CHAPTER FIVE

SMALL PARALLEL CORPORA
IN AN ENGLISH-ARABIC TRANSLATION
CLASSROOM:
NO NEED TO REINVENT THE WHEEL
IN THE ERA OF GLOBALIZATION

HAMMOUDA BEN SALHI,
UNIVERSITY OF CARTHAGE (TUNISIA)

Introduction

At the start of the 21st century, the translation market is not only witnessing a boom, but also new requirements and changes. These are both imposed and motivated by the overwhelming phenomenon of globalization and the rapid and substantial development of information technologies (IT). New socio-economic conditions generated by globalization, advancement of global communications, new horizons opened up by the web as well as the frequent and critical social, technical, cultural and financial crises\(^1\) are having an impact on the language services profession, mainly on the translation business. Translation business, and translation training, are now called upon to satisfy the high demand for multilingual documentation and meetings, with a special demand for Arabic language as SL and TL.

Because there has been an urgent need for the translation industry to cope with these changes together with new market and globalization requirements such as high quality and speed, a number of novel tools and methods have been developed to serve the translators’ community. One of the main advantages brought about by technology is the easy and quick

access to electronic texts (Zanettin, 2002) as well as their availability on a large scale, whereas in the past only printed texts were available. Electronic texts can be easily retrieved, parsed, concordanced and aligned not only for translation research and practice, but also for translation teaching and learning. Therefore, corpora and concordancing software programs as well as easy and quick access to the different sources of knowledge such as terminology databases and the Internet have now become of critical importance in the translator computerized workstation. Every single professional translator is now required to be equipped with these tools and resources.

Comparable and parallel corpora, which were developed first to serve machine translation, are increasingly being used as reliable translation resources from which translators, translation scholars and trainee translators learn not only about the nature of translated language but also how to understand source texts, write fluent target texts, familiarize themselves with text genres and deal with collocations and idiomatic language use. The use of corpora has been a response to and an offshoot of these new requirements.

The above are some of the changes that the profession of translation has undergone as a response to the many new requirements the market and globalization are imposing. It is at this critical stage of progress in the translation market and profession that an upgrading of the translator training classroom should be performed in the Arab world, if we really want to keep pace with current and future needs in the translation market. In Western countries, and as a result of these changes, the use of parallel (or translation) corpora in translation classrooms has now become common practice amongst translation teachers and students. These rich resources of fresh data help make students familiar with the genres of texts to be translated and, thus, to enhance their understanding of source texts and improve their usage of target words (and terminology). In fact students are now confronted with copious amounts of both original texts and their associated translations\(^2\), which will make their translation tasks easier, save them time, make their assignments more interesting, and, most importantly, prevent them from reinventing the wheel.

The present paper will look into the challenges that translation students face now and in the future for these new changes and requirements not to increasingly frustrate them and, thus, not to hamper the efforts of educators/trainers aimed at alleviating the many different linguistic, pragmatic and cultural problems and challenges that still haunt their

translation assignments and widen the Academic and Professional Gap (herafter APG). Likewise, it will explore the opportunities that a collaborative approach\(^3\) could offer to students. More concretely, it will report on an experiment that was carried out at the Faculty of Arts and Humanities of Sousse with undergraduate students to demonstrate how small parallel (and comparable) corpora can be used to enhance the fluency and professionalism of trainee translators and facilitate teaching and learning processes on the basis of a collaborative approach to the translation exercise. Small English-Arabic parallel corpora and other small comparable corpora taken, mainly, from United Nations texts, which can be regarded as reliable in an era where reliability has become questionable, are presented as an example of translation corpora used in the translation classroom. The findings showed an improvement in students’ translation skills and translation output.

**Translation and Technology: the New Frontier**

In the past, translation was basically hand-written and the translation of phrases, sentences and texts hinged on the creative skills an individual professional translator might be endowed with, but such creativity together with the data that such a translator could collect have always been “limited to what a single individual could experience and remember”\(^4\). With technology, a number of human shortcomings could be overcome, such as the very limited capacity of human memory and, obviously, limited skills of the one against those of the many, where each translator can profit from the skills of other translators, together with automated translation memories. Technology, therefore, can be described as the new frontier of the current translation industry.

