultaneous interpreting experiment that will take place next Monday 25th May at 3.30 p.m. in the Simultaneous Interpreting Room.

The exercise will be:

Simultaneous interpreting from English into Spanish
Topic: 'Spas: health benefits'.

You will be in the booths (one person per booth) and when you are ready, I will start reading the text in English which you will have to interpret into Spanish. The duration of the exercise is approximately 8 minutes.

In order to prepare for this exercise you will have to look for documents related to the topic of the interpreting assignment. With documents about the topic, both in English and in Spanish, and with the use of the computer programme that was explained to you last week, you will be able to prepare the vocabulary that might be used during the exercise.

I'll see you on Monday 25th.

Thank you all very much in advance.

Pablo Pérez

P.S. Attached please find the 'Stoplists' in English and in Spanish.

As can be seen in these emails, the instructions were exactly the same for both groups with the only difference relating as to how students were asked to prepare for the vocabulary that might be used. Also, because finding the right 'Stoplist' or even preparing one is time-consuming, the 'Stoplists' were provided to Group 2 students, which meant that they could then focus on looking for documents related to the topic and preparing the vocabulary (as Group 1 students who were not asked to use the programme). The same wording was used in emails sent to each of the groups of students for the other topics.

The outputs of the students were recorded as individual mp3 files, so that a transcription could be done of each of the students' performances in order to carry out the analysis. Transcribing the outputs of the students was a necessary stage of the process prior to the analysis. The transcription of the individual mp3 files of each of the students took a lot of time and care, because, as Meyer (2002, p. 66) points out: "Transcribing speech is an extremely lengthy process, requiring the transcriber to listen to the same segments of speech over and over again until an accurate transcription is achieved". In order to allow for a rigorous subsequent analysis, we had to be extremely careful when transcribing the students' outputs.

4.3. Results of the experiments

To assess the impact of the use of the corpus management programme in the interpreting assignments, for each of the three topics used in the experiments, a number of terms were analysed to see the number of matches of each of the students. In the case of the first topic 'Spas: health benefits', 87 terms from the text used during the interpreting exercise were analysed to see how many had been interpreted correctly by the students in Group 1, in comparison with the students in Group 2. In the second topic 'Solar energy: solar panels', 73 terms

from the presentation were used for the analysis and to compare the results between both groups of students. And in the case of the third topic ‘Iridology’, 63 terms were used to analyse the results.

Once all the experiments were completed and all student outputs were transcribed and analysed, we were able to determine whether the use of a corpus management programme had had an impact on the outcome of the interpreting assignment. To do so, we compared the number of matches of each of the students in each group, of the terms analysed in each topic.

4.3.1. Results of experiment number 1: ‘Spas: health benefits’

The following table shows the results of the experiments of the first topic, ‘Spas: health benefits’. 16 students took part in the interpreting assignment of this topic, 8 of them were in ‘Group 1’ and did not use the corpus management programme to prepare the vocabulary, and the other 8 were in ‘Group 2’ and used the programme to prepare the vocabulary for the interpreting assignment.

Table 3. Results for ‘Group 1’ (no use of the programme) and for ‘Group 2’ (use of the programme) for the 87 terms used for analysis in this topic:

GROUP 1	Number of matches	Percentage of matches
Student 1	25 out of 87	29 %
Student 2	32 out of 87	37 %
Student 3	30 out of 87	34 %
Student 4	35 out of 87	40 %
Student 13	34 out of 87	39 %
Student 14	32 out of 87	37 %
Student 15	27 out of 87	31 %
Student 16	19 out of 87	22 %
	Group 1 Average →	34 %
GROUP 2	Number of matches	Percentage of matches
Student 5	27 out of 87	31 %
Student 6	31 out of 87	36 %
Student 7	38 out of 87	44 %
Student 8	41 out of 87	47 %
Student 17	28 out of 87	32 %
Student 18	30 out of 87	34 %
Student 19	31 out of 87	36 %
Student 20	47 out of 87	54 %
	Group 2 Average →	39 %

The average percentage of matches in the analysed terms, in the case of the students who had prepared the topic about spas without the help of the corpus management programme was 34 %, while the average percentage in the case of the students who had prepared the same topic with the use of the corpus management programme was 39 %.

