Hermits with Duden, tipex and an Olivetti Studio 45

Ricardo Muñoz Martín

Ricardo Muñoz Martín
MC2 Lab & *Alma Mater Studiorum* Università di Bologna;
ricardo.munoz@unibo.it;
ORCID: 0000-0001-6049-9673

#tradumatica20_Q6

What working ways have totally changed compared to the ones you used to adopt when you started that you would not give up under any circumstance?	
Quina manera/procés del treball com a traductor ha canviat des que vas començar i ara no hi renunciaries de cap de les maneres?	EN / CA / ES
¿Qué manera/proceso del trabajo como traductor ha cambiado desde cuando comenzaste al que no renunciarías ya bajo ningún concepto?	

Abstract

This little article summarizes the changes in translators' work from the 1990s as a tribute to the *Revista Tradumàtica* on its 20th anniversary. No surprises: text processing, the Internet as a source for information, the use of CAT tools and neural MT, and changes in the market and labor relations stand out as the prominent features that have changed (or not) in the last 30 years.

Keywords: text processing; Internet; CAT tools; MT; market culture.

Resumen

Este articulillo resume los cambios producidos en la labor de los traductores a partir de la década de 1990, como homenaje a la *Revista Tradumàtica* en su 20º aniversario. Nada nuevo: el procesamiento de textos, Internet como fuente de información, el uso de herramientas TAO y la TA neuronal, y los cambios en el mercado y las relaciones laborales predominan como los elementos más destacados que han cambiado (o no) en los últimos 30 años.

Palabras clave: procesamiento de textos; Internet; herramientas TAO; TA; cultura de mercado.

Resum

Aquest breu article fa un repàs dels canvis produïts en la labor dels traductors a partir de la dècada de 1990 com a homenatge a la *Revista Tradumàtica* en el seu 20è aniversari. Res de nou: el processament de textos, Internet com a font d'informació, l'ús d'eines TAO i la TA neuronal, i els canvis en el mercat i les relacions laborals són els elements més destacats que han evolucionat (o no) en els últims 30 anys.

Paraules clau: processament de textos; Internet; eines TAO; TA; cultura de mercat.

When Mikhail Gorbachev announced the end of the Cold War at the United Nations, I was starting my PhD in California. I had been a part-time Spanish teacher in Munich, where I also used to type away translations daily and mail them weekly to Granada to meet the requirements for some courses I had to finish in order to graduate as a translator. At the same time, I was also moonlighting as a freelance translator to make ends meet. I had brought my Olivetti Studio 45 with me from Spain and would cross the Atlantic with it too, but right away I started going to campus on the weekends to discreetly teach myself how to use the IBM PCs (AT) at the department and to develop a course to teach Spanish phonetics in Hypercard—let's say, PowerPoint's ancestor—on one of the Macintosh IIs in the open access labs.

I was still a typewriter guy. I wrote the titles of my texts in capitals and underscored them with a matching line of equal signs. It took me three months to discover the caps lock key and stop typing my headings with my left pinky stiff on the shift key to keep it pressed down. That day, upon my finding, I walked into a Radio Shack store to come out with my first computer, a Tandy 1400LT laptop. That technological bulky beauty had no hard drive. You had to take out diskette 1 of WordPerfect 4.2 to introduce diskette 2 with the spellchecker, run it on the open file, and save your work on the diskette in the other drive. Still rudimentary, but already powerful enough. Like databases. Back in 1990, a university librarian could not understand the disappointment in my face as she handed me results a-plenty of several days of patient and meticulous searches, where she had ingenuously combined the terms from my oh-so-long keyword list. Mind you, all references of translation scholarship ever written fitted in two of those 3.5" diskettes!

Those were daring times. Madonna, Bon Jovi, Phil Collins, Paula Abdul were often on the billboards but I preferred Tracy Chapman and longed to go to Nelson Mandela's 70th Birthday concert in Wembley. *Rain Man* was number one in the box office, but I cherished *A Fish Called Wanda*. Act Up was denouncing the governments' passivity facing the AIDS pandemic (they still do). As Reagan and Thatcher were pushing their neocon involution, I was learning from Prof Luis Fernando Lara that languages are nowhere, that they are something all their speakers share and nobody owns. Fukuyama (1989) proclaimed the end of history at the end of a very short 20th century (Hobsbawm 1994) but the eyes of language scholars and translators were rather turned to the horizon of a new dawn. Email became popular thanks to clients like Eudora, and soon after the Internet started hosting the world wide web (motto: *Let's Share What We Know*), which simply changed the world we used to live in. And us.

