QUALITY TRANSLATIONS WITH THE HELP OF TERM BASES

Imre Attila

Assoc. Prof., PhD ", Sapientia" University of Tîrgu Mureș

Abstract: The present article focuses on the possibility of quality translations with the help of term bases. As present day computer-assisted translation tools make extensive use of term bases and translation memories, professional translators with high quality specialized term bases have the leverage in producing time- and cost-effective specialized translations. Dictionary entries are compared to online term bases regarding Romanian–English and English–Romanian terms regarding both quality and quantity.

Keywords: term bases, dictionaries, quality, quantity, CAT-tools.

Introduction

Since the (r)evolution of technology, the (r)evolution of translation is connected to computers and software (Imre, 2013, pp. 102–174), and with the passing of time this is more and more obvious. There are many studies discussing the competences of a translator, and – preserving the 'classical' ones, such as linguistic or cultural – they consider that computer skills and technical know-how in the field of translation is the guarantee of remaining competitive on the translator's market.

Before the technical revolution a person dealing with translations could have been referred to as a translator, but ever since the term itself underwent changes, as machine translation (MT) or a person using MT may be called a 'translator'. As a consequence, everybody may be regarded as a translator once we are connected online and can feed a text to be translated into an MT software. The outcome may be called 'translation', disregarding the debates connected to it: the importance of pre- and post-editing (Bowker, 2002, p. 4; Allen, 2003, pp. 297–300) or its quality.

Although the issue of quantity may be considered as solved this way, the second aspect of translation, quality, needs discussion. The ALPAC report in 1966 (Hutchins, 2003)

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

was rather bleak about the use of MT, which delayed the development of the industry. However, since the appearance of computer-assisted translation (CAT) tools, the situation changed. The triangle of translation would have three vertices, MT, CAT and the human translator (HT), which were each trying to demonstrate their own supremacy over the other two. Yet, in the long run, this 'brawl' did not prove to be beneficial, so CAT tool vendors began to highlight an important feature of their products, namely 'aid' or 'assistance', and a giant leap in developing their products was to incorporate MT and offer it as a plugin that can be enabled or disabled during the translation process. The proof of the success is demonstrated by the fact that while in 2008 CAT tools had no built-in MT, year-by-year they began to build in various MT options. For instance, the leading CAT tool worldwide, SDL Trados Studio, now offers SDL Language Cloud Machine Translation[1], whereas memoQ translation environment has eleven MT plugins: Crosslang Gateway MT, Google MT, IP Translator MT, iTranslate4.eu MT, KantanMT API Connector, Let's MT, Microsoft MT, Pangea MT, memoQ pseudo-translation, Systran MT, tauyou machine translation, and all of them can be further configured (e.g. language pairs), set as preferred or enabled or not:



Figure 1. memoQ and MT options

Interestingly, memoQ had no MT plugin in 2008, but Kilgray Company started to add the best MT possibilities to their software. They state that memoQ "does not have its own

[1] http://www.translationzone.com/products/sdl-trados-studio/professional/industry-engines.html, 29.11.2015.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

machine translation engine but it is connected to the main engines through plugins"[2]. To sum up, the following illustrations (Imre, 2014a, p. 68) may help:

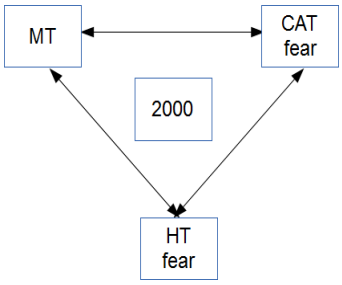


Figure 2. MT, CAT and HT in 2000

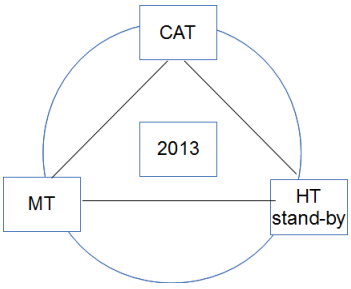


Figure 3. MT, CAT and HT today

Figure 2 and 3 reflect the fear of CAT and HT manifested towards MT at the turn of the century. However, today the majority of professional translators understood that both MT and CAT are here to stay and they should turn them to their advantage. Gouadec stated as early as 2007 that the age of pencil and paper assisted translation (PRAT) is to be replaced with CAT (2007, p. 109) and those who were able to switch made the first step towards professional translation. And this is where globalization enters the picture: McLuhan's 'global village' brought about changes in dealing with customers, focusing on 'here and now' once the customer pays for it, often to the detriment of quality. Translation vendors realized soon enough that the ever growing demands on a competitive market would result in a reduced timespan dedicated to text production and preserving previous translations is the key for that, leading us to the next section.

