The volume Specialized Languages and Conceptualization includes in-depth studies on terminology and terminography in an interdisciplinary approach. Specialized communication and translation problems cover multiple languages, such as English, French, German, Romanian, Hungarian and Russian. Diachronic sources and term trajectories have been taken into consideration for a better explanation of synchronic behavior. The denotative mobility of terms correlated with contrastive studies (English–French, Russian–Romanian, Hungarian–Romanian, etc.) confers a unique flavor to this volume. The uniformity stems from the topic as well as from the level of approach: term analysis and specialized senses, conceptual-semantic dimension, as well as syntagmatic and paradigmatic levels. The studies in the volume go beyond the lexical semantic analysis, as the authors also pursuit problems of contextual terminology, specialized translation and interculturality.

Doina Butiurcă

Doina BUTIURCĂ is an associate professor at Petru Maior University, Târgu-Mureş, Romania, publishing studies on linguistics. Attila IMRE is an assistant professor at Sapientia University Cluj-Napoca, Romania. Inga DRUȚĂ is an associate professor and researcher at the Institute of Philology of the Academy of Sciences of Moldova, Chişinău.

978-3-659-37328-2

Specialized Languages

Butiurcă, Imre, Druță

Doina Butiurcă
Attila Imre
Inga Druță

Specialized Languages and Conceptualization

LAP LAMBERT Academic Publishing

Doina Butiurcă
Attila Imre
Inga Druță

Specialized Languages and Conceptualization

LAP LAMBERT Academic Publishing

Attila Imre (2013) Machine Translation Reloaded, in Butiurcă, D. & Imre, A. & Druță, I. (eds) Specialized Languages and Conceptualization. Lambert Academic Publishing, pag. 135–154. ISBN 978-3-659-37328-2. CNCS

Attila IMRE: Machine Translation Reloaded

Abstract

The article offers an introduction about machine translation (MT), focusing on various definitions as well as a short history of MT. As there is a lot of controversy about MT, we have tried to offer both advantages and disadvantages of, presenting a translation task with the help of *Google Translate* carried out in 2011 and repeated in 2012. The results are discussed, followed by certain allegations and conclusions about MT.

Key-words: machine translation, definition, quality, human intervention, gisting.

Introduction

It is in fashion to measure everything from the point of view of profit. Language services or translation activities enjoy an increasing interest, as a multitude of companies have budgets for these activities, especially machine translation programs after the Second World War (Lambert & Hermans, 2006, p. 149). According to European Union studies, the growth rate of the language industry is estimated at 10% minimum over the next few years...

Machine translation (MT) probably constitutes one of the hottest topics regarding both the present and the future state of translation industry, bugging the mind of people hoping for an effective universal translator. The rapid technological developments resulted in deep and lasting impressions regarding this field as well, and we tend to believe that spectacular improvements are yet to come.

A very simplistic formula of the translation process includes decoding the source text on the one hand, and on the other hand recoding

'it' in the target language. Now, the problem of 'it' opens the evergreen problem: should the translation focus on the *meaning* or the *words*? If we focus on the meaning, the problem of synonyms will inevitably appear, but if we focus on the words, we will never be able to correctly render idioms or metaphorical expressions. This is why we can consider translation a complex mental activity during which the entire interpretation of the source text is necessary. This should involve lexical, grammatical, semantical, syntactical, idiomatical, cultural and contextual analysis; cognitive linguistics even discusses background- or world-knowledge (Imre, 2010) , and we are not sure that all important aspects of text-translation are listed here.