4.3.2. Results of experiment number 2: ‘Solar energy: solar panels’

Again 16 students took part in the experiment for the second topic, ‘Solar energy: solar panels’: 8 in ‘Group 1’, who did not use the programme for the preparation of the vocabulary, and 8 in ‘Group 2’, who did.

For this topic, the average percentage of matches in the analysed terms, in the case of the students who prepared the vocabulary without the help of the corpus management programme was 26 %, while, this time, the average percentage in the case of the students who prepared the topic with the use of the corpus management programme was 34 %.

Table 4. Results for ‘Group 1’ (programme not used) and ‘Group 2’ (programme used) for the 73 terms analysed in this topic:

GROUP 1	Number of matches	Percentage of matches
Student 2	28 out of 73	38 %
Student 9	23 out of 73	32 %
Student 4	25 out of 73	34 %
Student 10	18 out of 73	25 %
Student 21	14 out of 73	19 %
Student 22	19 out of 73	26 %
Student 23	13 out of 73	18 %
Student 24	13 out of 73	18 %
	Group 1 Average →	26 %
GROUP 2	Number of matches	Percentage of matches
Student 5	23 out of 73	32 %
Student 6	27 out of 73	37 %
Student 7	29 out of 73	40 %
Student 11	29 out of 73	40 %
Student 13	30 out of 73	41 %
Student 25	13 out of 73	18 %
Student 20	24 out of 73	33 %
Student 26	21 out of 73	29 %
	Group 2 Average →	34 %

4.3.3. Results of experiment number 3: ‘Iridology’

The following table shows the results for the third topic, ‘Iridology’. 16 students took part in this experiment, again 8 in each group: one with the assistance of the corpus management programme and the other without.

Table 5. Results for ‘Group 1’ (no use of the programme) and ‘Group 2’ (programme used) for the 63 terms analysed in this topic:

GROUP 1	Number of matches	Percentage of matches
Student 12	13 out of 63	21 %
Student 5	19 out of 63	30 %
Student 7	19 out of 63	30 %
Student 8	9 out of 63	14 %
Student 17	8 out of 63	13 %
Student 20	21 out of 63	33 %
Student 25	19 out of 63	30 %
Student 19	22 out of 63	35 %
	Group 1 Average →	26 %
GROUP 2	Number of matches	Percentage of matches
Student 6	25 out of 63	40 %
Student 9	28 out of 63	44 %
Student 4	26 out of 63	41 %
Student 2	38 out of 63	60 %
Student 27	11 out of 63	17 %
Student 13	30 out of 63	48 %
Student 16	11 out of 63	17 %
Student 26	24 out of 63	38 %
	Group 2 Average →	38 %

The average percentage of matches in the analysed terms, in the case of the students who prepared the vocabulary for this topic without the use of the corpus management programme was 26 %, while in the case of the students who used the corpus management programme to prepare the vocabulary for the topic, the average percentage of matches was 38 %.

4.4. Analysis of the results of the experiments

For this case study, a total of 48 interpreting performances were recorded, transcribed and then analysed to see if the number of matches with the terminology used in the presentations of the three topics was higher in the group of students who had not used the programme to prepare the vocabulary for the assignment, or in the group of students who had used the programme. Because three different topics were used in the experiments, this also allowed us to see if the results changed if the topic was different. Also, by using three different topics, students could take part in different experiments, i.e. first in experiments where the vocabulary for a topic had to be prepared in a traditional manner, and then after receiving the explanation about how to use the corpus management programme, in experiments where the vocabulary for a different topic had to be prepared using the programme. This has allowed the possibility of analysing whether the results of one same student were different with and without the use of the tool.