Quality Assurance and CAT tools

The basic idea of CAT tools is to create a modern framework for professional translations; translators using these software create translation projects, together with translation memories (TM) and term bases (TB). TMs will preserve the entire translated text divided in pairs of source and target segments (basically sentences), while TBs all the pairs of terms in the project. Both TMs and TBs can be:

- Created at the beginning of the project; in fact, this is a must if we wish to reuse any part of the project later (the essence of CAT tools);
- Imported from sources outside the project (the Internet, fellow translators, own previous translations);

[2] https://www.memoq.com/integration-with-machine-translation-systems, 29.11.2015.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

- Modified during the project in more ways: new entries added or previous ones completed, corrected or deleted, thus assuring quality (e.g. the same term is translated consistently throughout the entire project).

Although there were many who were reluctant to accept that CAT tools may offer both reduced work or editing time and enhanced quality, today this is not an issue any more. Furthermore, the Run QA function of memoQ will make use of the Spelling and grammar settings, which makes it possible to use a spell-checker while typing or when proofreading a translation:

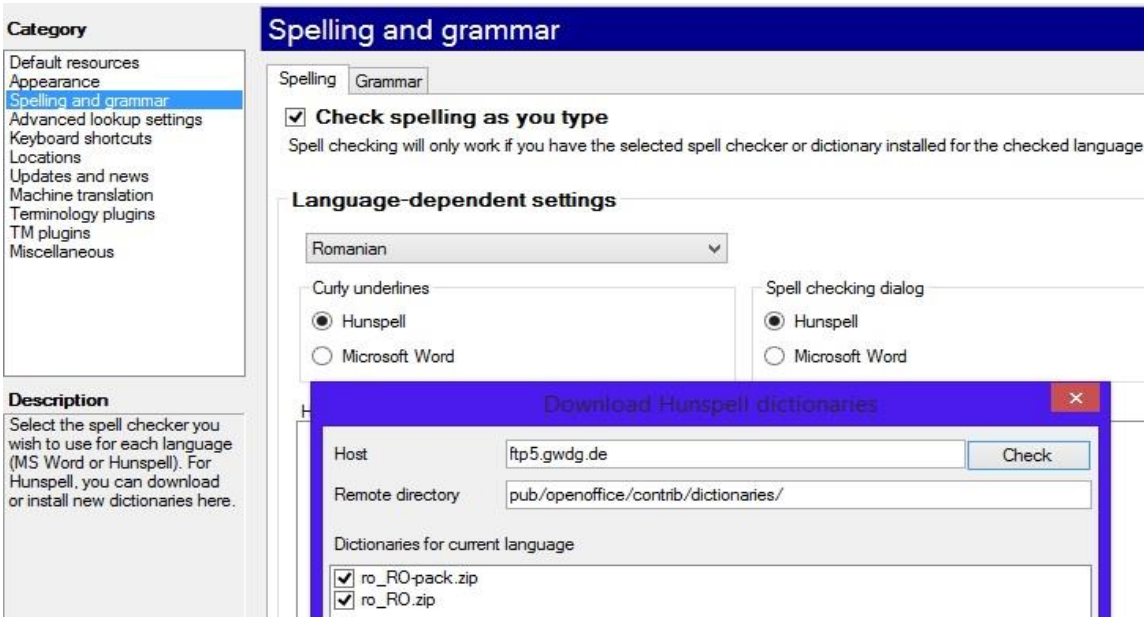


Figure 4. Spelling and grammar in memoQ

As a result, we strongly believe that a proper selection and use of the target language options will contribute significantly to both a much better quality and reduced working time, boosting productivity.

Term base quality of dictionaries

While carrying out various translations, it came to our attention that there were complaints about the poor quality of Romanian–English legal dictionaries.[3] We consider that in a globalized world and unified Europe there is an ever growing need for high quality multilingual term bases, such as IATE.[4]

[3] http://www.proz.com/forum/romanian/66770-dictionar_juridic_roman_englez.html, 19.08.2013.

[4] http://iate.europa.eu/, 30.11.2015.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

A post-doctorate project[5] (POSDRU/159/1.5/S/133652) enabled us to collect the majority of available Romanian–English legal dictionaries, trying to create a term base for CAT tools such as SDL Trados or memoQ. After having checked around 330,000 entries more than a dozen dictionaries, we could identify around 100,000 non-repetitive entries, although many of them belong to non-legal fields, such as geography, history, IT, economics, etc.

