[*Translating and the Computer 27*, November 2005 [London: Aslib, 2005]

HOW MUCH DOES USING CONTROLLED LANGUAGE IMPROVE MACHINE TRANSLATION RESULTS?

Nathalie de Preux
Multilingual Processing Department,
Ecole de Traduction et d'Interpretation, University of Geneva, Switzerland
Nathalie.DePreux@eti.unige.ch

ABSTRACT

In many companies, especially those relying heavily on technical documentation, the demand for translation heavily outstrips what the translation services can supply. Thus, many translation departments feel the need to make use of machine translation systems (good, bad and indifferent, depending on what is available) and to look for ways to improve the quality of their output. One favourite solution is to use a controlled language. This article reports on a small experiment designed to see to what extent using a controlled language does really improve the output of a machine translation system.

INTRODUCTION

Many companies now make use of machine translation (MT) in order to respond to the ever increasing demands of the translation market. The primary reason for turning to MT is to increase productivity whilst keeping costs as low as possible. This implies, in turn, keeping post-editing costs *as* low as possible. However, despite considerable progress in computational linguistics, MT is still a long way from being able to offer results that can be used without any revision. A major concern then, of those companies that use MT, is to improve the quality of the system produced raw translation as much as possible. One promising approach seems to be to influence the input text, especially by constraining the lexical items present and the grammatical constructions used - in other words, restricting input text to a controlled language.

In the work reported here, we were interested in quantifying to what extent controlled languages helped to improve MT output. We thus carried out an experiment with the aid of translation students from the University of Geneva translation school and of professional translators from European Centre for Nuclear Research (CERN). The purpose of the experiment was to find out whether the use of a controlled language did in fact result in improved translation, and, if this was the case, the extent of the improvement. We carried out a quantitative and qualitative evaluation of the translation produced for two parallel texts, one in a controlled language, the other not. The MT system used was Reverso. Furthermore, we used two different types of controlled language in order to see whether the nature of the controlled language also influenced the results.

Here we briefly present the experiment, followed by a description of the corpus used for the experiment. Then we follow through the various stages of the evaluation. In particular, we discuss the classification of errors, as well as the qualitative criteria chosen as part of the evaluation of the translation. Finally, we present the results, followed by some conclusions that could be drawn from them.

1. The Experiment

We assembled a corpus made up of parallel examples of controlled language texts and free texts. All the texts came from two companies who use controlled language (CL) as an authoring aid. We translated these texts with a MT system in order to determine whether the CL texts gave better results than the free texts. In the interests of being as objective as possible, we carried out an evaluation of the results in two stages. In the first step we counted the number of errors in each translation, taking into account how serious we considered the error to be. This method provided us with a precise quantitative indication of which version of the text (controlled or free) produced the best results. Secondly, we carried out a qualitative evaluation by asking people with experience of technical translation to evaluate the translations produced by the two versions.

Preparing the experiment posed two main problems: the search for CL examples and the choice of the MT system. We were fortunate in being able to obtain Simplified English (SE) texts from Boeing and Caterpillar Technical English (CTE) texts from Caterpillar. These are amongst the most well-known and most frequently used CLs for technical translation. As we hinted earlier, the CL rules for the two are rather different. CTE defines a sub-set of the English lexicon and grammar adapted to the authoring of the maintenance manuals, user manuals and servicing manuals produced by Caterpillar (Hayes et al, 1996: 84). It can be thought of as a pre-editing language, oriented essentially to the needs of the computer system. Simplified English was developed by AECMA during the 70s as a way of improving documentation produced within the aeronautics industry (Gingras 1987, in Shubert et al 1995). Its main purpose is to make texts more easily comprehensible for a human reader (Allen 1999).

As a MT system for the experiment, we chose Reverso. The primary reason was that this system is easily accessible to a large number of users and produces relatively acceptable results on the whole.

1.1. The corpus

The parallel controlled and free texts used come from Boeing user manuals and Caterpillar maintenance manuals. Since we thought it important that the experiment should reflect real working conditions in the two companies, it was critically important to find examples of sentences which had already been checked by the technical authors at Boeing and Caterpillar. In choosing example phrases for testing, we tried to take phrases which were long and complex enough to be interesting, but still simple enough to allow our test subjects to read and understand them rapidly (Holmback et al 1996).