Ardent defenders of machine translation usually highlight these aspects, benignantly overlooking the idiomatical, cultural, contextual or other knowledge parts. Thus human "understanding" seems to be unrivalled in this regard, although huge improvements have been made in the field of machine translation. At present it is questionable whether humans can program a computer to create an "authentic" target text based on a source text, although there are different approaches. These may be of various or combined types (some even running on cross-platform, such as *IBM*) detailed below:

- *Rule-based interlingual machine translation*[21] approach: the source text is transformed into an interlingua, i.e., an abstract language-independent representation, out of which the target language is generated; this may offer good results only in a very specific domain (e.g. *SYSTRAN*, *Eurotra*, *Apertium*[22], *GramTrans* for Scandinavian languages and English[23]);

[21] Source: http://en.wikipedia.org/wiki/Interlingual_machine_translation, 10.02.2013.
[22] Further details: http://en.wikipedia.org/wiki/Apertium, 12.02.2013.
[23] Further details: http://en.wikipedia.org/wiki/Gramtrans, 12.03.2013.

- *Rule-based transfer machine translation*[24]: the source language is transformed into an abstract representation out of which an abstract target language representation is obtained, then the target text is generated, taking into account linguistic rules of the involved language pairs; this transfer may be either superficial (syntactic, for languages belonging to the same/similar family) or deep (semantic, between distant languages):
    1. *shallow-transfer machine translation*[25] assumes no previous knowledge of the text, thus they simply apply statistical methods to the words surrounding the ambiguous word; this approach has been more successful until now than deep-transfer;
    2. *deep-transfer machine translation* presumes a comprehensible knowledge of the word (e.g. *Matxin*, *OpenLogos*);
- *Rule-based dictionary machine translation*[26] uses a method based on dictionary entries, thus the words will be translated as a dictionary does: string of words with little correlation of meaning between them.
- *Direct machine translation*: words are translated directly;
- *Statistical machine translation*[27] (SMT) tries to generate translations based on bilingual text corpora (e.g. Canadian Hansard corpus, record of the European Parliament or Google's SYSTRAN). These results may be noteworthy, and the development of these corpora may lead to a more unequivocal success of machine translation overall (among the most known examples is *Google Translate*);

[24] Source: http://en.wikipedia.org/wiki/Transfer-based_mchine_translation, 10.02.2013.
[25] Source: http://en.wikipedia.org/wiki/Machine_translation, 10.02.2013.
[26] Source: http://en.wikipedia.org/wiki/Dictionary-based_machine_translation, 10.02.2013.
[27] Source: http://en.wikipedia.org/wiki/Statistical_machine_translation, 10.02.2013.

- *Example-based machine translation*[28] (EBMT) is in fact an analogy translation making use of a bilingual corpus.
- *Hybrid-based machine translation* (HMT) tries to balance the advantages and disadvantages of both rule-based and statistical MT either by rules post-processed by statistics or by statistics guided by rules (e.g. *Wordlingo*).

The list is probably not complete[29], but the basic idea is visible: rule-based methods parse the source text up to a certain (below-human) level, and due to this drawback the generated target text needs "human intervention." As Prószéky concludes, rule-based machine translation is characterized by low recall and high precision, whereas statistical systems tries to handle each and every one case, but it is prone to error even in trivial cases (Prószéky, 2005, p. 80).

Anyway, at present we are less interested in the machine translation method than in its output, so in the following we will try to offer definitions of machine translation, looking into its history, offering both pros and cons. Furthermore, we are going to test it before enlisting some allegations and offering some concluding remarks.

Definition of MT

There are more possibilities to define machine translation. The simplest definition describes MT as a procedure by which an activated computer program analyses the source text and produces a target text *without further human intervention.*[30] However, there are more detailed approaches: *machine aids* for translators, *machine-aided* translation and *machine*

[28] Source: http://en.wikipedia.org/wiki/Example-based_machine_translation, 10.02.2013.
[29] Further details:
http://en.wikipedia.org/wiki/Comparison_of_machine_translation_applications, 12.02.2013.
[30] Source: http://en.wikipedia.org/wiki/Machine_translation, 12.02.2013.

translation (cf. Blatt's definition in Freigang, 2001, p. 134). We knowingly avoid the term *computer-aided* translation, as this refers to something different today. Machine aids may include word processors, e-dictionaries, various term banks without performing the translation task; Blatt defines machine-aided translation as a tool to help the translator, whereas machine translation is fully automatic. Naturally, human post-editing is more than desirable in this case.

