Real voices: What translators do and why we need to keep doing it

Nicholas Hartmann
611 N. Broadway, Suite 509
Milwaukee, WI 53202
nh@nhartmann.com

Abstract

Speaking both as an experienced independent technical translator and on behalf of the 11,000 members of the American Translators Association (ATA), the Association's president will explain the training, standards, objectives and working methods of the people who convey meaning from one language into another as their chosen profession. The role of machine translation in that undertaking will be discussed, but the presenter's main objective is to give the translator's view of human translation, describing a process that conveys authentic voices based on much more than words, that in order to be faithful cannot be merely generic and interchangeable, and that gives individual attention to every human utterance. The truth about translators will also be revealed: they are early and eager adopters of language technology when it serves their purpose, and they welcome any opportunity to advance the cause of communication. The goal of this presentation is to contribute to better understanding on both sides, and to help establish a productive and mutually beneficial relationship between MT professionals and the entire community of translators and interpreters represented by ATA.

1 Introduction

This presentation was written in response to an invitation, from Alon Lavie, Laurie Gerber and Mike Dillinger, to speak to the Ninth Conference of AMTA in Denver both as an experienced independent translator and as a representative of the American Translators Association (ATA). As an individual, I have made my living for the last 30 years translating technical and scientific material, mostly from German. My two specialties are patent applications, in which strict accuracy and adherence to the letter of the source text are paramount; and technical marketing, which often shades over into copywriting and demands an ear for tone and nuance. In the one case the principal purpose is to inform; in the other, to persuade. In the mid-1980s—many generations ago in cyber-time—I replaced my Selectric with a computer, and shortly thereafter acquired a fax machine. Ever since then, like most other translators, I have kept up with technology that I believed would help me work faster or more productively.

I am also halfway through my term as President of ATA, which was established over 50 years ago and now has more than 11,000 members in all 50 states and in 90 other countries. We are a broad church that embraces the entire spectrum of players in the translation business: translators and interpreters, teachers of translation, researchers, private- and public-sector language professionals, in-house employees, translation companies and their owners, project managers, and many others. Approximately 70% of our members describe themselves as independent contractors or freelancers. Many different specialties and levels of experience are represented within ATA, along with almost a hundred different language combinations, but it is safe to say that we also have much in common.

I propose to convey to you, both in my own voice and with the combined voices of ATA, some of the attitudes and viewpoints shared by this group of language professionals. I will be frank about our concerns with regard to machine translation; I will also demonstrate that translators possess and exercise unique and unduplicatable skills, and that we must keep exercising them because our work is not just useful but essential. Lastly, I will suggest a future in which our two communities continue doing what each of us does best, for our mutual benefit and in order to serve our ultimate customer, namely human communication.

2 Gripes and nightmares

Before I begin, some unfinished business between us needs to be addressed. In the bad old days about 15 to 20 years ago there was a flare-up of animosity and misunderstanding between human translators and MT professionals, feeding on some overly optimistic claims about MT that had been made decades earlier. Some MT developers believed at that time that FAHQT (fully automatic high-quality translation) was right around the corner; they also seriously overstated MT capabilities in terms of cost and suitability, and were not shy about predicting that human translators would soon be replaced by universal language-processing software and hardware. Some of those human translators, including influential ones within ATA, responded (or perhaps overreacted) by denigrating translation technology in general, reserving special demonization for MT. The upshot was that the two communities went off in their respective huffs and for a long time did not speak to each other; those years of isolation have allowed misconceptions to flourish.

In the last few years ATA and AMTA have begun working at an organizational level to get us back on speaking terms, most recently by co-locating our conferences in Denver so that we could coordinate programs and exchange speakers. This inter-association cooperation is just the beginning; despite the decades of misperception and prejudice, most individual translators have already begun using language technology in a wide variety of ways, and no doubt more will soon embrace the idea of working with the MT community to advance our understanding of how your area of technology may affect what we do in our profession. I believe that from this common ground we can begin to move forward together.

