Translation Quality Assessment:
Can a Corpus-based Approach be Effective?
Mª Beatriz Rodríguez Rodríguez
Universidad de Santiago de Compostela

Este trabajo pretende subrayar el papel del corpus en la evaluación de textos traducidos. Despúes de esbozar el análisis evaluativo, se analiza la relevancia de las herramientas de corpus que se pueden manejar en la evaluación de traducciones, como por ejemplo corpus bilingües, alineamiento de textos paralelos, herramientas Woodsmith y ParaConc, concordancias y listas de palabras. Evidentemente, la utilidad de estas herramientas varía en relación con el tipo de texto. De hecho, varios programas se han aplicado con éxito a la evaluación de traducción automática en detrimento de otras traducciones publicadas. Como cabría suponer, los mayores problemas surgen a la hora de evaluar traducciones literarias. Finalmente, cabe destacar el uso de corpus en la didáctica de la traducción.

En conclusión, el manejo de corpus puede reducir significativamente el elemento subjetivo en la evaluación de traducciones, aunque el análisis se debe completar con otros recursos para lograr conclusiones objetivas y fiables.

This paper aims at highlighting proposals concerning corpora tools to be applied to translation assessment. After discussing on translation assessment analysis, computational tools and methods are analysed, including bilingual corpora, parallel text alignment, Wordsmith and ParaConc tools, concordances, and word lists. Obviously, the utility of these tools varies accounting for the text type. In fact, computer programmes have been successfully applied to machine translation assessment. Concerning literary translation, a great number of drawbacks can be encountered due to the nature of these texts. Interestingly enough, corpora in translation training allow students to be aware of available translation strategies and to assess errors, which improves their competence.

Drawing in these premises, it is argued that adopting a corpus-based approach can significantly reduce the subjective element in translation quality assessment, although it has to be complemented with other resources concerning the evaluator's competence and critical judgement to reach accurate conclusions.

Palabras clave: traducción; evaluación; análisis; herramientas; corpus.
Key Words: translation; assessment; analysis; corpora tools.

1. INTRODUCTION

Given that Translation Quality Assessment entails making objective judgements concerning translation, an exhaustive analysis of the target and source text is required to reach reliable and accurate conclusions. This paper aims at highlighting the valuable role of corpora in this contrastive assessment analysis. After the central stages of Translation Quality Assessment Analysis are explored, corpora tools to be applied to translation assessment are outlined. Needless to say, approach varies accounting for the target text. Effectiveness of corpora is examined concerning the assessment of published translations, professional translators' work and translation training.

2. TRANSLATION QUALITY ASSESSMENT

Corpora are considered to allow a practical and objective approach which contributes to bridging "the gap between scholarly approaches to translation quality assessment and practical quality assessment" (Lauscher 2000: 164).

Broadly speaking, to assess a translation it is essential to perform a detailed descriptive and contrastive analysis of the target and the source texts at all levels, so as to proceed to apply certain assessment criteria. The analysis must be flexible enough to be adapted to the specific features of each text.

To a certain extent, the central step of this contrastive analysis covers the identification of units of analysis or segments in order to establish the shifts between the target and the source text to be further assessed. Following Toury segments are: "units in comparative analysis would always emerge as couple pairs of target- and source- text segments, 'replacing' and 'replaced' items respectively" (1995: 89). Shifts or deviations between the two texts can be classified into categories of shifts (Leuven-Zwart 1990) so as to account for the translation devices or strategies which they involve such as expansions, reductions, modulations and transpositions (Delisle et al. 1999). The implementation of assessment criteria proves essential to evaluate the appropriateness of the translator's decisions concerning those strategies or devices selected (Brunette 2000: 180). Undoubtedly, the notion of translation quality involves "fuzzy and shifting boundaries" (Bowker 2000: 347). Indeed, there is not consensus on the definition of a good or appropriate translation. Even current definitions appear to be too broad and blurred. Consider, for instance, Halliday's proposal: "A translation (i.e., is equivalent) in respect of those linguistic features which are most valued in the given translation context and perhaps also in respect of the value which is assigned to the original (source language) text" (2001: 17). Within this framework, absolute evaluation in terms of good or bad translation is to be refused (Chesterman 1997: 118); a specific frame of reference is required. In addition, comments on translation quality cannot be restricted to an enumeration of translation mistakes which, although worth considering, do not include any further appreciation. To my knowledge, it is undeniable that translation quality depends on a range of factors (Lausher 2000: 150), which vary accounting for the specific features of each text (Schäffner 1988). Comments concerning translation quality assessment must be accomplished taking these parameters as a point of reference.