Roughly 30 years later, by the end of 2022, five billion Internet users spend about seven hours a day on average on at least 200 million active websites. The web is described as a modern, larger library of Alexandria. Google and Baidu, currently the

major web engines, witness eight out of ten web surfers using their search services more than three times a day. However, an important share of lookups are simply seeking to access other sites, such as YouTube, Facebook, Amazon and Instagram. About half of the searches are seeking local contents (like the weather) and ca. 60% come from cell phones and tablets. Regarding searches, the trumpeted Information Age seems to boil down to comfortably finding your way to that new grocery store or checking whether the rain will ruin your plans for next weekend.

Of course, that was the flip side but this coin has another one. *Wikipedia* offers information in 329 languages and reports 271 billion lookups in the last year—nearly 34 per human on this planet. *Encyclopædia Britannica* went online in 1994 and only online in 2011.[1] English dominates the Internet, so figures for Spanish are much more temperate. Still, in 2021 the RAE dictionary alone surpassed one billion lookups (RAE 2021). The authoritative dictionary was first freely accessible online in 2001, just about 10 years after the web had opened up. The Spanish Royal Academy still offers its products both in print and online, as does Duden for German. Yet in most developed countries, information search and management—especially, for workers in the language industries— has gone digital and, for years, print sources have been collecting dust on the shelves of nostalgics, as my old Duden books do. I feel a silly comfort in seeing them there.

The Internet does not only offer a wealth of information. It frequently does it for free and from nearly anywhere, through the same means (sometimes through browser plugins, such as those listed in the Mycroft project), and in formats that even spare the users the need to type what they find and paste into their texts. It's the translators' panacea. Nevertheless, the possibilities are drastically reduced once you work with languages other than the most spoken ones (Joshi et al. 2020). The creation of language resources is usually driven by profit prospects and, to a lesser extent, public policy. For instance, four American languages with 1.5–60 million speakers have fewer resources than Estonian, spoken by just about one million people (Hedderich et al. 2021: 2546). Sure, commendable efforts are on the rise (Magueresse, Carles & Heetderks 2020) but there is not enough interest; not enough money is put there.

Switching to digital means for finding information has other, perhaps deeper consequences. First, searching styles have changed. For instance, rather than contextualizing a notion in our minds, classing it in a library classification system, and then looking for it in that place, as we used to do, full-text searches quickly made comparable Internet approaches obsolete, like the Yahoo and the DMOZ directories. Now there is no need for abstraction and perspective, just the token word. We started our journey in a world where information was difficult to find and we sharpened our critical thinking to make sure that we did not waste our time, that what we found was good and right. Now we have literally too much information and disinformation, and now and then we cannot tell the difference between them (Mitchell et al. 2020). We have grown used to the mantra that memorization is a matter of the past but there is no way around the fact that knowing something entails having it in our memories (Sanger 2010:

---

[1] It was Microsoft's encyclopedia Encarta, however, and not *Wikipedia* that killed the printed *Britannica* (Greenstein 2017), only to be itself discontinued in 2009.

s.p.), even if in partial, schematic ways. We face the paradox that we now have increasingly open access to all kinds of documents but seem less and less ready to understand them.

Second, we can now choose from an ever-vaster array of world-wide cultural products in digital form, which we can often enjoy when we want. Personal preference and chance combine to make people's cultural profiles—their connectomes—more different from each other's than before. We, do not share as many experiences and knowledge that in the good old analog times made up part of our daily experience and our identities. In an age of paradoxical nationalist revivals, global communications are also ironing out national differences. Of course, this is only a matter of degree. Yet irony is now more difficult because we cannot be sure that we share the same values and cultural clues. The implicit images that as translators we build of our readers and writers are now fuzzier, we seldom meet our clients, we are not as sure that our students will understand what we say or mean. The notion of digital isolation is being used to describe the situation of people who cannot access the Internet as much as others do (we will come back to that) but the hikikomori lifestyles of the new digital hermits that technology helps to sustain (Rooksby, McLeod & Furuhash 2020), are all but dismissed as a collateral damage of our Zeitgeist. Sadly, some digital nomads learn nothing about the place they squat in, it is nothing but an exotic backdrop.