We finally settled on five units of CL and five units of free language for each of our two CLs. A unit is made up of one or several phrases that constitute a single instruction. The length of a unit is between

13 and 47 words. The total number of words in the corpus is 124 for the free version of the Caterpillar text, 138 words for the controlled version of the same text, 158 words for the free version of the Boeing text, and 172 words for the controlled version of this same text.

2. EVALUATION STEPS

Before trying to translate the texts, we first created a personal dictionary containing the terms and the technical verbs appearing in the texts but not in Reverso's dictionary. We also provided target language equivalents. One of our reasons for choosing Reverso as an experimental system is that we thought it important to be able to modify the system in such a way as to be able to produce the best results possible. Reverso makes this possible to any user, without having to go through the manufacturer.

Next, we needed a reference translation for each of our chosen sentences. This reference translation would be used in assessing what errors were present in the results. We machine translated the example sentences, corrected the translation ourselves and then asked experts in technical authoring and translation to check our corrected version.

2.2. Quantitative evaluation

The first step of the evaluation is purely quantitative. We simply counted the errors present in each translation in order to determine whether the translation of controlled sentences contained fewer errors than the translation of free sentences. We did not consider all errors to be equal: we weighted each error using a weighting scheme based on the model proposed by Arnold et al. 1994.

2.2.1. The errors

In order to classify the errors appropriately, we made use of proposals made in the context of human translation by Jean Delisle et al (1999) and Hurtado Albir (2001). But since these proposals were made in the context of human translation, we simplified the classification somewhat and also added three categories specific to certain types of errors occurring in MT but not in human translation. We also chose the categories to be informative in the context of our own experiment. The final classification covers eight types of errors (definitions of category 2, 3, and 4 come from the work of Delisle et al. 1999):

Punctuation errors: use of punctuation signs that do not correspond to the rules of the target language.

Under-translation error: omission of any compensation or explicitation required to obtain an idiomatic translation.

Over-translation error: explicitation of elements of the source text that ought to be implicated in the target text.

Solecisms: producing a syntactical structure that does not conform to grammatical conventions of the given language.

Faulty syntactic analysis: error in the syntactical category attributed to a word resulting from an incorrect source text analysis.

Lexical incorrectness: error in the semantic meaning given to a word.

Stylistic clumsiness: when the translation is grammatically correct but is not idiomatic.

Constructions having no sense: an error in which a whole part of the sentence is incorrect, making the translation impossible to understand. The cause of this mistake is very difficult to determine, as there is usually more than one cause.

2.2.2. *Classifying the errors*

According to Delisle (1993, in Hurtado Albir 2001: 190), translation errors can be grouped into two categories. The first is "language faults", which are due to an imperfect knowledge of the target language. The second are "translation errors", which are caused by a "poor interpretation of the source text". These two groupings correspond to the analysis and generation modules of a MT system. We therefore grouped our eight error categories into two larger groups as follows:

a) Faults occurring during the translation process: over-translation, under-translation and faulty syntactic analysis.

b) Language errors: punctuation errors, stylistic clumsiness, lexical incorrectness and solecisms.

2.2.3. *Relative severity of errors*

Intuitively, the errors made by a MT system differ in the repercussions they produce on the quality of the output translation: they cannot all be considered as of equal importance. Each type of error should be multiplied by a precise weighting factor which reflects the seriousness of the error (see Arnold et al 1994: 173-75). Each phrase is given a score which depends on the number of errors it contains, as well as the weighting factor attributed to each error. Taking as a starting point the work of Arnold et al (1994: 173-75), Sager (1989, in Hurtado Albir 2001: 295) and Dancette (1989, ibid: 303), we first classify the errors as **minor,** or superficial (errors which only have repercussions at the level of the form of the text) and **serious** (errors which have an incidence on the sense of the phrase).

Independently of this distinction, we classified errors due to faulty syntactic analysis of the source text and lexical incorrectness errors into four groups (definitions come from the work of Delisle et al. 1999):

False nuance: minor errors of sense which make little or no difference to how the phrase is understood.

Incorrect meaning: a sense is attributed to a word or a segment from the source text that is does not have in the context in which it appears.

Misinterpretation: a word or segment from the source text is given an entirely erroneous sense from that intended by its author.

Nonsense: illogical formulation in the target text due to misinterpretation of the source text or wrong translation methodology.