Our presupposition was that printed dictionaries would offer a rich and high quality data over distrustful online dictionaries or term bases available free of charge. However, during a period of eighteen months we slowly but surely changed our mind, at least partially. First of all, many printed dictionaries seem to reuse the entries from previously published dictionaries, not mentioning the unimaginably many errors regarding spelling, grammar, layout, translation, not suitable entries, etc. (e.g. Imre, 2014b). Being able to select from different printed dictionaries requires a considerable effort (time, money, energy), which proved to be less rewarding than expected.

In the meantime we started to check various entries in Romanian–English databases and dictionaries available online, and our attention switched more and more towards them. However, instead of drawing far-reaching consequences, we tested a few relevant entries. The present article will bring new results regarding the possible translations of an important legal term. As we have in mind to offer an (improved) English and Hungarian translation of the new Romanian Penal Code (Noul Cod penal; Noul Cod de procedură penală, 2014, p. 15), we checked it for recurrent terms, such as infracțiune, fraudă, falsificare, pedeapsă, or tentativă. As we have limited space here, we will analyze the results regarding the English translations of the Romanian term tentativă and all the other expressions containing this term, which is clearly defined and described in the New Romanian Penal Code, Articles 32 and 33.

Fourteen printed dictionaries containing legal and economic terms were searched for this entry, out of which none of the economic dictionaries contained the term (Ionescu-Cruțan, 2006; Năstăsescu, 2009; Imre, Butiurcă, Druță, & Bakos, 2014), whereas only one legal dictionary (Voiculescu, 2008) did not contain this term at all. The following nine legal dictionaries and one general dictionary contained tentativă:

[5] http://www.postdoc.commscie.uaic.ro/index.php/component/content/article/2-uncategorised-uncategorised/68-burse-commscie-2014, 30.11.2015.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

1. (Lozinschi, 2008, p. 602)
2. (Hanga & Calciu, 2009, p. 159)
3. (Lister & Veth, 2010, p. 571)
4. (Dumitrescu, 2009, p. 221)
5. (Botezat, 2011, p. 157)
6. (Voroniuc, 2011, p. 357)
7. (Voroniuc, 1999, p. 107)
8. (Jidovu & Nitu, 2010, p. 198)
9. (Mezei, 2006, p. 695)
10. (Nedelcu, Murar, Bratu, & Bantaș, 2000, pp. 974–975)

plus one more for general use contained, detailed below:

RO	**EN**	**Source: pg.**
1. tentativă	attempt	1:602; 2:159; 3:571; 4:221; 5:157; 6:357; 7:107; 9:695; 10:974-975
2. tentativă	attempted delict	9:695;
3. tentativă	bid	10:974-975;
4. tentativă	endeavor US	10:974-975;
5. tentativă	endeavour	10:974-975;
6. tentativă	trial	5:157; 10:974-975
7. tentativă de asasinat	attempted murder	10:974-975;
8. tentativă de atac	attempted assault	1:602;
9. tentativă de căsătorie	attempted marriage	1:602;
10. tentativă de contact	approach theory	6:357; 7:107
11. tentativă de crimă	attempt to murder	4:221;
12. tentativă de crimă	attempted felony	4:221;
13. tentativă de evadare	attempt of escape	4:221;
14. tentativă de evadare	attempt to escape	10:974-975;
15. tentativă de evadare	escape attempt	3:571;
16. tentativă de evadare	escape bid	10:974-975;

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

17. tentativă de fraudă vamală	attempted customs fraud	8:198;
18. tentativă de furt	attempted theft	3:571;
19. tentativă de furt	theft attempt	3:571;
20. tentativă de mituire	attempted bribery	3:571;
21. tentativă de monopolizare	attempted monopolization	1:602;
22. tentativă de omor	assassination attempt	3:571;
23. tentativă de omor	attempted murder	3:571; 4:221
24. tentativă de sinucidere	attempted suicide	1:602;
25. tentativă de trecere a frontierei	att(e)mpted[6] border crossing	8:198;
26. tentativă de trecere ilegală a frontierei de stat	attempted illegal state border crossing	8:198;

Table 1. Translations for tentativă in printed dictionaries

A rough analysis shows that 26 Romanian entries were found, 15 Romanian non-repetitive terms resulting in 22 non-repetitive English terms, which is a very good result, covering various types of attempts, such as escape, murder, or theft. Yet, we would like to compare these findings with available online Romanian–English sources, requiring no financial investment and the least possible effort.