As can be seen from the results tables, the average percentage of matches is higher in the case of the students who used the corpus management programme to prepare the vocabulary for the interpreting assignments. This has been the case in each of the three topics used in these experiments. Table 6 below shows a summary of the results of the average percentage of matches of the students:

Table 6. Summary of results of the average percentage of matches of the students from each group in each topic

Topic	GROUP 1 (prepared vocabulary traditional way)	GROUP 2 (prepared vocabulary using corpus management programme)	Difference
Spas	34 %	39 %	+ 5 %
Solar Panels	26 %	34 %	+ 8 %
Iridology	26 %	38 %	+ 12 %

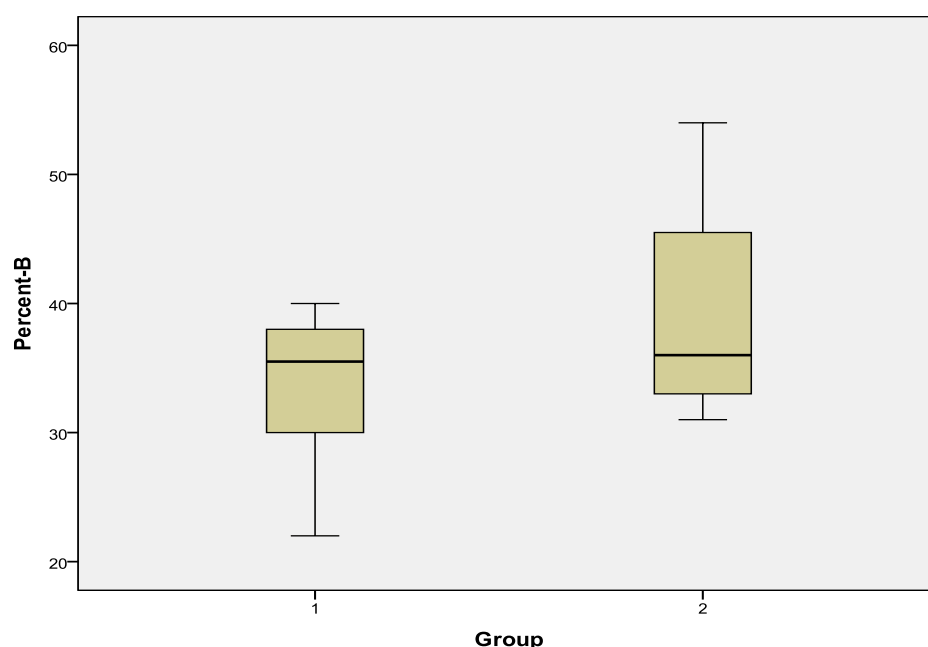
Some students took part in different experiments, as three topics were used. For some experiments they prepared the vocabulary without the use of the corpus management tool and for other experiments they then used the programme. They were students numbers 2, 4, 5, 7, 8, 9, 13, 16, 17, 19, 20 and 25. When comparing the percentages of matches in the terms analysed in each topic for these students, it can be seen that, in general, there is a higher percentage of matches in the cases where the corpus management programme was used for the preparation of the vocabulary for the interpreting assignment. Only students number 16 and 25 had a lower percentage of matches in the experiments where they used the programme to prepare the vocabulary. All the other students who took part in experiments with different topics and with different ways of preparing the vocabulary (with and without the use of the programme) had higher percentages of matches in the analysed terms when preparing the vocabulary with the corpus management programme. These were students 2, 4, 5, 7, 8, 9, 13, 17, 19 and 20.