Finally, we fixed the weight of each type of error by drawing inspiration from the five criteria which Hurtado Albir (2001: 304) puts forward as determining the severity of the error:

1. the importance of the error with respect to the original text as a whole
2. its importance with respect to the coherence and cohesion of the target text
3. the degree to which the sense deviates from the original text
4. its importance with respect to the communicative aspects of the target text
5. its impact on the purpose or functionality of the text

Table 1 below gives the categories of errors, their degree of severity and the weight attributed to them.

Table 1

Error category	Type of error	Relative severity of errors	Weight
Translation error	Under-translation	a. Minor b. Serious (omission)	0.2 0.4
	Over-translation	a. Minor b. Serious	0.1 0.3
	Faulty syntactic analysis	a) Minor i. False nuance b) Serious i. Incorrect meaning ii. Misinterpretation iii. Nonsense	0.1 0.2 0.4 0.6
Language error	Lexical incorrectness		
	Solecism	a. Minor b. Serious	0.2 0.4
	Punctuation error	(minor)	0.1
	Stylistic clumsiness	a. Minor b. Serious	0.1 0.3
Ambiguous category	Constructions having no sense	(serious)	0.8

2.3. Qualitative evaluation

In the second part of the evaluation we applied qualitative criteria to the evaluation of the translations produced by Reverso

2.3.*1. Criteria for the qualitative evaluation*

We used three different models in defining the criteria for qualitative evaluation, that of Arnold et al (1994: 169-71), that of Hutchins and Somers (1992: 163-64) and that of Holmback et al (1996: 171). Taking these models into account, we picked out four criteria to assess machine translations: the comprehensibility or clarity of the translation, the precision or fidelity in terms of the source text, respect for the form of the source text (style, register, level of language used), and the usefulness of the translation produced.

We added this last criterion as a way of measuring the usefulness of the raw translation when considered as a starting point for a human producing a final version. We constructed a questionnaire using these criteria which served as a basis for the qualitative evaluation.

The questionnaire itself opens with a brief account of the experiment and of the reasons for carrying it out, as well as instructions on how to complete the questionnaire. Then the four criteria are set out in such a way that the test persons can give a score ranging from one to five four each criterion. As an example, the score 1 for comprehensibility indicates that the test person considers the translation incomprehensible, and 5 that he considers it perfectly clear.

2.3.2. *The test persons*

For our results to have any validity, it was important that the test persons should have previous experience of technical translation. We thus recruited test persons amongst students of the translation school following a course on technical translation. It should be noted that the students were towards the end of their translation training (between two and four years of translation studies already successfully completed), and therefore far from being neophytes in translation work. They had French as their mother tongue, with English amongst their passive languages. We also recruited a second group of professional technical translators from CERN, which meant that the level of experience amongst our test persons was quite high. In all, 12 students and 8 professionals made up our group of twenty subjects.

2.3.3. *Carrying out the experiment*

We divided the translations produced by Reverso into four groups:

1. Translations of the CL version of the Boeing sentences
2. Translations of the free version of the Boeing sentences
3. Translations of the controlled version of the Caterpillar sentences
4. Translations of the free version of the Caterpillar sentences.

Each participant received one group, which contained five sentences. We then distributed the questionnaire. In order to preserve validity of the first criterion, the comprehensibility of the translation, the subjects should not have had access to the source text. Thus, the scoring of this first criterion was done before the source text was distributed. Subsequently, subjects scored precision, respect of the form of the original and the usefulness of the translation produced referring to the source text as they wished. The evaluation using the student group took place during a technical translation class which was part of their normal studies and lasted twenty minutes. The questionnaires were distributed to the technical translators of CERN at their work place. Given a total of 20 subjects, each translation of a given version was scored by three students and two translators.

3. RESULTS

3.4. Results of the quantitative evaluation

Once the questionnaires were filled in, we counted the errors produced by the MT system in each sentence and calculated the total number of points for each sentence taking account of the weighting function explained earlier. The sentence with the highest number of points was the sentence with the most errors or with the most serious errors. Thus, according to this metric, the lower the score, the better the translation.

Table 2 below shows an example of a sentence annotated with the errors it contains and their type. The free version of the text has been placed in parallel with the controlled version of the text in order to facilitate comparison. Table 3 gives a list of the acronyms used in annotating the errors. Table 4 gives a summary of the structure of table 2. Thus, table 4 tells us that line 1 of table 2 contains the original version of the sentence. Line 2 contains the translation produced by Reverso. Finally, line 3 contains the translation as corrected by the revisors (the reference translation).