Term base quality of online Romanian–English term bases

At present there are a few bi- or multilingual online dictionaries and term bases[7], and we also resorted to ten online sources:

1. http://hallo.ro;
2. https://glosbe.com/ro/en;
3. http://ro.linguee.com;
4. http://www.dictionarromanenglez.ro/;
6. https://www.dictio.ro;
7. http://www.eudict.com;
8. http://ro-en.gsp.ro/;
9. http://ro.bab.la/dictionar/romana-

[6] The dictionary contained a misprint: *attmpted instead of attempted.

[7] For the present research it is not really relevant whether the source is a dictionary or a term bases, as we are primary interested in the number of translations rather than their grammatical features.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

5. http://memorg.com/dictionary; engleza;

10. http://www.wordreference.com/roen;

The initial findings contained 38 entries, but for obvious reasons we excluded terms not beginning with tentativă, and terms not belonging to a standard legal terminology. The following 17 terms remained, out of which there are 10 non-repetitive Romanian entries:

RO	EN	Source
1. tentativă	attempt	hallo, glosbe, ro-en.gsp, dictionarenglez.ro, Eudict, ro.bab.la, Linguee[8], memorg, dictionarromanenglez, dictio, WordReference
2. tentativă	bid	hallo, memorg
3. tentativă	effort	hallo, glosbe, memorg
4. tentativă	endeavor	glosbe, Eudict
5. tentativă	endeavour	glosbe, Eudict, ro.bab.la, memorg, dictio
6. tentativă	tentative	hallo, memorg
7. tentativă	try	hallo, glosbe, memorg
8. tentativă de asasinat	assassination attempt	Linguee
9. tentativă de asasinat	attempted assassination	Linguee
10. tentativă de fraudă	attempted fraud	Linguee
11. tentativă de omucidere	homicide attempt	Linguee

[8] Linguee uses of the database provided by the European Parliament, accessible at http://www.europarl.europa.eu/.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

12. tentativă a unui atentator sinucigaş	attempt by a suicide bomber	Linguee
13. tentativă comunitară	community attempt	Linguee
14. tentativă de a transfera sarcina primirii solicitanților de azil	temptation to transfer responsibility for welcoming asylum seekers	Linguee
15. tentativă de aplicare a forței	attempted use of force	Linguee
16. tentativă de manipulare	attempted manipulation	Linguee
17. tentativă de minimizare sau de negare a genocidului	attempt to minimise or deny the genocide	Linguee

Table 2. Online translations for tentativă

One can easily observe that there are similar elements in Table 1 and 2, and it may prove difficult to jump into conclusions regarding their compared quality. Yet one thing is sure: ten years ago the dictionaries could have won easily, but today the mushrooming of multilingual online sources may prove more and more efficient. If we take into consideration the invested time, money and effort in creating a Microsoft Excel file from online sources (around two hours, not counting the fact that we had already known these Romanian–English sources before the testing), we can say that we obtained quite a remarkable result and we are more than satisfied with it.

Conclusions

At this stage it is obvious that both printed dictionaries and online sources have their advantages and disadvantages. Although our initial presupposition that printed dictionaries offer quality beyond any doubt proved to be false, we cannot deny the feeling of comfort and happiness when an entry is found in these dictionaries.

An extra advantage is that once they contain legal entries, one can hardly find terms belonging to sub-standard language (slang, taboo, swear words, etc.). As a result, a TB deriving from these sources will surely offer valuable matches in case annoying errors are eliminated (especially typographical ones). Yet this TB has a high financial and temporal price.

Online sources offer the comfort of quick access to a large database of entries, but this also comes with a price: entries are not filtered and non-legal terms contain more than 50% of the matches, taking extra post-editing time. If we add the unreliability of the majority of these

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

sources, no wonder that many translators question their use. But if we take a closer look, (online) electronic TBs have a huge advantage in the hands of a professional translator specialized in a particular field.

While printed dictionaries cannot be further improved (no possibility to add, correct, delete, modify entries), TBs can be modified any later time, partially due to a (hopefully) ever-growing expertise of the user and partially due to the changes in language (cf. new entries) and technology. For instance, memoQ could not handle files with .xls or .xlsx extensions when one wanted to import a file to a TB three years ago. As a result, one had to convert these files to .csv format to enable the import, but when this happened, due to encoding standards (basically UTF-7 or UTF-8) at least two Romanian letters with diacritical mark (ș and ț) damaged and it was a strenuous activity to restore them. Today there is the possibility of a direct import from an Excel file (.xlsx, .xls) and the problem of diacritical marks ceased to be.