But first there are some things you need to know about the way in which many translators view MT, and about how we regard our own work. Let's begin with a stereotypical translator's nightmare about MT:

1. The entire language services industry shifts almost completely to MT, and all that is left for the majority of human translators is tedious post-editing at burger-flipping wages.
2. A very small group of human translators is still needed, but only for poetry, literature, advertising, and some sensitive business and diplomatic documents.
3. Nobody in category 1 ever has an opportunity to build their experience and skills and move up to category 2, so category 2 translators gradually become extinct and everything ends up being done by machines anyway.
4. On their deathbeds, the last surviving category 2 translators look at what is being cranked out by the computers and feel the same sad emptiness expressed by T.S. Eliot in *The Love Song of J. Alfred Prufrock*: "That is not what I meant at all. That is not it, at all."

Rational and well-informed translators realize that this scenario is exaggerated, but there are nevertheless real differences between our two communities that we must face up to if we are going to have a productive conversation. Some of these issues may not seem important to the MT community, but they are very serious when seen through translators' eyes.

The first is this: with all due admiration for the skill and imagination of MT professionals, there is a significant difference between what MTers and translators define as acceptable translation "quality" in terms of accuracy, word choice, syntax, and general linguistic fidelity. There are good reasons for this: technology people in general seem to be optimistic and risk-tolerant, tending to focus on possibilities and potential rather than on limitations. We translators, on the other hand, have a professional ethos that requires us never to oversell ourselves or our capabilities, or to accept work for which we know we are not qualified. Translators are also exceptionally sensitive to the need to prevent errors and misunderstanding in language. Never mind actual outright mistakes: a translation that merely sounds "off key," or that has an inappropriate intonation or style, grates on our ears like fingernails on the blackboard.

Furthermore, in translators' terms and according to our own carefully considered definition of how we think our work ought to be done—and, much more importantly, in the view of many discerning clients who realize the central importance of excellent translation—MT has not demonstrated equivalence with, let alone superiority to, human translators except in a very few contexts.

One such context is when some translation (as opposed to no translation at all) of large volumes of text must be produced in a short period of time, either to serve vital strategic or economic needs, or in instances where expectations are low and a market or user would otherwise remain entirely unserved. Even then, we believe that such output is usable only for "gisting," and that a "real" translation of the truly interesting material can be produced only by humans.

It is true that highly controlled linguistic environments can be established in certain cases (such as the automobile industry), and that restrictions and restraints that are impractical elsewhere can be enforced in such environments, for example limited source-language vocabulary, strict composition rules, and careful pre- and post-editing. Within these walled gardens a custom-designed MT system can yield a useful and understandable result—but it remains a fact that outside them, in the real world, almost everything that has to be translated is written not in any "controlled language" but in real and unpredictable Spanish or Chinese or English.

We are also fundamentally skeptical about the term "machine translation" itself, since what we see most MT systems doing is more like "information-technology-assisted language transfer based on collective integrated terminology management and correlation." We share the view of Martin Kay that language is much more than "... a probability distribution over strings of letters or sounds." We therefore believe that most MT systems perform, or at worst merely simulate, only one of a very large number of activities performed by human beings when they translate, namely the discovery and comparison of terms and phrases based on large volumes of previously translated material—that material having been previously translated, by the way, by us.

We vigorously dispute the notion that MT is a universal solution. Every individual professional and craftsman acquires different tools and develops unique ways of using them. No computer desktop, no carpenter's toolbox, and no researcher's bookshelf is identical to any other. Some MT promoters, however, would have us believe that their system—their tool kit—can be used with equal effectiveness by anyone. We often counter this idea by contrasting a lumberjack wielding his chainsaw (an MT system for gisting large volumes of text very quickly) with a surgeon and her scalpel (a human translator with experience and sensitivity in a particular language and subject): each can be optimally effective in the right context but inefficient or lethal in the wrong one. And while an MT system can function interchangeably, like a vending machine that responds to your coins with a candy bar no matter who you are, human translators are (or aspire to be) unique designers and providers of specialized intellectual services to specific clients, who have equally specific messages to convey.