**Advantages brought about by Technology**

The relation between translation and technology can be divided into three main phases. In the first phase there is simple or no use of technological tools where pens, papers and printers were the only technology used in

---


that craft. The second is the phase where technology is used as a mere support to the translator and the third and current phase where technology is considered as an integral part of translation practice. One of the key factors behind these transitions is the internet. This research tool offers unparalleled advantages to translators, translation researchers, translation teachers and trainee translators. Chief among these advantages are the increasing availability of information, a wealth of data now emerging from past translations, less dependence on local sources of information, enabling worldwide human collaboration in an unprecedented degree.

A translator, according to Al-Jahidh⁵, needs to have good command of idiomatic language, creative expressions and syntax as well as the customs and traditions of SL and TL speakers and their means to achieve mutual understanding and interchangeability. In addition, he has to have an encyclopedic knowledge, which allows him to understand whatever information, whether emanating from mathematics, business, philosophy, chemistry, or any other field of knowledge. It is thanks to a tool like the internet that much of these requirements and obstacles can be overcome today. With a simple click, the user can get answers to all of his questions accompanied with illustrative explanations. These illustrations may comprise written texts, sounds, pictures, or even movies and documentaries. The internet, therefore, has created the basis for cooperation and collaboration between translators on the one hand, and experts, philosophers, native speakers, and other translators, on the other. For instance, some web sites, such as the search engine Google, are launched online for the simple sake of facilitating such collaboration and exchange of information between people worldwide.

Other technological tools are also of paramount importance in the translation workstation and have brought other advantages which complemented the aforementioned resources, especially those which are required in the corpus analysis process. The main tools are:

a. Translation memory systems: they are computer applications that allow translators or any other user to store past translations associated with their originals in a database, so that when they are needed in new translation projects they can be easily retrieved and re-used. These systems work with other tools such as the concordance and aligner programs;

---

b. **Terminology databases:** Terminology databases are technological pillars of modern translation. Terminologists, translators, and linguists as well as experts collaborate together to establish reliable databases containing all the terminology used in a specific field. Such a database will be useful for future users and help the fluidity of the technical terms and the lexicon in general be understood. The most famous and reliable terminology databases are those established by the United Nations in a good number of domains. These help unify terminology usages inside the UN system. They contain novel terms and are updated regularly. But, unfortunately, Arabic translators, terminologists and linguists have not sufficiently profited from these databases. A situation worsened by the overwhelming phenomenon of technology illiteracy across the Arab world.

c. **Electronic corpora:** corpora, whether they are comparable or parallel, consist of original or translated texts that serve as useful resources for translators, main corpora include:
  - British National Corpus (BNC) - 300 million words on-line /100 million on CD-ROM
  - International Corpus of English (ICE)- 1 million fully parsed words
  - International Corpus of Learners’ English (ICLE)
  - European Language Resources Association (ELRA) - collection of corpora in many languages
  - ICAME - CD-ROM of several corpora

d. **Concordancers:** “a concordancer, which finds and displays, in an easy-to-read format (e.g. KWIC — key word in context) all occurrences of a search term (and minor variations thereof).” This tool is especially useful when it comes to the translation process as it helps show all the translational options available in a parallel corpus.

e. **Alignment tools:** they are programs which are used for automated alignment of words or sentences in a restrictive or ideal domain and context.

---


New Changes and Requirements in the Profession

The translation industry and the profession of translators and interpreters have been deeply transformed by technology and the internet. Now computational tools and information resources, internet-based marketing strategies, and e-mail are turning the business into a tech-intensive globalized industry that both favors and promotes technology know-how and impels traditional translators, who might be very competent sometimes, to take a back seat in a rapidly changing market. The transformation is basically generated by the high demand for translation and interpreting services.  

In fact, the changes in the size, nature and type of the demand have had a major impact on the translation profession, which has become more industrialized. A number of researchers have addressed these changes and provided some concrete recommendations on how to better keep pace with the new changes and requirements that the translation market imposes.  

Two of the main changes and requirements in the translation industry are the use of large comparable and parallel corpora as rich resources of translational knowledge and the resort to the Virtual Meeting Collaboration (VMC) as a powerful collaboration tool.  