It is worth highlighting the results of Student 2. This student seems to be an outlier, already performing well, and then showing an even more marked improvement after using the corpus management programme. He/she took part in two experiments where the vocabulary was prepared without the use of the

programme, and the percentages of matches were 37 % and 38 % respectively, whereas in the experiment where he/she prepared the vocabulary with the use of the programme, the percentage of matches was 60 %. The results of Student 7 also stand out, as in the experiment without the programme he/she had 30 % of matches, and in the other two experiments for the other two topics prepared with the use of the corpus management tool, he/she obtained a result of 40 % and 44 % of matches. As can be seen from these results, both experiments carried out by this student with the help of the programme had a higher percentage of matches in the terminology than the experiment prepared without the corpus management programme. And finally we have the case of students numbers 8, 9 and 17, who took part in two experiments and who carried out a traditional manual preparation of the vocabulary for the first experiment and then used the computer programme for the second, also with a better performance as an outcome. Their results were as follows: Student 8 had 14 % of matches in the analysed terminology in the first experiment (without the use of the programme), and then 47 % of matches in the second (with the assistance of the programme), a performance that is three times better than the first one; Student 9 had 32 % of matches without the use of the programme and 44 % of matches in the experiment prepared with the programme; and Student 17 had 13 % of matches without the programme, and 32 % with the assistance of the tool (more than double the score in the first instance).

4.4. Statistical representation of the results

A statistical representation of the results also shows how these are better in the cases when the corpus management programme was used for the preparation of the vocabulary, in comparison to when it was not. A box plot^v has been used for the statistical representation of the results as represented in Figures 1, 2 and 3:



^v A box plot (Rodriguez, 2007, p. 111) is a graphic developed by John Tukey used to summarize visually a distribution of scores by using boxes and whiskers.

Figure1. Statistical representation for the percentage of matches on the experiments about “Spas”.

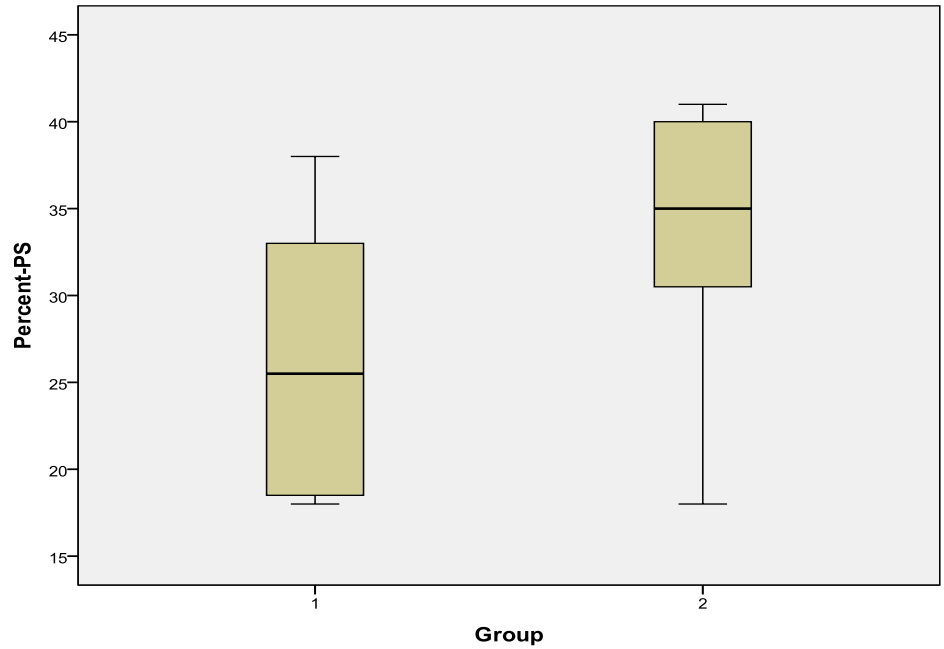


Figure 2. Statistical representation for the percentage of matches on the experiments about “Solar Panels”.