The parts of the sentence which have been corrected are in bold characters. The lower part of table 2 gives a summary of all the errors found in the sentence. These errors are multiplied by their weighting (See Table 1 for the weightings of each error); then the results are totalled to give an overall score for the sentence. Thus, direct comparison of the final score makes it easy to see whether the CL version has resulted in a better translation.

Table 2

	Free version	Controlled version
1	Inspect all components for wear or damage and replace the components if it is necessary.	Inspect all components for wear or for damage. Replace the components, if necessary.
2	Inspecter tous les composants pour **[UT+]** **l'**usure**[SC-]** ou **[UT-] endommager[FSA (NS)]** et remplacer les composants si **c'est[UT-]** nécessaire.	Inspecter tous les composants pour **[UT+]** **l'**usure**[SC-]** ou pour **des[SC-]** dégâts. Remplacez les composants**.[P]** si nécessaire.
3	Inspecter tous les composants pour **verifier qu'ils ne présentent pas d'**usure ou **de dégâts** et remplacer les composants si nécessaire.	Inspecter tous les composants pour **verifier qu'ils ne présentent** pas d'usure ou **de** dégâts. Remplacez les composants si nécessaire.
	Score	**Score**
	UT+: 1*0.4 = 0.4 UT-: 2 * 0.2 = 0.4 SC-: 1 *0.1 =0.1 FSA (NS): 1 * 0.6 = 0.6 **TOTAL SCORE: 1.5**	UT+: 1 * 0.4 = 0.4 SC-: 2 * 0.1 = 0.2 P: 1*0.1=0.1 **TOTAL SCORE: 0.7**
	Comments : In this example, the MT system has produced a better result out of the controlled version. It was able to recognize the word « damage » as a noun and not as a verb, as it does in the free version. The reason for this mistake in the free version is the ambiguity which is solved in the controlled version.	

UT- : minor under-translation **UT+** : serious under-translation (**OM** : omission)	
OT- : minor over-translation **OT+** : serious over-translation	
FSA : faulty syntactic analysis	**(FN)**: false nuance **(IM)** : incorrect meaning
IC : lexical incorrectness	**(MS)**: misinterpretation **(NS)**: nonsense
S - : minor solecism **S+** : serious solecism	
P : Punctuation error	
SC - : minor stylistic clumsiness **SC+** : serious stylistic clumsiness	
CNS : Construction having no sense	

Table 3

Line 1	Original sentence
Line 2	Translation by Reverso
bold + underlined	Errors made by the system
Line 3	Reference translation
bold	Parts that were corrected

Table 4

3.4.1. Overall results of the quantitative evaluation

Table 5 summarizes the overall results of the quantitative evaluation. For the Caterpillar text, the translation of the five units of sentences in their free language version obtained a score of 9.7, which should be compared with the score of 7.3 obtained by the CL version. The difference of 2.4 between the two scores indicates that the CL versions gave rise to translations of better quality. To say this a little differently, the score given to the errors produced went down by 24.8% in the case of the CL version. It is interesting to note, however, that the total number of errors is the same for both the free and the CL versions: 36 in each case. This tends to indicate that using a CL does not reduce the number of faults overall, but does tend to reduce the gravity of the errors. Being able to show this is, of course, one of the reasons for introducing a weighting factor onto the calculation of error scores. Although not discussed here, the weighting function should in practice be determined by considering the users of the raw translation and their expectations of the output. Looking now at the Boeing texts, we see that the free version resulted in a score of 8.5 for the translations, and the CL version in a score of 6.4. Once again, this could be re-expressed by saying that applying a CL produced a 24.7% reduction in the error rate, a figure almost identical to that obtained for the Boeing texts. The difference between the two types of text is clearer, on the other hand, if take into account the total number of errors produced irrespective of their gravity: the free version produces a translation with 40 errors, and the CL version a translation with only 30 errors. Thus, in this case, using a CL has produced a 25% reduction in the number of errors in the translation.