Thus our definition of machine translation sounds like this: a computer program either separately installed on an operating system or accessed online, which is capable of reading a source text in a(ny) language and – *without* human intervention – is also capable of transforming it into a(ny) different language in a comprehensible way by a target speaker. However, the prerequisites include human intervention in the form of preparation (pre-editing):

1. choosing a natural source language;
2. preparing the source text in machine-readable form (e.g. scanned handwriting excluded);
3. selecting the target language.

The modus operandi does not constitute the active part of the definition, and neither does the analysis/appreciation of the result. Nevertheless, at present, without post-editing one should not expect very good results, which may be explained by understanding how machine translation works. Albert explains that machines do not "translate", they do not search for equivalents or look for meanings, and they cannot "read between the lines." Instead, they recode a language system into another with formal equivalents, operating with word-meanings, hence we can conclude that this is nothing else than mere code-switching (Albert, 2011, p. 81) without the real problem of polysemy. However, when we would

like to translate anything with machine translation (e.g. *Google Translate*) we are already offered possible variants.

History of MT

Somers' article (Somers, 2001a, pp. 140–143) succinctly summarizes the history of machine translation, starting with the first patents from 1933, then mentioning the Cold War period after the Second World War. Although Alan Turing suggested nonnumeric applications, the "founding father" of machine translation is Warren Weaver.

Yehoshua Bar-Hillel became the first full-time researcher at MIT in 1951, and in 1956 the first MT conference was held at MIT, where USA, Britain, Canada, the Soviet Union sent delegates, although Japan also started work on MT at Kyushu University. It goes without saying that the funding had mostly military reasons and the predominant technique was 'dictionary-based direct replacement.'

The Georgetown experiment[31] in 1954 was considered to be a huge success (Russian–English translation), and the authors claimed that within the next three or five years MT would be a solved problem. Evidently, this did not happen; moreover, in 1966 the (in)famous ALPAC report (Hutchins, 1996) resulted in a dramatic cut in funds, as according to the expectations regarding MT were not met. The conclusion was that "MT was slower, less accurate and twice as expensive as human translation" (Somers, 2001a, p. 140). The first serious attempts in developing MT systems were supported by mainly the United States of America and the USSR due to the fact that they wanted to move out the rocket technology from German, then to spy on each other (Biau Gil & Pym, 2006, p. 16).

After this first 'direct' wave a second followed, which may be called 'indirect', using a transfer-based approach or the previously mentioned

[31] Source: http://en.wikipedia.org/wiki/Georgetown-IBM_experiment, 10.02.2013.

interlingua approach. The most notable success of this period is the *MÉTÉO*^{TM} system, which was used to translate weather bulletins from English into French.

A further very successful system was *SYSTRAN*, which was used even by the USAF and NASA as it could carry out rough translations from Russian into English. The English–French version of *SYSTRAN* was also used by the Commission of the European Communities. Although we could witness the appearance of the first commercial MT systems, potential customers in USA did not see the possibilities in them, whereas the standards were far below the expectations of European customers, as Somers observes (Somers, 2001a, p. 141), not mentioning that they were expensive and translators felt (and still feel) threatened by them. However, the eternal problem of MT, namely how much 'understanding' of a text, was not solved.

Nevertheless, at the beginning of the 1990s MT was supported again (companies like *Philips*, *Siemens*, *IBM* and mainly Japanese electronics companies), although the hope in the development of artificial intelligence (AI) and its support in MT never came true. The latest developments focus on the use of corpora instead of linguistic algorithms, and statistics in this respect seems to be more successful.