Assessment approach varies, as translation quality assessment covers areas such as published translation (mainly literary translation), assessment of professional translators' work (technical, economic, scientific, legal, commercial, etc.), and assessment in translation teaching (Martínez & Hurtado 2001).

The use of corpus as a resource for assessing translations is argued to offer clear advantages during the analysis. Corpora enable us to encode in compact and efficient forms to access and interrogate vast quantities of data. Obviously, electronic texts can be gathered and consulted more

quickly than printed texts. In addition, as qualitative and quantitative techniques are combined, a consistent and reliable empirical analysis can be reached (Bowker 2001).

3. CORPORA TOOLS

Computational tools and methods for processing the corpus and arranging the data are considered to contribute to facilitating the evaluator's work. The handling of corpora provides evidence of how the translator faced translation problems, for segments and shifts can be analysed in a systematic, objective and non time-consuming way. Proposals concerning corpora tools to be applied to translation assessment include parallel or multilingual corpora, text alignment, Wordsmith tools, Parallel Concordancers, and word lists.

Following Mona Baker, parallel corpora (1995: 230) consist of original source texts and their translations. Similarly, multilingual corpora include translations into several languages of the same source text. Corpora must be processed by means of sentence alignment, tokenization and annotation. Despite the fact that the relevance of these processes cannot be underestimated, text alignment has been proved to be crucial for processing a translation oriented parallel corpus as equivalents between target and source texts are covered. Thus, parallel bilingual corpus must be aligned (Véronis 2000) at least at the sentence level to reach more reliable conclusions. Given the fact that we are dealing with units of analysis, a paragraph-correspondence alignment has to be adopted. However, in some cases texts must be "adjusted" manually to have the whole paragraph perfectly aligned in order to make up for the length of sentences, punctuation patterns, word order and syntax. Interestingly enough, multiple versions of texts can be also aligned, which facilitates the contrastive analysis and assessing of different translations of the same source text. As to alignment programmes for parallel texts, several proposals can be applied such as plug word aligner[1] in bilingual parallel corpora for Windows and Linux and Multicord (by David Wools).

In addition, Wordsmith Tools (Scott 1997)[2] and Parallel Concordancer (Barlow 1995)[3] are especially relevant for the alignment and processing of parallel corpora. Wordsmith produces word lists and key word lists of the corpus as well as concordances of key words

(specially worthy at the phrase level) (Bowker and Pearson 2002). Word frequency lists reveal how many different words occur in the corpus and how often each appears. Information provided in these lists aids the evaluator to decide whether the term is used by experts or whether is one author's idiosyncratic preference, which implies a central point of reference in the assessment analysis. That aside, "intercalated text" (Munday 1998) can be applied to word frequency lists so as to obtain a text by manually keying in the target text between the lines of the source text, which enables an easy and systematic classification of segments and their consequent shifts. Interestingly enough, Wordsmith tools enable users to consult multiple corpora at once, which allows the evaluator to contrast different versions of the same source text. Although this type of analysis is often said not to evaluate the quality of a given translation, but to understand the decision process underlying the product of translation and to infer translation norms adopted by the translator, it seems to cover a step especially required in the further assessment. In fact, the contrastive analysis must be contextualised, so that shifts can be accurately examined.

Similarly, ParaConc provides translation equivalents and their lexical collocations in translation corpora. Parallel concordancers identify all examples of a phrase along with all their corresponding translated sentences. Concordances analyse and list the occurrences of a particular search term and its context. Corpus Frequency Data and Collocate Frequency Data are the commands used to carry out these tasks. However, concordance lines generally used as an analytical tool do not always offer enough linguistic context to investigate features of whole texts and semantic phenomena. That aside, ParaConc aligns documents at several levels and accepts up to four parallel texts, which facilitates the comparison and assessment of several translations of the same source text.

As research advances, new current proposals attempt to reduce possible drawbacks. What is more, certain tools have been specifically designed for being applied to specific parallel corpora. XTRA-BI, for instance (University of Deusto) (Abaitua et al 2001), enables the user to reach bilingual segments in parallel corpora, which results in a systematic and reliable analysis. A format TMX is adopted for importing segmented and aligned corpora to translation memory software segments.

The relevance of translation memory software must be outlined as aligned segments (usually sentences of previous bi-texts) are stored.

The vast amount of information to be handled must be highlighted. In addition, information from previous translations can be exploited. Despite the problems and restrictions in the processing of corpora, it is possible to investigate the appropriateness of specific terms or expressions. Having the exact correspondence of texts in both languages definitely contributes to the process of tracing equivalents. Consequently, reliable results can be obtained. Thus, segments can be identified, and their consequent shifts classified and assessed provided that assessment criteria are further applied.