A parallel corpus, for instance, is now used to serve as a bilingual lexis resource. By means of concordancers and aligner software programmes, terminological and lexical equivalents can be easily and quickly retrieved. The advantage of parallel corpora over bilingual dictionaries is that they provide equivalents in real contexts. Verbal, cognitive and situational contexts are all there in a parallel corpus. Therefore, the use and building of such corpora, in addition to translation memories, terminology databases, and related computational tools and systems have become an integral part of the work of the 21st century translator.  

Also, in order to be more competitive in the translation market, whose businesses and transactions are increasingly taking place virtually on the web, a modern translator is now required to resort to VMC. Requirements pertaining to the speed of delivery and the quality of the output can be met through this tool. Virtual meetings of professional translators with

---

different but complementary language pairs abound now. We can mention both the (http://www.translatorscafe.com) and the (http://www.proz.com), which are virtual forums designed to serve translators. They form spaces where translators, interpreters, translation companies and end clients can meet and work efficiently and profitably. Major advantages of web sites like these include:

- **Help**: any professional or novice translator can give or receive help in translating tough technical terms.
- **Conferences and training**: any user can attend online (or offline) conferences or training sessions making the community of translators closer to each other. These events are designed to disseminate new translational skills and know-how.
- **Discussion**: in these spaces, translators can discuss issues of immediate interest to translators like localization, subtitling and use of special computational tools.

These are some of the new transformations and requirements that trainee translators have to be aware of in order to acquire the necessary tools to meet them. Let us see now in more detail the characteristics of the workstation of the current professional translator.

### The Professional Translator Workstation

It has become a fashion to resort to the Massive Online Collaboration (MOC) in the new computerized translator workstation. MOC is now changing the rules of the game. Professional translators are increasingly marking their presence in the virtual world of the internet for the many advantages it offers. The final output of the translation process is no more the product of a single (individual) translator and is never restricted to the translation task, in the narrow sense of the word. Most of the translation assignments are carried out collectively now as they undergo some proofreading by specialized proofreaders or experts, some editing, reviewing and revising by specialized skillful revisers, which are all services provided now online either by individual freelancers or agencies.

Didawi\(^{11}\), for example, puts an emphasis on the critical importance of the revision step and shows how it is necessary to enhance the translation

---

output. He provides many examples to illustrate that, they include the following:

✓ “Violations of sexual integrity”
Before revision: الانتهاكات ضد الحرمة الجنسية
After revision: انتهاك الحرمة الجنسية

The phrase sounds more Arabic after revision.

✓ “Military armed cars arrived at his home with over 50 men”
Before revision: جاءت سيارات مدرعة إلى بيته يرافقها أكثر من 50 رجلا
After revision: وصلت إلى بيته سيارات (مصفحة) وعلى متنها أكثر (ما يربو على) 50 نفر.

Before revision, the translator used a false collocation in Arabic. He made the word سيارات (cars) collocate with يرافقها (accompanied by), while in Arabic سيارات (cars) collocates rather with على متنها, a prepositional phrase used when one takes some means of transport such as cars, trains and plains.

Through the internet, online consultation of experts in every specific technical domain is also possible in order to make translation output of scientific and technical texts scientifically accurate. Such a consultation is sometimes recommended as linguistic differences between languages might have their impact on the accuracy of the scientific information.

**State of the Translation Classroom in Tunisia**

Though it is one of the most educative and comprehensive subjects, translation is still not accorded its worthy status by Tunisian academic institutions and students still hold bad misconceptions about translation assignments and regard the subject as one of the bêtes noires.

Most of the translation teachers are not specialized in translation studies and a good number of them have never practiced translation out of the classroom environment. Most departments do not set their own objectives and lack well-planned translation teaching materials. Teachers still follow an out-of-date and unsuitable method that is more likely to generate “disabled” students who will find it very difficult to access the language (or translation) market if the doors of secondary school teaching are still closed in front of them. A typical method adopted by most translation teachers is that they photocopy whatever article or any short excerpt of a literary text and hand it out to students as a work assignment for the next lesson. Lacking the necessary motivation and enslaved by those misconceptions, students are always reluctant to do such assignments
and the following lesson the teacher finds himself working alone or, in best cases, with some two or three motivated students. As soon as they obtain their diplomas, students start to suffer from the big problem the Academic and Professional Gap (APG).