3 What do translators do?

What exactly do we human translators do that makes us so special?

Translators don't translate words, because real communication is not just about words. For example, a sign on the fence around a construction site in Italy might read: "L'ingresso è vietato ai non addetti ai lavori." The words of this text can be "correctly" translated as "Entry is forbidden to those not associated with the works," but what the sign really means is "Unauthorized entry prohibited." Or consider this invitation in a tourist brochure for a small town in Québec: "Arpentez la rue principale et découvrez la nature amicale et chaleureuse des résidents." Literally this would be "Stroll the main street and discover the friendly, welcoming nature of the local residents." But my colleague Grant Hamilton has turned this into an English sentence that is much more effective at drawing in the reader: "Soak up the cozy friendliness of small-town Québec with a walk through the old town."

Two things are worth noting: First, these two sentences, which are so similar in information content and so different in flavor, share only a single fragment of a single non-trivial word ("friend"). Second, this is writing that is optional and discretionary: nobody absolutely must read it. A machinist cannot put a new lathe into service without reading the manual, but the reader of the Québec brochure is just passing by, and will step into the text only if its flow and rhythm are appealing. This is a piece of writing that has to persuade as well as inform: if it fails to do either, a potential customer will go elsewhere and the translation is a waste of money at any price. (Cynics have noted that while human translation is considered essential in order

to make a sale, in certain markets MT seems to be entirely "sufficient" and "economically sensible" once the sale has been made.)

Even when the purpose of a text is simply to convey information, automatic lookup and matching can produce words that only seem correct. A human utterance that looks simple can in fact get very complicated: English phrases containing "can," "get," "do," "make," "still," "right," and thousands of other chameleons are ubiquitous, easily misunderstood even by humans (especially those who lack a deep and broad knowledge of the language), and impossible to translate without an equally broad consideration of context. Similar subtleties lurk in every other language, even in perfectly straightforward texts. For example, I recently translated a German technical document in which the words "auslösen" and "durch" each took on two different meanings: "durch die Öffnung ausgelöst" means "dissolved out through the opening," but "durch die Steuerpulsfolge ausgelöst" means "triggered by the control pulse sequence." A computer program that "knows" only one translation for "durch X ausgelöst" has at best a 50% chance of being right; a very careless or inexperienced human translator might also confuse the two, but the incongruity of the "wrong" translation would be immediately evident to anyone minimally qualified to work in that language and subject field.

The problem goes far beyond words. Here is another purely descriptive and unsubtle sentence, again about a technical system: "Wie die Sendeeinheit mit einem im Inneren einer Geschirrspülmaschine, insbesondere in einer Geschirrschublade angeordneten Dosiergerät zusammenwirkt, wird nachfolgend an Hand der Figuren erläutert," meaning "The manner in which the transmitting unit interacts with a dispensing device arranged in the interior of a dishwasher, in particular in a tableware rack, will be explained below with reference to the Figures." What a popular online translation tool produces, however, is this: "As the sensor is cooperating with an inside of a dishwasher, especially in a dish drawer arranged dosing, subsequent to the figures explained"—a "translation" that an expert reader of German would not understand without referring to the source text, and that would be completely useless to a monolingual engineer or scientist. Human translators simply do not make these kinds of mistakes.

We further disagree with the idea that when a source text is an exact match with a previously translated passage, it can be translated correctly into exactly the same target text the next time around. Not necessarily: the content and syntactic context may seem identical, but the client may be different, the underlying technology may have changed, or a new attorney may simply have other preferences. It can be argued that MT works very well for boilerplate, and it does—provided the boilerplate hasn't been repurposed and the expectations for the translated boilerplate have not changed. So the claim that translation technology can "leverage" from past translations to future ones holds true only if <u>everything</u> about both the source and target environments—context, author, register, intent, audience, and much more—is identical. Such instances are far less common than some boosters of MT (or indeed of TM) would have us believe. As a result, we contend that MT systems based on translation memories and corpora routinely violate a basic principle, namely that "past performance is no guarantee of future results."