Third, digital communications are also gathering groups that would have been unthinkable before, like fansubbers and scanlators, who often act as prosumers simply because their interests cannot be monetized. The Internet itself fosters a continuous growth of translation demands, and today's more prominent machine translation (MT) investors are not in the language industries, like Facebook and Amazon. In this increasingly digital society, unpaid translation is not only for hobbies. Public administrations, scholars (Bowker & Buitrago 2019), clinical doctors (Khoong & Rodríguez 2022), patent professionals (Nurminen 2020), college students (Zhou, Zhao & Groves 2022) and many others use MT for their daily tasks (see Orrego-Carmona, this issue). Their needs have become part of the goals and object of study of many research projects in translation & interpreting studies—e.g., Projecta U, which focuses on MT users. [Non-professional] translation is now the fifth language skill (Pym, Malmkjær & Gutiérrez-Colón 2013).

Another area of radical change is in the tools of the trade, in the amazing possibilities that computers and their applications have brought to our fingertips. Text processors are still by far the best and most typical tool translators use. They may not seem remarkable if you are under 40. But those of us who had to retype full pages, use real carbon copies for our records and xerox our translations to hide lots of tipex from some client's suspicious glare, certainly feel otherwise. Computer keyboards are softer than those of typewriters and offer many more options. Texts can constantly be modified. Now we can offload the rendering we have in mind onto the virtual page on the screen and then comfortably re-read it and ponder whether it is the best option for its wider neighborhood. We now routinely format texts and graphics and can offer much more sophisticated products than 40 years ago. Furthermore, in 2007 Microsoft changed the file formats of

its hegemonic Office products to actually turn them into zip files containing several XML files. Formats for print and screens merged, and overnight our offices became paperless. Many of our translations will actually never be printed out.

Machine Translation, the Holy Grail of processing languages with computers, rides on centuries-old dreams of universal dictionaries, although modern attempts to develop MT started ca. 1930s. In contrast, translation memories were a development of the 1990s. Steps towards developing software that could reuse translations had been taken from the 1970s, but translation memories probably kicked off when Brian Harris' (1988) proposed the notion of *bitext*. Harris' concept was psychological—text units to be aligned were assumed to reflect the ways translators naturally chunk texts—and entailed alignment at several levels (word, phrase, clause, sentence, paragraph), so current bilingual text memories are simplified derivatives, but the applications using them are clearly the CAT tools *par excellence.* Thanks to them, the productivity of many translators has rocketed, consistency is higher, large source texts can be split into sections whose translations can later be seamlessly merged.

MT lagged behind until it finally dropped the notion of reproducing grammar and other language rules in the minds of ideal speakers and focused on statistically imitating texts. A recent advancement, so-called neural MT, predicts likely word sequences in a language based on patterns observed in bilingual and multilingual training datasets. The results are spectacular. Even newer is the tendency to integrate MT within CAT tools, which yields even better results. MT and CAT tools do not exhaust the list of the perks of digitalization. Customizing tools through macros, use of generic but useful apps like CATcount make translators' lives much easier, but there is no more room here to dwell on their advantages, overwhelmingly ignored by my students and novel graduates until they get a job. Instead, we need to focus on some consequences, and on its downsides.

Perhaps we should start by clarifying that professional translators now mix translating, *strictu sensu*, with MT output editing ("post-editing"). However, post-editing seems to be more mentally demanding.[2] We are seeing the same hints in the emerging field of computer-assisted interpreting tools. Rather than making the tasks easier, some supportive technologies seem to make it possible that people other than seasoned professionals will in some circumstances reach acceptable levels of output quality. However, post editors feel the task is more repetitive, more demanding, and worse paid—they tend to be paid per word, and prices do not vary depending on MT output quality. Not that translators are too happy either. In their quest to grab an ever-larger market share, prominent language service providers have sometimes played automation off against translators to keep their heads and their earnings down. Business hype has insisted on naïve sales pitch lines such as human replacement and human parity and have all but pronounced human translation dead. Who of us was never asked whether human translation days are over?