The Academic and Professional Gap (APG) in Translator Training in Tunisia

The majority of students still have major and inherent problems in the translating process. They cannot grasp the fact that translating involves more than just replacing a word or a phrase with its equivalent in another language. They are still unable to deal with sentences, ideas and perhaps the author’s intentions so as to reflect the same coherence found in the source text to come up with a target text that reads as if it originated in the target language. Though teachers always advice students not to make words their ultimate worry, they face major problems in applying pieces of advice on the ground. Market demands and requirements must necessarily find their way into the translation classroom; otherwise the serious APG will continue to generate “disabled” translators.

The APG is also the outcome of technology illiteracy in the field of translation. Translation students use little or no technology in their training. Bowker\textsuperscript{12} found that 60.5\% of the translation-related job advertisements require computer literacy. Employers who advertise for translation-related positions in Canada seem to be relatively in tune with the skills required to work as a translator in the 21st century. Computer skills – both general and specialized – are now considered to be an integral part of the profile of the 21st century translator. Tunisian universities, however, are still undermining the importance of computers and the internet in the translation classroom. Access to the internet and other technological tools is available in many other fields in Tunisia but not in the classroom.

In fact, technology should not be taken as an obstacle or a challenge but rather as an opportunity available for translators and especially for translation students. Today, nearly all translation work is done on a computer, and most assignments are received and submitted electronically. This enables translators to work from almost anywhere, and a big proportion of the translation assignments are carried out at home. The

internet provides advanced research capabilities, valuable language resources and access to reliable corpora, specialized dictionaries and glossaries. In some cases, memory tools, which provide comparisons of past translations with current work, help save time and reduce repetition. Therefore if we really want to meet the market needs, ICTs must occupy the right place in translation classroom.13

The Experiment

It all started in the academic year of 2005/2006 with the launching of a web site (http://www.freewebs.com/hsalhi) designed to facilitate communication between teachers and students, extend the teaching and learning processes beyond class hours and attract students to partake in the very real pleasure that the world of translation, linguistics and the web provides. Therefore the first work was to motivate students and help them fault those misconceptions about the translation exercise. Theoretical lectures are delivered to all undergraduate students of 3rd year on the importance of the contextual cues and the corpus-based approach. These lectures are complemented by regular applied ones wherein students have the opportunity to translate, from and into Arabic, textual documents representing domains of current concern and in line with subjects dealt with by the different United Nations agencies. In fact, students are provided with a manual containing all the translation assignments along with small comparable or parallel texts aimed at assisting students in their endeavors to overcome linguistic, pragmatic and cultural difficulties they may encounter in these assignments. In addition, a collaborative approach has been introduced after a good proportion of students have shown a clear interest in translation.

With the small comparable and parallel texts included in the manual, students are also encouraged to collect their own small corpora from United Nations texts basically and then build their own glossaries of original words, collocations, and technical terms along with their associated equivalents and translations. Some students were even motivated to regularly consult the web on how target words, collocations and idiomatic expressions are properly used. Working in small groups, students could share their findings with their colleagues and the translation classroom has become rather a student-centered lesson.

Assessment of the Students’ Translational Skills

To assess the progress students achieved in acquiring new translational skills thanks to corpora and the collaborative approach adopted, we made a comparison between the grades of students in the final exams of both academic years of 2005/2006 and 2007/2008 with one academic year in between. Though mindful of the fact that exam grades do not really reflect the real competence of students, I will extensively rely on them to make the necessary comparison. In fact grades can be taken as a reliable indicator of their level if exams are well prepared for, several coordination meetings have taken place between the graders before the actual grading work takes place and a careful and comprehensive grading scheme is adopted.

Coordination Meetings

In addition to the informal meetings that translation teachers have held, three formal coordination meetings have been organized.

- **A meeting before courses start**: teachers have met before the courses start off to decide on the texts to be included in the manual and set the teaching objectives and identify the up-to-date needs of students.
- **A second meeting before the exam**: teachers have met some two months or so before the date of the exam to select the texts to be translated and formulate the exam questions.
- **A final meeting after the exam**: teachers take a random sample of exam copies to jointly grade them following the grading scheme that has been adopted in order to create some harmony on the given marks.