In our case, it is clear that the combination of both major sources (Table 1 and 2) would be the most felicitous, even if it requires a certain effort in the short run to select the entries one may need. While at present CAT tools make extensive use of TBs and TMs, professional translators with high quality specialized term bases have the leverage in producing time- and cost-effective specialized translations.

The future may still hold a few surprises. One of them could be the acceptance of the fact that language belongs to everybody, or all spoken languages belong to the human race and their preservation in high quality TBs and TMs for the benefit of humanity will be of utmost importance. Although some may say that this would mean the end of human translators, similar allegations have always existed either in sci-fi movies or scientific papers (cf. Imre, 2013, pp. 223–229), none of them proving true even in the age of Google Translate.

Acknowledgements

The research presented in this paper was supported by the European Social Fund under the responsibility of the Managing Authority for the Sectoral Operational Programme for Human Resources Development (Sistem integrat de îmbunătăţire a calităţii cercetării doctorale şi postdoctorale din România şi de promovare a rolului ştiinţei în societate), as part of the grant POSDRU/159/1.5/S/133652.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

REFERENCES:

Allen, J. (2003). Chapter 16 Post-editing. In H. L. Somers (Ed.), Computers and Translation: A Translator's Guide (pp. 297–317). John Benjamins Publishing Company.

Botezat, O. (2011). Dicționar juridic român-englez / englez-român (2nd ed.). Bucureşti: C.H. Beck.

Bowker, L. (2002). Computer-Aided Translation Technology: A Practical Introduction. University of Ottawa Press.

Dumitrescu, D. (2009). Dicționar juridic român-englez. Bucureşti: Akademos Art.

Gouadec, D. (2007). Translation as a Profession. John Benjamins Publishing.

Hanga, V., & Calciu, R. (2009). Dicționar juridic englez-român și român-englez. Bucureşti: Lumina Lex.

Hutchins, J. (2003). ALPAC: The (In)Famous Report. In S. Nirenburg, H. L. Somers, & Y. A. Wilks (Eds.), Readings in Machine Translation (A Bradford Book, pp. 131–136). Cambridge, MA; London, England: The MIT Press.

Imre, A. (2013). Traps of Translation. Brașov: Editura Universității "Transilvania."

Imre, A. (2014a). Fordítások menedzsmentje vagy fordításmenedzsment. In V. Bárdosi (Ed.), Szövegalkotó gyakorlatok, nyelvteremtő praktikák (pp. 119–132). Budapest: Tinta Könyvkiadó.

Imre, A. (2014b). Ways to enhance legal dictionaries. Communication, Context, Interdisciplinarity, (3), 519–527.

Imre, A., Butiurcă, D., Druță, I., & Bakos, L. (Eds.). (2014). Dicționar de termeni economici II. (Vol. 2). București: Editura Universitară.

Ionescu-Cruțan, N. (2006). Dicționar economic englez-român, român-englez. Bucureşti: Teora.

Jidovu, I., & Nitu, A. (2010). Mic dicționar terminologic pentru domeniul Schengen. Bucureşti: Universul Juridic.

Lister, R., & Veth, K. (2010). Dicționar juridic englez-român / român-englez. (R. Dinulescu, Trans.). Bucureşti: Niculescu.

Lozinschi, S. (2008). Dicționar juridic Român - Englez. Bucureşti: Editura Smaranda.

Mezei, J. (2006). Magyar-román-angol jogi, közgazdasági és üzleti szótár. Bucureşti: C.H. Beck.

Năstăsescu, V. (2009). Dicționar economic englez-român, român-englez. Bucureşti: Niculescu.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)

Nedelcu, C., Murar, I., Bratu, A., & Bantaș, A. (2000). Dicționar român-englez. Bucureşti: Teora.

Noul Cod penal; Noul Cod de procedură penală. (2014) (4th ed.). Bucureşti: Editura Hamangiu.

Voiculescu, C. (2008). Dicționar juridic englez-român / român-englez - terminologie UE/SUA. Bucureşti: Niculescu.

Voroniuc, A. (1999). Dicționar de termeni economici și juridici (român-englez). Iași: Institutul European.

Voroniuc, A. (2011). Dicționar englez-român / român-englez de termeni economici și juridici. Iași: Polirom.

BDD-V2012 © 2015 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.34 (2026-11-13 23:59:27 UTC)