	Caterpillar				**Boeing**			
	Number of errors		**Score**		**Number of errors**		**Score**	
	Free version	CL version	Free version	CL version	Free version	CL version	Free version	CL version
Sentence 1	5	4	1.5	0.7	5	3	1.2	0.7
Sentence 2	3	3	0.9	0.8	4	6	0.4	1.2
Sentence 3	11	11	3.1	2.4	12	7	2.5	1.4
Sentence 4	6	6	1.0	1.0	8	9	1.7	1.7
Sentence 5	11	12	3.2	2.4	11	5	2.7	1.4
TOTAL	**36**	**36**	**9.7**	**7.3**	**40**	**30**	**8.5**	**6.4**

Table 5

An important factor that we have not so far mentioned in discussing the total number of errors is the number of words in each version. As mentioned when we were discussing the corpus, the free version of the Caterpillar text contains 124 words. If we analyse the total number of errors, we can conclude that an error is produced for every 3.4 words in the input text. On the other hand, for the CL version of

the same text, containing 138 words the rate is one error for every 3.8 words. Once again, we observe that even if the total number of errors is the same for both versions of the original text (36), the relation between the total number of words and the number of errors produced is not the same. So we can again say that the translation of the CL version of the Caterpillar text is better than the translation of the free version of the same text.

With the Boeing texts, (158 words in the free version and 172 words in the CL version), the free version produces a translation with one error for every 3.9 words, and the CL version one error every 5.7 words. Here too then, the language controlled version produces a translation which is clearly of better quality than that produced by the free version.

One of our initial interests was to look at the differences in translation quality in relation to the CL used. On the basis of these results we might be tempted to conclude that the application of Simplified English results in a translation superior to that resulting from the application of Caterpillar Technical English. However, we shall see below that the results of the qualitative evaluation seem to indicate the opposite conclusion.

Figure 1 and 2 below illustrate the difference between the total number of errors and the final score attributed to each version. These diagrams show that the number of errors is the same for the translation of each of the two versions of the Caterpillar text, whilst for the Boeing text there are fewer faults in the CL version. On the other hand, the score given to the errors in the free version of the two texts is indisputably higher. This means that the translations resulting from the language controlled versions are of better quality in the two cases, notwithstanding that the total number of faults is the same for free and CL versions of the Caterpillar text.

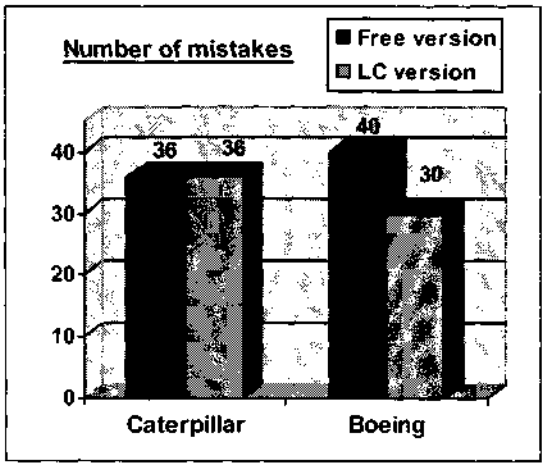

Figure 1

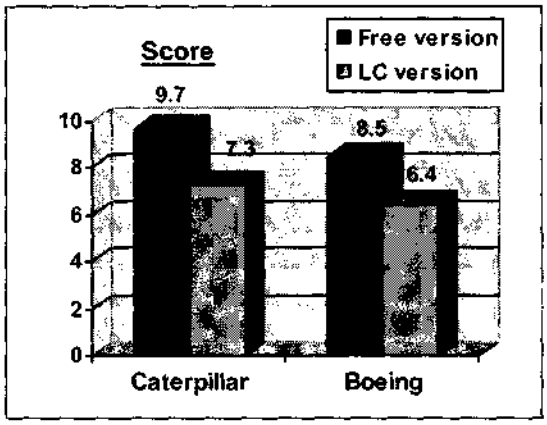

Figure 2

3.5. Results of the qualitative evaluation

In considering the results of the qualitative evaluation, we kept apart the scores given by the students and the scores given by the professional translators, suspecting that differences in experience and background between the two groups might influence the results. We first calculated the average of the translation resulting from each version according to the four criteria chosen (comprehensibility,

accuracy, respect of source form and usefulness). Tables 6 and 7 below show the results for the student group and the professional group respectively, and also show the differences between the results produced by the free version and by the CL version. The lower part of each table shows the overall score attributed to each version, taking all the criteria into account. Bold figures show the cases where the CL version produced a better score than the free version.