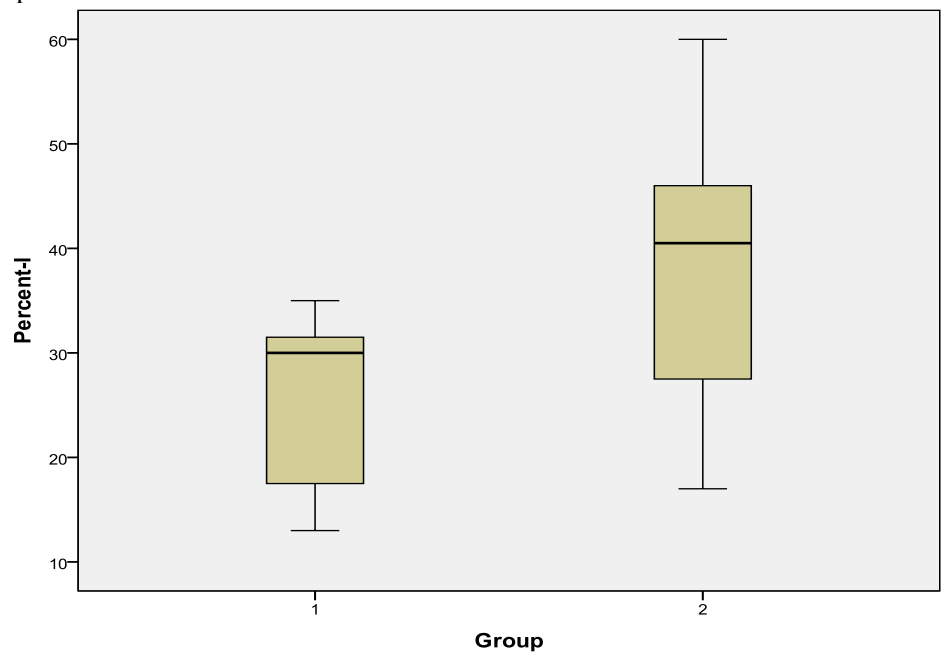


Figure.3. Statistical representation for the percentage of matches on the experiments about “Iridology”.

Figures 1, 2 and 3 show Group 1 students (who did not use the corpus management programme to prepare the vocabulary) represented on the right side, and Group 2 students (who used the programme) on the left side. As can be observed in these graphs, data values in the box plots for Group 2 are

Translation & Interpreting Vol 10 No 1(2018) 148

higher than those in the box plots for Group 1. In other words, vocabulary-related preparation using the corpus management programme proved beneficial.

5. Conclusion

This paper discussed how a corpus management programme could help student interpreters prepare the vocabulary for a conference interpreting assignment, with a description of the features of the programme that can be useful for the interpreter in the preparation. The paper reflected on a series of experiments carried out to analyse the impact of the use of a corpus management programme. Our hypothesis was that the use of the programme for the preparation of the vocabulary would have a positive impact on interpreting assignments. In order to test the hypothesis, a series of simultaneous interpreting experiments were carried out with interpreting students of the final year of the Degree in Translation and Interpreting at the University of Malaga. Some students prepared the vocabulary without the use of a corpus management programme, and other students used the programme to prepare the vocabulary for the exercises. A total of 27 students took part voluntarily in the experiments. Some of them only participated in experiments where they did not have to use the computer programme to prepare the vocabulary, i.e. they prepared the vocabulary in a traditional manual way; other students took part first in experiments without using the programme, and then in experiments using the programme; and some students only took part in experiments where they had to use the computer programme to prepare the vocabulary. A total of 48 simultaneous interpreting performances were carried out involving three different topics.

Once the experiments were completed, all 48 interpreting performances of approximately 8 minutes were transcribed. Then a series of terms included in each of the presentations of the three topics were analysed to see if they had been interpreted correctly by the students, so as to compare the number of matches from students who had not used the corpus management programme to prepare the vocabulary for the assignment, with the number of matches from students who had used the computer programme.

The results were then available in number of matches and in percentage of matches, by topics and by groups of students (with and without the use of the computer programme). The results show that, in general, the number of matches of the students who had prepared the vocabulary with the corpus management programme was higher than the number of matches of the students who had prepared the vocabulary in a traditional manner. This supports our hypothesis that the use of a corpus management programme has a positive impact on the outcome of the interpreting assignment.