At present *SYSTRAN* powers online both *Google Translate* and *Babelfish* of Alta Vista, and they state that the output is "reasonable". Anyway, throughout the last sixty years many scholars seriously doubted the possibilities of MT, especially when high quality was at stake. Since the ALPAC report (which, by the way, mentioned many positive facts about MT) there has been a tendency not to accept MT without human intervention, especially in fields of high stake, such as legal translations or medicine (Kis & Mohácsi-Gorove, 2008).

Interestingly, where formal or formulaic language is used, machine translation of government and legal documents more readily produces usable output than conversation or less standardized text.[32]

Although translation is more tedious and almost impossible without the help of computers nowadays and MT is getting better and better (Prószéky, 2005), the outbreak is yet to come.

Advantages of MT

We have already mentioned that the limited success of the second 'indirect' approach was partially due to the high expectations in Europe. However, if we focus on positive aspects of MT, we can mention Somers' analysis of rough/raw input of the fully automatic MT (Somers, 2001b, pp. 137–8). He claims that this "may still be useful, even though it may lack in style or even accuracy" when the source text is an 'exotic' language (cf. unfamiliar writing system) or "the consumer may be a (perhaps amateur) translator, or a subject specialist", as the 'quick and dirty' first draft offers a clue for the translator about the relevant parts or the scientist wants "to know only roughly what the article says."

Others correctly observe that even human translation is "subject to revision" (Somers, 2001b, p. 138); remarkable facts are that the revision of MT does not hurt feelings, and it is prone to commit recurrent mistakes which may be easily post-edited with further interactive tools or near synonyms by accessing dictionaries and thesauri. Indeed, proper terminology work and serious post-editing of MT can produce results in a shorter period of time. However, this is only a possibility by itself, and at present the tendency is to have an MT system integrated into a computer-assisted translation tool (cf. *MemoQ*).

[32] Source: http://en.wikipedia.org/wiki/Machine_translation, 10.02.2013.

A further real advantage of MT is "gist translations from languages you know nothing about. It allows users to identify the texts or fragments of interest, which they can then have translated by other means" (Biau Gil & Pym, 2006, p. 16). They also mention that MT produces "high quality translations in very restricted contexts", but in our opinion this is not real-life situation. However, gist translation functions very well when in the 21st century people keep buying from *eBay* (technical gadgets or common, everyday products).

Yet, there are real-world applications; in this respect we can mention the customized version of *SYSTRAN* by the European Commission, the English–Danish *PaTrans* belonging to *Lingtech A/S* (combined with *SDL Trados* commercial CAT tool), or the Spanish–Catalan MT system for the Spanish daily newspaper, *Periódico de Catalunya*.

Last but not least we can mention the efforts against terrorism and international warfare, which excessively make use of the benefits of MT, for instance US Air Force has awarded a $1 million contract to develop a language translation technology.[33]

If humour is considered an advantage, then in particular cases MT systems should be awarded the special prize of *the most "authentic" and funniest mistranslation providers*, to be discussed below (cf. *Testing MT* section).

Disadvantages of MT

From the outset we would like to present the most negative evaluation of MT, to be found on Wikipedia's Hungarian page of what Wikipedia is not. It is explicitly stated (point 24) that Wikipedia does not collect machine translations, as texts translated by MT – in their experience – are simply not comprehensible in Hungarian, and they are not worth correcting

[33] Source: http://en.wikipedia.org/wiki/Machine_translation, 10.02.2013.

either.[34] However, nothing similar is to be found on either the corresponding English or Romanian Wikipedia page.

Another problem (initially raised by Yehoshua Bar-Hillel) is *word sense disambiguation* in case a word has more than one meaning. In the 1950s this was an insurmountable problem, but today this may be (at least partially) solved with either the shallow (no knowledge of the text) or the deep approach (comprehensive knowledge of the word). As human (natural) language is contextual, seemingly little chance is given MT to "understand" it, but sometimes even humans misinterpret contexts (and not only in case of puns, jokes but everyday conversations as well).