The stubborn reality, however, is that translator hirings are experiencing a sharp rise. The US Bureau of Statistics states that the projected growth in translator employment from 2021 to 2031 is 20%, three times higher than the average. The pitfalls, on the

---

[2] Inconclusive and contradictory results about cognitive effort (demand, load) may be related to researchers' sometimes really measuring different things, but we cannot develop this line here.

other hand, can perhaps be better seen in Europe, which has more translators than anywhere else (about 200,000, nearly 30% of the world's workforce). Here, hybrid working schedules are the new black. Both freelancers and employees of language service companies complain that they cannot reach an adequate work-life balance, even though— or precisely because—half of their work is done at home. Nearly all in-house translators use translation memory tools. Companies tend to impose the use of both MT and CAT tools on their freelancers, and this may help explain why only 12% of freelancers never use them. Actually, more than 70% of independent professionals are using neural MT, although half of them only occasionally. But more than half of the translators using MT are actually using DeepL, a free, open, web-based tool. MT is only used in one out of every four translation projects, whether by companies or freelancers. The truth is, 30% of the European translators in the ELIS survey need to top up their income with a second activity (data from DGT 2022). Just as I did in the mid-1980s. This is happening in a very fragmented industry—70% of translation companies only have between 1 and 10 employees—with good profit margins where translators associate but do not unionize: *homo homini lupus.* Not everything has changed.

We know that machines do not translate. They just process symbols, but they do so incredibly fast and well. They also transmit them consistently in a flash, but they do not communicate. People do. Nobody in their right mind would put the machine-produced version of a sensitive text out without a human reading and approving it. Yet nobody in their right mind would go back to the tipex times. No longing for times past, no room for memories *lost like tears in the rain.* The elder in the tribe have actually contributed to develop these technologies, seen them grow, and grown used to them, and the juniors would not want to work without them, let alone know how. What we need is a different corporate culture, one that does not make translators feel constantly watched over their shoulder by Big Brother (cf. Vieira, Ragni & Alonso 2021), just a tiny cog in a large machine (Moorkens 2020). This is going to demand time and effort, but we need to transform the information society into a knowledge society. Communications (and universities, for that matter) are more necessary than ever before and we need to nurture and pay for them appropriately. We cannot start again from zero. You know, there was no key for zero in my Olivetti Studio 45.

Acknowledgments

References

Bowker, Lynne; Buitrago Ciro; Jairo (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community.* 1st ed.

Bingley, UK: Emerald. <https://doi.org/10.1108/9781787567214>. [Accessed: 20221222].

Directorate General for Translation (2022). *European Language Industry Survey 2022: Trends, expectations and concerns of the European language industry*. European Comission. <https://elis-survey.org>. [Accessed: 20221222].

Fukuyama, Francis (1989). The end of history? *The National Interest,* n. 16 (Summer), pp. 3–18. <https://www.jstor.org/stable/24027184>. [Accessed: 20221222].

Greenstein, Shane (2017). The Reference Wars: Encyclopædia Britannica's Decline and Encarta's Emergence. *Strategic Management Journal*, v. 38, n. 5, pp. 995–1017. <https://doi.org/10.1002/smj.2552>. [Accessed: 20221222].

Harris, Brian (1988). Bi-text, a new concept in translation theory. *Language Monthly,* n. 54, pp. 8–10. <https://www.academia.edu/34281085/Bitext_the_original_article_of_1988>. [Accessed: 20221222].

Hedderich, Michael A.; Lange, Lukas; Adel, Heike; Strötgen, Jannik; Klakow, Dietrich (2021). A Survey on recent approaches for natural language processing in low-resource scenarios. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.* Association for Computational Linguistics, pp. 2545–2568. <https://doi.org/10.48550/arXiv.2010.12309>. [Accessed: 20221222].