Special attention is paid to how these corpora tools can be applied in the assessing. Needless to say, approaches and advantages vary accounting for diverse types of translations.

4. ASSESSMENT OF PUBLISHED TRANSLATIONS

This area focuses essentially on literary texts assessing. Given the characteristics of these texts more shortcomings are to be faced in the processing of corpora. Firstly, not all published translations are available within a corpus or are even in electronic format. As a result, a parallel corpus has to be built, which implies a more complex process concerning the design criteria, and the building process itself (tagging, formatting, etc). Secondly, despite the fact that proposals investigating the style of literary translators, such as Baker's (2000), attempt to facilitate the approach, the use of a corpus does not appear to be the most appropriate way of evaluating the stylistic features of a literary translation (Bowers 2001: 360); further critical judgement or a partially manual approach is obviously required. Eventually, the implementation of assessment criteria to literary translation seems impossible to be done by means of a computer.

However, after the corpus is aligned and segments identified, databases can be especially designed in an attempt to classify and analyse data in an accurate and easier way, which would facilitate the assessing and the drawing of objective conclusions. Qualitative and quantitative analysis can be carried out in an efficient way provided that criteria and judgements cover the next step. As has just been stated, the use of aligned parallel corpora enables the critic to discover translation shifts and assess the strategies or devices handled to render the source

words. Equivalents between the two texts can be traced using the tools mentioned above, mainly Intercalated Text, ParaConc and XTRA-BI. The handling of multilingual corpora leads to the comparison of shifts within different translations on the basis of a reliable and systematic approach. Translation memories contribute to facilitating the contrasting among different translators' proposals. That aside, it is undeniable that shifts and data can be organised in databases so as to facilitate the tackling and classification of data covering the categories of shifts, even features concerning the translator's style such as repetitions, puns and metaphors. Evidently, a manual approach seems required to design this classification, as well as to implement assessment criteria covering the next step of the assessing.

In any case, other shortcomings are supposed to rise during the research. Needless to say, extratextual factors play an essential role in literary translation. Although extralinguistic information can be provided in the header of the corpus, a lot of information is not patent at all. The translator's prologue, for instance, has to be closely examined to verify that the information provided is objective. Additionally, historical factors affecting the text are to be considered after a first corpora approach to data. Evidently, the practice of translation at the moment the text is written conditions the translator's decisions. As a consequence, the reader to whom the text is addressed is also worth being taken into consideration. Literary conventions, the handling of translator's footnotes and their possible relevance to clarify cultural and linguistic problems deserve being analysed. Eventually, the translator's purpose seems essential to justify the translator's decisions, for his perspective is supposed to have affected the target text. The impact of these parameters or criteria on the target text seems to be beyond the scope of a computer. However, the evaluator's task can be facilitated by a systematic non time-consuming analysis of data and statistic analyses.

New current corpora approaches aim at eliminating shortcomings. Indeed, corpora can be especially designed for the analysis of literary translation. *Palimpset* for windows, for instance, (Maczewski) proposed by the University of Göttingen is a tool for computer assisted literary translation which is thought to browse the texts phrase by phrase in order to display existing links between source and target literary texts. As a result, statistical routines produce database tables containing the selected segments.

5. ASSESSMENT OF PROFESSIONAL TRANSLATORS' WORK

This area covers the translation of technical, economical, scientific, legal and commercial texts. Due to the nature of these translations the definition of quality is supposed to imply more precise features. In fact, several computer programmes have been successfully applied to machine translation assessment to the detriment of other published translations.

As has just been stated, assessment criteria and the notion of quality itself seem to be clearly identified. Standards of quality like DIN2345 and ISO-9000 have been applied in certain approaches. Apart from that, L&L, a Netherlands-based translation and localisation services company suggests the Quality Assurance Model of the Localisation Industry Standard Association (LISA QA Model) published in 1995 and revised in 1999 (Ling Koo and Kids 2000: 47). Eval/Trans (Nieben et al 2000) is a computer programme which aims at the assessment of translation quality in Machine Translation as the comparison of system performance and rate progress seems necessary. It applies the criteria of subjective sentence error data, information item error rate, and average error rate to a corpus which includes source sentences and all evaluated target hypothesis. In these proposals scales to assess the seriousness of the mistake can be directly created in the computer, which results in accurate statements concerning the quality of the target text. Among other current projects it is worth mentioning the Blue Method for Automatic Evaluation of Machine Translation (Kishore et al. 2001) and the NIST MT Evaluation Facility[4]. As can be concluded, the benefits of handling a corpus in professional translations and specifically in machine translation evaluation are beyond any doubt.