A certain disadvantage of MT comes from proofreaders scared about the future possibilities of MT (to be discussed later). They tend to highlight all possible errors and they often approach the MT-produced target text subjectively. Hence the criticism that MT is only "gisting translation", often requiring an experienced reader with good knowledge of both languages. Yet, the basic idea of gisting translation is to offer a rough translation of a language the reader has 'no idea about'. No wonder that Claude Piron concludes (Piron, 1994):

MT, at its best, automates the easier part of a translator's job; the harder and more time-consuming part usually involves doing extensive research to resolve ambiguities in the source text, which the grammatical and lexical exigencies of the target language require to be resolved."[35]

We have already mentioned that the translation of idiomatic expressions, metaphors, puns is extremely problematic for MT, often resulting in 'patent nonsense' and undeniably authentic humour. Beaugrande and Dressler explain that "a computer working only with a

[34] Source: http://hu.wikipedia.org/wiki/Wikipédia:Mi nem való a Wikipédiába?, 10.02.2013.
[35] Source: http://en.wikipedia.org/wiki/Translation, 10.02.2013. English summary on Piron's book in French.

grammar and lexicon (both virtual systems) was found unable to operate reliably, because it could not evaluate context" (Beaugrande & Dressler, 1981). In the following we would like to present two tests carried out in 2011 and 2012.

Testing MT

Machine translation is now available in many languages (cf. www.word2word.com), but it has become popular among people looking for gisting translation thanks to the Internet (*Babel Fish*, *Babylon* or *StarDict*). *Ectaco* produces pocket translation devices, which use MT (www.ectaco.com).

Google also uses MT (Google translation tools), which is getting more and more widespread. In order to test its efficiency, we selected a real conference program in 2011 and 'gave' it to Google translator. The source text was in Hungarian and we asked for a Romanian translation. The table below enlists 7 highly incorrect cases:

Hungarian source text	**Romanian target text**
Marosvásárhely 2011 március 4-5.	Targu Mures 2011 martie 4-5. *Missing Romanian diacritical marks (Târgu-Mureş); non-typical form of date (4-5 March 2011).*
Péntek, 2011 március 4.	Vineri 2011-3-patru *Date misunderstood, it is translated as a string of numbers.*
9. -- Megnyitó, dékáni köszöntő	Nouă - - Ceremonia de deschidere, Bine ati venit Dean *Instead of preserving the opening time (9 o'clock), it is transcribed into a word, which is unfortunately homonymous and disambiguation is needed. Furthermore,*

	dean *is in English.*
9.20 -- Veress Károly: Valóságépítő virtualitás	9.20 - Charles Veress: Utilaje constructii virtualitate Reality *Proper name is translated, but not into Romanian (Carol). The worst misinterpretation: instead of constructing reality one can find construction equipment due to the 'unfortunate' word-boundary in Hungarian.*
Haller Piroska: Számítógépek ott és egykor	Little Red Riding Hood Haller: Computere şi există un timp *Proper name is translated into English, resulting in a well-known character of a tale; moreover the title of her presentation is nonsense.*
14.40. -- Tőkés Gyöngyvér: Digitális magatartásformák a virtuális térben	14:40. - Blood Pearl Mallard: comportament digitale în spaţiul cibernetic *Proper name is translated into English. This is a very unfortunate translation as the first name was identified as a compound name (pearl and blood), whereas the last name is turned into a common name (wild duck).*
Zsigmond István	Stephen King Sigismund *Proper name is translated into half-Romanian, half-English, resembling either a famous historical name or the name of an American bestseller-writer.*

Table 1. Google Translate Test 1, 2011

Our first remark would be that although the 'job' was to translate from Hungarian into Romanian, some words are translated into English;

this might mean that whatever language combination we choose, Google translator will first use English as the mediating language. And indeed, Jost O. Zetzsche supports our observation (Zetzsche, 2009). When no corresponding word is found, the English intermediary word is preserved, and in case the initial Hungarian is preserved (rare/longer/compound words), then there is no English translation for that word. As of the beginning of 2013, there are more than 60 source and target languages available for Google translate[36], although Zetzsche mentions more than 300 languages, Latin included.