Grading Scheme

The exam consists of two parts: a short text to be translated on 16 and a commentary of the translation of some specific spots in that text on 4. As for the first part, we have adapted the correction system applied in the United Nations Competitive Examination for Arabic Translators with some changes introduced. Marks have been given according to the following scale:
➢ **Understanding**: errors pertaining to problems of understanding are marked by an (x) mark to the right and 5 points out of a total of 100 points are deducted for each single mistake.

➢ **Language**: language mistakes are marked by the letter (L) and punished by deducting 4 points out of a total of 100 points for each single mistake.

➢ **Wording, expressions, and structures**: mistakes of this kind are marked by an (-) mark. These are punished by deducting 1 point out of a total of 100 points for each single mistake.

➢ **Omission**: any omission or missing part having a serious impact on the overall meaning of the target text is punished by 5 points.

This mistake-based grading is complemented by another scale of positive marks. A number of spots in the SL text have been determined in advance as difficult spots. These spots might be difficult words, idiomatic expressions, collocations, structures of sentences. A (+) mark is awarded to the student who overcomes an obstacle or excels in using collocations, idioms and so on. Any (+) is worth 3 points awarded to the student in question. Students are also asked to comment on four underlined excerpts in the texts in a separate question for 4 points. A provisional score of points is calculated on the basis of positive and negative points:

**Provisional score** = total of positive points – total of negative points.

After the third coordination meeting, teachers decide on which provisional score should be the pass mark - worth 8 out of the remaining 16 points.

## Quantitative Results

By comparing results of both academic years, we found that mean marks are as follows:

- **2005/6**: the mean mark of a total of 501 students is 8.11 out of 20
- **2007/8**: the mean mark of a total of 439 students is 9.82 out of 20

Though there is a clear progress of the mean mark but still the real progress is even better, as the number of absent students who got zeros, is bigger during the second academic year.

- **2005/6**: out of the 501 students, 9 did not sit for the exam and got zero
2007/8: out of the 439 students, 28 did not sit for the exam and got zero

The below figure shows that the results of the 2007/8 academic year, represented by red lines, is generally better than the other year, represented by blue lines. The red lines are in most cases higher than the blue ones even though the 28 zeros are taken into account. Only students in the 2007/8 academic year could reach the mark area above 13.8 with the highest mark of 16 out of 20.

Source: Computer Science Division of the Faculty of Arts and Humanities of Sousse
Qualitative Results

Qualitative results can be summed up in the following developments:

- **Motivation**: by the end of the three-year experiment, students started to express their positive attitudes about the utility of the translation exercises and that some of the serious problems are now manageable thanks to corpora and the clear set objectives and the available manual. These attitudes helped raise the degree of motivation among students.
- **The bête noire misconception**: A good proportion of students became aware that the “bête noire” attitude is a fallacy.
- **Attendance**: there is an increasing willingness to attend translation classes.

Although the qualitative and quantitative results are generally positive, they hide some other facts and weaknesses pertaining to some other students and to the performance of the experiment itself. Still a good number of students are not motivated enough and still hold the view that translation is a difficult subject and can never excel in it. The experiment relied on exam marks which, as I said earlier, are not a good indicator of the competence and skills of a student. Also, technology which is highly required by the future translation market is not heavily used and not well introduced to students, especially concordancers and aligner software programs. This is in fact hindered by a global academic and social environment that does not really foster the introduction of computational tools like these. In spite of these shortcomings the experiment allows us to make some concrete recommendations to students and teachers alike.

Recommendations and Conclusions

As for students, they are called upon to try to have a minimum of formal training in computer science. It is high time now they took the initiative to access online virtual forums, such as translatorscafe.com, and to create their own forums. Teachers must be aware that students now need less education and more training. They are called upon to design courses which respond to changing conditions in the world of professional translation.

This article has demonstrated how small parallel (and comparable) corpora and a collaborative approach can be used to motivate students and therefore enhance their fluency and professionalism. But, as a final
message, I can say that corpora, technology and collaboration are enough to attract students to the world of translation and save them from reinventing the wheel. Alone, however, they are not enough to turn themselves into good and professional translators.