6. ASSESSMENT IN TRANSLATION TRAINING

Parallel corpora in translation training allow students to be aware of the available translation strategies and devices to be used in the rendering of source texts (Botley et al. 2000). In addition, translation errors are identified and assessed. As a result, translation students' competence concerning diverse fields and text types (literary, scientific and technical translation, etc.) is assessed and improved as students are supposed not to commit the same mistakes. By means of handling

multilingual corpora and translation memory software, segments and shifts covering different types of texts or different translators' proposals can be contrasted. This fact allows students to be conscious of translation strategies handled to tackle translation problems. If learners are given access to corpora, they can be encouraged to observe translation and language in use (Aston 2000). Translators' decisions are examined and assessed, which enables students to face all steps of the translation process and factors affecting translation. As this suggests, linguistic information is identified during the assessment to be further examined. As can be inferred, the use of corpora in translation training can enhance students' communicative competence in both languages, strategic competence, and transfer competence (Zanettin et al. 2003) by means of judging a wide variety of translation shifts and strategies[4].

Finally, it is worth highlighting that Lynne Bowker, among other authors, assures that students' translations can also be appropriately assessed by means of corpora (2000: 354; 2003). Bowker proposes an Evaluation Corpus, which consists of three subcorpora: the Quality Corpus, the Quantity Corpus, and the Inappropriate Corpus. The Quality Corpus consists of texts that have been hand-picked primarily for their conceptual context; their purpose is the same as conventional parallel texts. The Quantity Corpus in combination with a corpus analysis tool, such as Wordsmith, is used to investigate the appropriateness of specific terms and to identify appropriate collocations, etc. Finally, the Inappropriate Corpus entails 'inappropriate' parallel texts, that is, texts which deal with the subject matter in question but which differ from the source text in criteria such as text type or publication date. It is reasonable to assume that Bowker's proposal implies an useful resource for helping trainers to identify and correct errors in student translations.

7. CONCLUSIONS

Drawing in these premises, it is argued that adopting a corpus-based approach significantly reduces the subjective element in translation quality assessment, although it has to be complemented with other resources concerning the evaluator's competence and critical judgement in order to reach accurate and reliable conclusions. In any case, it is clear that the approach must be corpus-based, not corpus-bound (Bowker 2001: 361).

It has been proved that the effectiveness of parallel and multilingual corpora varies accounting for the nature of translation. Techniques of text alignment and corpora tools seem to be especially significant in Machine Translation Evaluation and in the assessment of professional translators' work due to the characteristics of these texts. Segments and shifts can be easily identified and analysed by means of alignment techniques. Translation errors can be assessed by means of scales accounting for their seriousness. Even standards of quality can be directly applied to the processing of data. As this suggests, assessment criteria and parameters are rather precise and accurate; consequently, an objective and systematic approach can be reached.

That aside, parallel corpora appear to be essential in translation training so as to enhance students' translation competence, for students can handle a vast amount of information concerning translation resources and translation strategies. By examining segments, translators' errors can be analysed and contrasted. Different types of texts, periods and translators can be contrasted. Translation memory software proves especially significant as resources are available for trainers to advice students. In addition, students' translations can be assessed in a reliable objective way. In all cases both qualitative and quantitative approaches can be combined to increase the accuracy of results. Doubtless, an advanced competence in translation can be acquired or enhanced.

Although current advances in the research must be highlighted, important shortcomings can be appreciated in literary translation quality assessment. Not all texts are available in electronic format; building a parallel or multilingual corpus of one's own making can be required for a specific assessment. That aside, extratextual features of literary translations cannot be directly analysed. However, after having performed a manual exhaustive analysis of the central issues concerning both texts, corpus data and segments can be imported to especially designed databases in order to examine translation strategies or features of style. The advantages of using a parallel corpus in literary assessment are beyond any doubt. Alignment enables the critic to identify segments and classify them following the shifts which they involve. Despite the fact that a manual aid is required, frequency lists and statistic routines tackle data in an efficient form. Data can be processed in an accurate and non time-consuming way which is required before applying human judgement.

In any case, current and future research is hoped to enhance the advantages of corpora in translation quality assessment as their potential seems far from being fully explored.