Another argument for using a corpus management programme is to reduce interpreters' overall preparation time. As it is a computer tool, on first use, the process of uploading the texts and handling the software might take some time. Users will normally need to learn how to use any new computer programme, but once they become more familiar with it, the time required for such use is reduced. This means that, when the corpus management programme is used on several occasions, it is handled in a more 'mechanic' manner, and the time needed to operate the programme stops being an issue, as it is very much reduced as is the usual time required to prepare for a conference interpreting assignment.

This study showed how students undertaking conference interpreting studies benefited from the use of the corpus management programme. It would be very interesting to conduct further research where interpreting students are instructed in the use of such programmes, to assess how this might benefit them in terms of preparation of the vocabulary for formative or summative interpreting exercises. Such further studies could also include students working with other language pairs and with other types of assignments, including for instance court or health interpreting. Moreover, by learning about the tool, student interpreters may continue to benefit from the use of such programmes in professional practice.

Acknowledgements

The research presented in this work has also been partially developed in the framework of the research project ‘VIP: Voice-text integrated system for InterPreters’ (ref. no. FFI2016-75831, MINECO) and the Educational Innovation Project ‘INTERPRETA 2.0: Application of ICT tools for the teaching-learning process of interpreting’ (ref. no. PIE17-015, type B).

References

- Anthony, L. (2014). *AntConc* (Version 3.4.4) [Computer Software]. Tokyo, Japan: Waseda University.
- Barreiro, A., Monti, J., Batista, F., & Orliac, B. (2013). When multiwords go bad in machine translation. In J. Monti, R. Mitkov, G. Corpas Pastor & V. Seretan (Eds.), *Proceedings of the 14th Machine Translation Summit. Workshop on multi-word units in machine translation and translation technologies* (pp. 26–33). Allschwil: European Association for Machine Translation.
- Bendazzoli, C., & Sandrelli, A. (2009). Corpus-based Interpreting Studies: Early work and future prospects. *Revista Tradumàtica* 7. Retrieved from <http://www.fti.uab.cat/tradumatica/revista/num7/articles/08/08art.htm>
- Bowker, L., & Corpas Pastor, G. (2015). Translation technology. In R. Mitkov (Ed.), *The Oxford handbook of Computational Linguistics 2nd edition*. Oxford, England: Oxford University Press.
- Corpas Pastor, G. (2008). *Investigar con corpus en traducción: los retos de un nuevo paradigma*. Frankfurt, Germany: Peter Lang.
- Costa, H., Corpas Pastor, G., & Durán Muñoz, I. (2014). Technology-assisted interpreting. *Multilingual*, 143: 27–32.
- Ding, Y. L. (2017). Using propositional analysis to assess interpreting quality. *International Journal of Interpreter Education*, 9(1): 17–39. Retrieved from <http://www.cit-asl.org/new/using-propositional-analysis-to-assess-interpreting-quality>
- Jiménez Serrano, O. (1998). El intérprete de simultánea ante la terminología médica (inglés-español): preparación y dificultades. In L. Félix Fernández & E. Ortega Arjonilla (Coords.), *Traducción e interpretación en el ámbito biosanitario* (pp. 339–352). Granada: Comares.
- Meyer, C. F. (2002). *English Corpus Linguistics: An introduction*. Cambridge, England: Cambridge University Press.
- Oakes, M. P. (1998). *Statistics for Corpus Linguistics*. Edinburgh, Scotland: Edinburgh University Press.
- Pérez Pérez, P. S. (2013). *La enseñanza de la interpretación de conferencias: un estudio empírico con metodología de corpus*. PhD thesis, Universidad de Málaga.
- Rodriguez, W. (2007). Box Plot (Box and Whisker Plot). In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 111–112). Thousand Oaks, CA: Sage Publishers.
- Scott, M. (2012). *WordSmith Tools* (version 6) [Computer Software]. Liverpool: Lexical Analysis Software.
- Torres Díaz, M. G. (1998). *Manual de interpretación consecutiva y simultánea*. Malaga, Spain: Universidad de Málaga, Servicio de Publicaciones.
- Yannakoudakis, E. J. & Angelidakis, G. (1988). An insight into the entropy and redundancy of the English dictionary. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6): 960-970.