Table 6 is based on the student results. For the Caterpillar text, the CL version gained a better score only for the criteria of comprehensibility (0.4 better) and accuracy (0.5 better). For the criteria respect of the form of the original and usefulness of the translation, the score produced by the language controlled version is inferior to that produced by the free version. With the Boeing texts, all results are worse for the CL version. The overall scores are negatives for the CL version for both texts: -0.1 for the Caterpillar text and -0.3 for the Boeing text.

Table 7 shows that the results from the student assessment differ significantly from those of the assessment carried out by the professional translators. For the Caterpillar text, the language controlled version scores higher for all four criteria: 1.2 more for comprehensibility, 0.4 for accuracy, 0.6 for respect of form and 0.9 for usefulness. In the case of the Boeing text, the CL text scores 0.3 more on comprehensibility and 0.6 more on accuracy, but does less well with the other two criteria, scoring -0.2 for respect of form whilst usefulness gets the same result for both versions. Overall, both texts produce positive results for the CL versions: the Caterpillar CL text scores 0.8 more than the free version, and the Boeing CL text scores 0.2 more than the free version.

ANSWERS FROM STUDENTS		Score / 5					
		Translation of the Caterpillar text			Translation of the Boeing text		
		Free version	LC version	Difference	Free version	LC version	Difference
Criteria	Comprehensibility	2.6	2.2	**0.4**	2.2	2.4	- 0.2
	Precision	3.5	3.0	**0.5**	3.3	3.7	- 0.4
	Respect for the form of the source text	3.1	3.5	- 0.4	3.1	3.4	- 0.3
	Usefulness as draft translation	2.5	3.4	- 0.9	2.3	2.6	- 0.3
Average score		2.9	3.0	- 0.1	2.7	3.0	- 0.3

Table 6

ANSWERS FROM TRANSLATORS		Score / 5					
		Translation of the Caterpillar text			Translation of the Boeing text		
		Free version	LC version	Difference	Free version	LC version	Difference
Criteria	Comprehensibility	3.2	2.0	**1.2**	3.2	2.9	**0.3**
	Precision	3.5	3.1	**0.4**	3.6	3.0	**0.6**
	Respect for the form of the source text	3.1	2.5	**0.6**	2.8	3.0	- 0.2
	Usefulness as draft translation	3.4	2.5	**0.9**	2.8	2.8	0.0
Average score		3.3	2.5	**0.8**	3.1	2.9	**0.2**

Table 7

Figures 3 and 4 show the overall results for each version, taking all criteria into account. The scores attributed by the students appear at the right, those of the professional translators at the left.

According to the data of figure 3, in the students' judgement the CTE version is not as good as the non-CTE version. On the other hand, the professional translators reach exactly the opposite conclusion. Furthermore, the difference between the scores given to the different versions by the translators (0.8) is greater than the difference

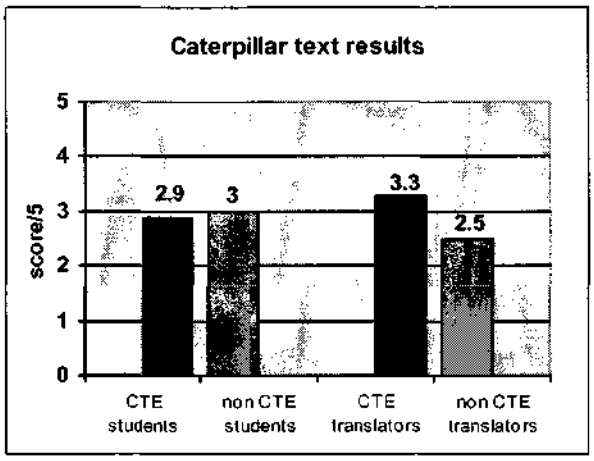

Figure 3

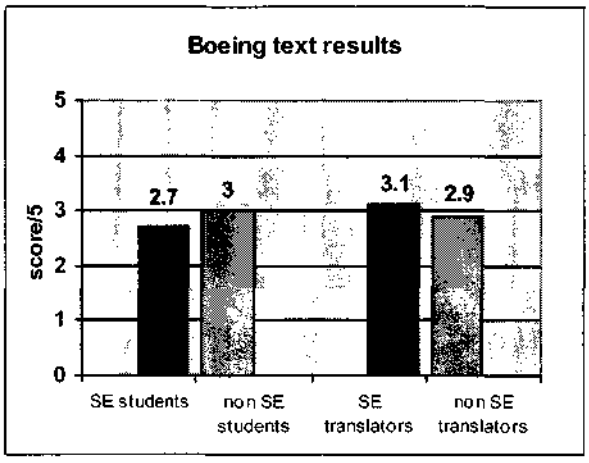

Figure 4

between the scores given by the students (-0.1). Thus, we can conclude that applying a CL to the Caterpillar text before its translation by Reverso results in a slight improvement in the quality of the translation.