NOTES

[1] See http://stp.ling.uu.se/~corpora/plug/pwa

[2] See http://www.lexically.net/worthsmith/version4/index.htm

[3] See http://www.ruf.rice.edu/~barlow/parac.html

[4] See http://www.nist.gov/speech/test/mt

[5] Disposable corpora, assembled from readily available sources have been suggested to be valuable learning resources on the translation and language classrooms despite the fact that designing and constructing them may at times require more time and effort than are available in the average teaching setting. However, the corpus building activity must be seen as a means in the pursuit of other aims. In addition, the corpus building procedure can be made faster by means, for example, of BootCat toolkit proposed by Silvia Bernardini and Marco Baroni (see http://sslmit.unibo.it/baroni/bootcat.html).

WORKS CITED

Abaitua, J., J. Díaz, J. Gómez, I. Jacob, K. Ocina. 2001. "XTRA-BI: Extracción automática de unidades de entidades bitextuales". *Procesamiento de la Lengua Natural*. 27: 205-306.

Aston, G. 2000. "Corpora and Language Teaching" in L. Bernard & T. McEnery, eds. *Rethinking Language Pedagogy from a Corpus Perspective*. Frankfurt: Peter Lang. 7-17.

Baker, M. 1995. "Corpora in Translation Studies: An Overview and some suggestions for future research". *Target* 7. 2: 223-243.

-------------2000. "Towards a Methodology for Investigating the Style of a Literary Translator". *Target*.12. 2: 241-266.

Barlow, M. 1995. *A Guide to Paraconc*. Houston: Athelstan.

Botley, S., A. M. McEnery and A. Wilson, eds. 2000. *Multilingual Corpora in Teaching and Research*. Rodopi: Amsterdam.

Bowker, L. 2000. "A Corpus-Based Approach to Evaluating Student Translations". *The Translator*. 6. 2: 183-210.

-------------- 2001. "Towards a Methodology for a Corpus-Based Approach to Translation Evaluation". *Meta*. 46.2: 345-364.

-------------- 2003. "Corpus-based applications for translator training: Exploring the possibilities". *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam and New York.

Bower, L. and J. Pearson. 2002. *Working with Specialized Language. A practical Guide to Using Corpora*. London: Routledge.

Brunette, L. 2000. "A Comparison of TQA Practices". *The Translator* 6. 2: 169-182.

Chesterman, A. 1998. *Contrastive Functional Analysis*. Amsterdam and Philadelphia: John Benjamins.

Delisle, J., H. Lee-Jahnke and M.C. Cormier, 1999. *Terminologie de la traduction. Translation Terminology. Terminología de la traducción. Terminologie der Übersetzung*. Amsterdam and Philadelphia: John Benjamins.

Halliday, M.A.K. 2001. "Towards a Theory of a good Translation", in E. Steiner and C. Yallop, eds. *Exploring Translation and Multi-lingual Text Production*. Berlin: Mourton De Gruyter: 13-18.

Kirshore, S. R., T. Ward, W. Zhu. 2001. "Bleu: A Method for Automatic Evaluation of Machine Translation". http://domino.watson.ibm.com/library/CyberDig.nsf/home

Lauscher, S. 2000. "Translation Quality Assessment. Where can theory and Practice meet?". *The Translator*. 6. 2: 149-168.

Leuven-Zwart, K. 1990. "Translation and Original: Similarities and Dissimilarities II". *Target* 2.1: 69-75.

Ling Koo, S. and H. Kinds. 2000. "A Quality-Assurance Model for Language Projects". *Translation into success. Cutting Edge Strategies for going with multilingual in a global Age*. R. C. Spring, ed. American Translators Association Scholarship Monographic Series. Vol. XI: 147-157.

Maczewski, J. "*Palimpsest* for Window: A Tool for Computer Assisted Literary Translation Studies". Poster Session. http://helmer.askis.uib.no/allcach96/Demos/maczewski.html

Martínez Melis, N. and A. Hurtado Albir. 2001. "Assessment in Translation Studies: Research Needs". *Meta* 46.2: 272-287.

Munday, J. 1998. "A Computer-assisted Approach to the Analysis of Translation Shifts". *Meta* XLIII, 4: 542-556.

Nieben, S., F. J. Och, G. Leusch, H. Ney. 2000. "An Evaluation Tool for Machine Translation: Fast Evaluation for MT Research". *Proceedings of the Second International Conference on Natural Resources and Evaluation*. Athens.

Schäffner, C., ed. 1988. *Translation and Quality*. Clevedon: Multilingual Matters.

Scott, M. 1999. *Wordsmith Tools*. Oxford: Oxford University Press.

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam and Philadelphia: John Benjamins.

Véronis, J. 2000. *Parallel Text Processing. Alignment and Use of Translation Corpora*.

Kluwer Academic Publishers: Dordrecht.

Zanettin, F., S. Bernardini and D. Stewart, eds. 2003. *Corpora in Translation Education*. St. Jerome: Manchester.