Hobsbawm, Eric (1994). *The Age of Extremes: The Short Twentieth Century: 1914–1991.* London: Michael Joseph; New York: Viking Penguin.

Joshi, Pratik; Santy, Sebastian; Budhiraja, Amar; Bali, Kalika; Choudhury, Monojit (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: July 2020. Association for Computational Linguistics, pp. 6282–6293. <http://dx.doi.org/10.18653/v1/2020.acl-main.560>. [Accessed: 20221222].

Khoong, Elaine C.; Rodríguez, Jorge A. (2022). A research agenda for using machine translation in clinical medicine. *Journal of General Internal Medicine,* v. 37, pp. 1275–1277. <https://doi.org/10.1007/s11606-021-07164-y>. [Accessed: 20221222].

Magueresse, Alexandre; Carles, Vincent; Heetderks, Evan (2020). *Low-resource languages: A review of past work and future challenges.* Cornell University. <https://doi.org/10.48550/arXiv.2006.07264>. [Accessed: 20221222].

Mitchell, Amy; Jurkowitz, Mark; Oliphant, J. Baxter; Shearer, Elisa (2020). *Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable.* Pew Research Center. <https://www.pewresearch.org/journalism/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/>. [Accessed: 20221222].

Moorkens, Joss (2020). A Tiny Cog in a Large Machine: Digital Taylorism in the Translation Industry. *Translation Spaces,* v. 9, n. 1, pp. 12–34. <https://doi.org/10.1075/ts.00019.moo>. [Accessed: 20221222].

Moorkens, Joss; Lewis, David (2019). Copyright and the reuse of translation as data. In: O'Hagan, Minako (ed.). *The Routledge Handbook of Translation and Technology*. London: Routledge. (Routledge handbooks in translation and interpreting studies), pp. 469-481.

Nurminen, Mary (2020). Raw machine translation use by patent professionals: A case of distributed cognition. *Translation, Cognition & Behavior*, v. 3, n. 1, pp. 100–121. <https://doi.org/10.1075/tcb.00036.nur>. [Accessed: 20221222].

Orrego-Carmona, David. 2022. Machine translation in everyone's hands—Adoption and changes among general users of MT. *Revista Tradumàtica. Tecnologies de la Traducció*, 20, 322-339. <https://doi.org/10.5565/rev/tradumatica.327>

Pym, Anthony; Malmkjær, Kirsten; Gutiérrez-Colón Plana, Mar (2013). *Translation and Language Learning: The role of translation in the teaching of languages in the European Union: A Study*. Luxembourg: Publications Office of the European Union. <https://termcoord.eu/wp-content/uploads/2013/08/European_Commission.pdf>. [Accessed: 20221222].

Real Academia Española (2021). *El «Diccionario de la lengua española» supera los mil millones de consultas en un año*. Madrid: RAE. <https://www.rae.es/noticia/el-diccionario-de-la-lengua-espanola-supera-los-mil-millones-de-consultas-en-un-ano-0>. [Accessed: 20221222].

Rooksby, Maki; McLeod, Hamish J.; Furuhashi, Tadaaki (2020). Hikikomori: Understanding the people who choose to live in extreme isolation. *The Conversation* (29 october). <https://theconversation.com/hikikomori-understanding-the-people-who-choose-to-live-in-extreme-isolation-148482>. [Accessed: 20221222].

Sanger, Larry (2010). Individual knowledge in the Internet age. *EDUCAUSE Review* 45, n. 2 (March-April), pp. 14–24. <https://er.educause.edu/articles/2010/4/individual-knowledge-in-the-internet-age>. [Accessed: 20221222].

Vieira, Lucas Nunes; Ragni, Valentina; Alonso, Elisa (2021). Translator autonomy in the age of behavioural data. *Translation, Cognition & Behavior*, v. 4, n. 1, pp. 124–146. <https://doi.org/10.1075/tcb.00052.nun>. [Accessed: 20221222].

Zhou, Shu; Zhao, Shuo; Groves, Michael (2022). Towards a digital bilingualism? Students' use of machine translation in international higher education. *Journal of English for Academic Purposes,* v. 60 (November), article 101193. <https://doi.org/10.1016/j.jeap.2022.101193>. [Accessed: 20221222].