Of course, there were successful parts during the translation (Hu: *Szemelvények a virtuális világtörténetéből*; Ro: *Extrase din istoria lumii virtuale*; En. *Extracts from the history of the virtual world*), but we should also take into consideration that the text was very easy: titles and names of a conference, without elaborated sentences or longer paragraphs. Yet, the aforementioned verdict is still valid: at its best, MT is (only) good for gisting.

However, we repeated the test in 2012, one year after the initial test, presented in the table below:

Hungarian source text	**Romanian target text**
Marosvásárhely 2011 március 4-5.	Targu Mures martie 2011 4-5. *Missing Romanian diacritical marks (Târgu-Mureş); 'a bit' closer to the typical Romanian form of date.*
Péntek, 2011 március 4.	Vineri 2011-3-patru *No change.*
9. -- Megnyitó, dékáni köszöntő	Nouă - Deschiderea, Bine ati venit lui Dean *No real change.*
9.20 -- Veress Károly: Valóságépítő	9.20 - Veress Karoly: constructii virtualitate

[36] Source: http://translate.google.com/, 12.02.2013.

virtualitás	Realitatea *Huge improvement: name not translated (bt Hungarian form is preserved without the diacritical marks), no misinterpretation of the compound word.*
Haller Piroska: Számítógépek ott és egykor	Piroska Haller: Calculatoare şi o dată acolo *Huge improvement in not translating the proper name and even correctly changet to Romanian style;the title is still puzzling/humorous.*
14.40. -- Tőkés Gyöngyvér: Digitális magatartásformák a virtuális térben	14:40. - Mallard Gyöngyvér: comportament digital în spaţiul cibernetic *Proper name is partially translated into English. First name is recognized as a proper name, last name is still translated as a proper name (wild duck), wrong name order.*
Zsigmond István	Stephen King Sigismund *No change.*

Table 2. Google Translate Test 2, 2012

We can conclude that within a year (!) considerable improvements have been carried out in three out of seven cases. The English imprint upon the Hungarian–Romanian is still visible, but proper names and word-boundaries were recognized, although not in all the cases. Still, we consider that this is an encouraging result, at least at the level of lexicon.

Allegations and Conclusions

The first allegation regarding the problem of MT probably comes from 1954, when it was already predicted that by 1957 or 1959 the problem would have been "solved" forever. Fifty years later (in 2004) MIT's

Technology Review stated that universal translation and interpretation will likely to become available "within a decade."

A recent prediction is signed by Raymond Kurzweil, who is an American author of books on health, artificial intelligence, transhumanism, technological singularity and futurism, but he is also inventor, futurist and director of engineering at Google, attracting "significant criticism from scientists and thinkers."[37] Kurzweil predicted that, by 2012, machine translation will be powerful enough to dominate the field of translation. [38]

Has MT become so powerful by now? According to our two tests, MT is rather far from having solved the "problem of translation", but let us check what the experts have to say about it. The ultimate question regarding MT may be verbalized this way: Who has to be satisfied? Somers mentions different users: end-user (i.e. the consumer of the translation), an intermediate agent, the translator, and the original author of the text to be translated (Somers, 2001b, p. 136), and we tend to believe that even the results of our brief tests are not satisfactory. Thus our verdict is compatible with Biau Gil and Pym, who state that "current systems are unable to produce output of the same quality as a human translator, particularly where the text to be translated uses casual language" (Biau Gil & Pym, 2006).