In contradistinction to the overall results for the Caterpillar texts and to the results produced by the quantitative evaluation, the translators' results show a slight improvement in the quality resulting from the SE version. The student results favour CTE, showing that the result

produced for the SE version is of lower quality. This difference in the student results (-0.3) is greater than the difference in the translators results (+0.2). We can therefore conclude that according to these results, the application of SE to the Boeing text has not produced a consequent increase in translation quality with a translation produced by Reverso. It is even possible that applying SE had a negative effect, and that the translation was better on the text to which SE had not been applied. However, the figures are too small for us to be bale to justify such a conclusion.

CONCLUSION

The main purpose of the study reported here was to discover whether applying a CL to a text before it was automatically translated was a good strategy to follow in terms of the quality of translation thus produced and any consequent reduction of post-editing time. We looked at the effect of two different controlled languages on the machine translations produced by Reverso. Extracts of parallel texts taken from two maintenance manuals produced by Boeing and Caterpillar were used in their free and language controlled versions as input to the MT system. We could thus evaluate the quality of the translations produced and subsequently compare the results produced as a consequence of free or CL input.

The results are relatively satisfying, especially when we consider the quantitative part of the evaluation. It proves to be the case that once we count the number of errors in the translation of each version and take into account the gravity of the errors, we see an improvement of about 25% for texts to which CL has been applied. On the other hand, the results of the qualitative evaluation, carried out with the help of students in technical translation and professional technical translators are not so positive, with improvement being assessed at only around 8%. Of course, the subjective nature of the criteria used make these results suspect. Thus, we believe that on the whole these results show, as one might expect, that texts produced with the aid of a CL lead to better translations (by MT systems) than do free texts.

Considering the two different controlled languages used in this experiment, the results of the quantitative evaluation indicate a better improvement with SE than with CTE. This result surprised us, since SE is not a machine oriented CL. But on the other hand, the results of the qualitative evaluation favour CTE, which show a higher improvement rate (16% according to the results of the translators' assessment) than the 4% due to SE (based on the same results). Considering the nature of CTE, a CL design for MT systems, this result was more in accordance to what we had expected.

BIBLIOGRAPHY

ALLEN, Jeffrey (1999) "Different Types of Controlled Languages", *TC-Forum magazine,* Vol. 1-99, pp. 4-5.

ARNOLD, Doug, Lorna, BALKAN, R.Lee, HUMPHREYS et al. (1994) *MT: An Introductory guide,* Oxford, NCC Blackwell Ltd.

DELISLE, Jean, Hannelore, LEE-JAHNKE, Monique C., CORMIER (1999) *Terminologie de la Traduction,* Amsterdam, John Benjamins B.V.

HAYES, Phil, Steve, MAXWELL, Linda SCHMANDT (1996) "Controlled English Advantages for Translated and Original English documents", *Proceedings of the First International Workshop on Controlled Language Applications (CLAW 96),* pp. 84-92.

HOLMBACK, Heather, Serena, SHUBERT, Jan, SPYRIDAKIS (1996) "Issues in Conducting Empirical Evaluations of Controlled Languages", *Proceedings of the First International Workshop on Controlled Language Applications (CLAW 96)*, pp. 166-177.

HUIJSEN, Willem-Olaf (1998) "Controlled Language - An Introduction", *Proceedings of the Second International Workshop on Controlled Language Applications (CLAW 98)*, pp. 1-13.

HURTADO ALBIR, Amparo (2001) Traducción y traductologia. Introducción a la traductologia, Madrid, Ediciones Cátedra.

HUTCHINS, W. John and Harold L., SOMERS (1992) *An Introduction to MT,* London, Academic Press.

SHUBERT, Serena, Jan, SPYRIDAKIS, Heather, HOLMBACK et al. (1995) "The Comprehensibility of Simplified English in procedures", Technical Writing and Communication, Vol. 25, n° 4, pp. 347-369.