Shields mentions that dictation to a computer is already used and voice can be generated from a computer, seriously affecting interpretation as well. However, rule-based MT cannot be extremely successful, as Nagao sees the key in the quality of samples, which should result from real-life situations not from extended rules (Prószéky, 2005, p. 81). Biau Gil and Pym are more convinced that MT are not replacing human translators as

[37] Source: http://en.wikipedia.org/wiki/Raymond_Kurzweil, 10.02.2013.
[38] Source: http://www.axistranslations.com/translation-article/what-is-translation.html, 10.02.2013.

fully automated MT is not a viable solution; furthermore, quality MT needs "serious attention to controlling writing of the input, which is an area that some translators may want to move into." (Biau Gil & Pym, 2006, p. 17). This means that the more texts we have, "the more texts will be processed, and the more work will be created for human translators." They also repeat that at present less time is needed to translate from scratch than correcting MT errors.

Post-editing may result in more work on behalf of the proofreader than a new translation, but we should not overlook the fact that the number of those who are satisfied with gisting translation is on the increase due to the presence of the Internet on a large scale (desktop computers, notebooks, smartphones, iPads, etc.).

Official statements report that Google improved their translation capabilities by inputting approximately 200 billion words from United Nations materials to train their system, thus accuracy of the translation has improved.[39]

Prószéky tries to extenuate the verdict by differentiating texts, and accepting that literary text are non-translatable by MT (lack of cultural background), but MT may function better with specific/technical texts (Prószéky, 2005, p. 79). He goes on and summarizes: our present MT systems should both translate and be tolerant with errors as the primary aim is understanding the content and definitely not to challenge the high quality human translations (Prószéky, 2005, p. 83). He also adds, that the greatest challenge is in fact the internet with extremely many incorrect instances, which cannot be corrected. MT cannot decide whether it is faced with a new word/phrase or it is just a typo (Prószéky, 2005, p. 82).

The basic level MT simply substitutes words for words (and it is much faster and rather reliable in case of nouns) and potential users may

[39] Source: http://en.wikipedia.org/wiki/Machine_translation, 10.02.2013.

take it 'as is.' Its speed will surely affect the translation industry, as Mike Shields (Anderman & Rogers, 2003, p. 43) explains it during a round table discussion:

translation work – for many, many translators – is going to be seriously affected, because, just as an example, most of us are paid per 1,000 words or per page, or per line, and this is fine, so long as it represents an average of the hard bits and the easy bits. Everybody knows that in some translations you find a bit that you're very familiar with ... And then you find other bits where you're not quite sure what it means, or what the word means in this context, and you can spend an hour thinking about it. And what's going to happen is that all the easy bits are going to be taken out of translation, and we're going to be left only with the difficult bits... I can see novels being banged out in machine translation systems and handed over to ghost writers to turn them into as good English as is necessary, and completely wipe out translators – and even interpreters.

And this is exactly what the late Claude Piron said: the harder and more time-consuming part involves doing extensive research to resolve source text ambiguities[40]:

Why does a translator need a whole workday to translate five pages, and not an hour or two? ..... About 90% of an average text corresponds to these simple conditions. But unfortunately, there's the other 10%. It's that part that requires six [more] hours of work. There are ambiguities one has to resolve.

In Piron's estimation only 25% may be automated, whereas the harder 75% is still done by a human translator. Boulton comments on the translation feature of Google Chrome[41]: "not all of the translations will be clean, crisp and accurate. But as with everything else Google does, Translate is an iterative technology that will Google will advance over time." (Boulton, 2010); Jost O. Zetsche is more concise: "neither the tool

[40] Source: http://en.wikipedia.org/wiki/Machine_translation#cite_note-piron-10, 13.02.2013.
[41] http://www.eweek.com/c/a/Web-Services-Web-20-and-SOA/Google-Chrome-Gets-Machine-Translation-New-Privacy-Features-839506/, 13.02.2013.

vendors nor translation agencies should be too worried" regarding MT (Zetzsche, 2009), although he remarks that whenever we upload an existing translation memory or perform any translation "the material will be used by Google for the training of its machine translation engine, even if you declare your translation memory to be »private«, and even after you »delete« it." If true (and we see no reason why it should not be), certain conclusions may be formulated:

a. MT will be better in time, including more and more languages;

b. MT will be used by an increasing number of people;

c. the more people will use MT (predominantly non-professionals), the less trustful MT will be for professionals, which is not in contradiction with point a.; in fact, the present-day 'balance' will be kept, as larger and larger database (partially unchecked) enters the MT translation memory.

Of course, those involved in the development of MT will highlight only the positive parts. For instance, *Ectaco* promises that anything you say will be translated by their speech translator (languages and quality are not mentioned!)

Other noteworthy endeavours are Microsoft Office's *Bing* translator[42], which offers instant translation from/to multiple languages without bearing any responsibility for the content and replaces with a single click the source text with the target text (built-in option of Revision tab), or the recently discovered free *Glosbe Online Dictionary*[43], whose name may be misleading: although it is a "dictionary", full sentences can be translated, even specifying the source (mainly the *Official Journal of the European Union* or *opensubtitles.org*). Allegedly, their translation memory

[42] Available at http://www.microsofttranslator.com, 12.02.2013.

[43] Available at http://glosbe.com, 13.02.2013. Many thanks to M. Popa for drawing my attention upon this.

is the largest online, containing more than 1 billion (10^9) sentences in many languages and possibly many *tmx* files from volunteers.

Finally, it would not be nice to leave the patient reader without a prediction about the future. Kis–Mohácsi-Gorove state that man cannot be fully replaced: the quality of machine translation will not reach the standard of a mediocre human translation for a very long time, so the bulk of translation is left for human beings. Legal, technical and other special texts are to be translated by human beings only, and even this has to be proofread (Kis & Mohácsi-Gorove, 2008, p. 13). And we fully agree with that.

References

Albert, S. (2011). *"A fövényre épített ház" A fordításelméletek tudomány- és nyelvfilozófiai alapjai*. Budapest: Áron Kiadó.

Anderman, G. M., & Rogers, M. (Eds.). (2003). *Translation Today: Trends and Perspectives*. Multilingual Matters.

Beaugrande, R. D., & Dressler, W. U. (1981). *Introduction to text linguistics*. Longman.

Biau Gil, J. R., & Pym, A. (2006). Technology and Translation. A pedagogical overview. In *Translation Technology and its Teaching (with much mention of localization)* (pp. 5–19). Tarragona: Intercultural Studies Group.

Boulton, C. (2010, March 2). Google Chrome Gets Machine Translation, New Privacy Features. *eWeek.com*. Retrieved from http://www.eweek.com/c/a/Web-Services-Web-20-and-SOA/Google-Chrome-Gets-Machine-Translation-New-Privacy-Features-839506/.

Freigang, K.-H. (2001). Machine-aided translation. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 134–136). London and New York: Routledge.

Hutchins, J. (1996). ALPAC: the (in)famous report. *MT News International*, 9–12.

Imre, A. (2010). *A cognitive approach to metaphorical expressions*. Cluj-Napoca: Scientia.

Kis, B., & Mohácsi-Gorove, A. (2008). *A fordító számítógépe*. Bicske: Szak Kiadó.

Lambert, J., & Hermans, J. (2006). From translation markets to language management: the implications of translation services (1998). In *Functional Approaches to Culture and Translation* (pp. 147–162). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Piron, C. (1994). *Le défi des langues - Du gâchis au bon sens*. Paris: L'Harmattan.

Prószéky, G. (2005). Egy új elvű gépi fordítási rendszer felé. In *"Mindent fordítunk, és mindenki fordít" Értékek teremtése és közvetítése a nyelvészetben* (pp. 79–83). Bicske: Szak Kiadó.

Somers, H. L. (2001a). Machine translation history. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 140–143). London and New York: Routledge.

Somers, H. L. (2001b). Machine translation applications. In M. Baker (Ed.), *Routledge Encyclopedia of Translation Studies* (pp. 136–139). London and New York: Routledge.

Zetzsche, J. O. (2009). Let's Talk: Trados and the Google Translator Toolkit. *The ATA Chronicle*, *38*(10), 18–21.