Translators also have a rigorous and not necessarily quantitative concept of "accuracy." Consider a one-paragraph medical report that has been translated into English. The concluding phrase of the translation reads: "..., the review group unanimously recommends that Mr. N not be subjected to a complete frontal lobotomy." If that entire 100-word paragraph contained only a single error involving one letter of one word, the translation could still be considered more than 99% accurate; but if the error resulted in the word "not" being changed to "now," the result would also be 100% wrong. If the person sending the translated report to the surgeon were unable to read English, that tiny but pivotal mistake would go completely undetected. As ATA has been saying for decades, to translation users and to the entire world, "mistakes can be costly" in terms of both money and human lives, and they are especially costly if you don't know where they are or that they even exist. In almost any context that matters, any MT-generated text that is not reviewed and corrected by a competent human can produce results that are just as quantitatively "accurate" and just as expensively or lethally wrong.

We have just seen that clients who purchase translations into a language they do not know are in the unpleasant position of being completely un-

able to judge the quality of what they are buying. Here is another illustration of the fact that when it comes to translation, what you don't know can definitely hurt you (or at least make you look very foolish): A few years ago, a town in northern Wales wanted to put up a new sign (bilingual in English and Welsh, as required by law) to redirect truck traffic away from a housing development. The message decided on was: "No entry for heavy goods vehicles / Residential site only." But because the department in question did not have a Welsh speaker in-house, they e-mailed the English text to a Welsh translator. What they got back was "Nid wyf yn y swyddfa ar hyn o bryd. Anfonwch unrhyw waith i'w gyfieithu." Only after the bilingual sign had been expensively manufactured and installed did they discover that the Welsh actually meant "I am not in the office at the moment. Send any work to be translated." Any unedited MT output into an unknown language can conceal similarly explosive mistakes.

We therefore assert that human perception of context and "rightness" is necessary not just for poetry and literature and advertising, but in every aspect of communication; and that real translation, which conveys genuine meaning, is a literally indescribable human activity and can only be simulated, not reproduced, by any system less complex than the human brain.

Every text is ultimately a dialog between brains, between a human author and a human audience, and translators have an ethical and professional obligation to be true to both of them. Our authors and audiences are, again, everywhere:

- the chamber of commerce of a picturesque village, attracting and welcoming visitors,
- an urban transit system giving guidance to Spanish-speaking riders, and doing so with a tone and register equivalent to that used for English speakers,
- a manufacturer of precision manufacturing equipment, making its sales case to potential customers at just the right level of technical sophistication,
- a pharmaceutical company preparing a package insert that not only instructs patients in the proper and safe use of its products but also gives clinical information to physicians and pharmacists, taking into account each audience's capabilities and expectations,

and many more in every field of human activity.

It is not enough to reproduce words and phrases accurately, although we have seen how difficult even that can be. Beyond that basic obligation, and especially in contexts where register and tone are crucial elements of what is being conveyed, failure to convey the authentic voice of the originator of a text constitutes distortion, scratches on the record, defects in reproduction that distract readers or hearers from the real message and may alienate them entirely.

4 What do translators really want?

We want to work together constructively with everyone who can contribute to the cause of genuine communication between people.

Believe it or not, we really do want technology. Translators are not inherently "anti-technology"; in fact, because we are on the whole intelligent and inquisitive people, we are often early adopters of appropriate technological tools. Even translators working at the highest level of our craft, meeting the most sensitive demands in terms of faithful communication between author and audience, routinely use technology such as computerized glossaries, translation memory (sometimes, and carefully), and web-derived terminology (increasingly often, but even more carefully).

Machine-assisted translation is therefore firmly entrenched among translators, and we love it. Almost every human translator already uses a computer to automate terminology management, finances, communication, and much more. MAT can be, and already is, effectively and efficiently used by translators for many different purposes at many different points along a long gradient of automation. This is technology that works for us: we select our language-technology tools because they enhance our productivity and improve our work.

We emphatically do not want tools that make our job harder. Technological progress is supposed to create tools that enhance the productivity of skilled professionals: CAD/CAM systems for aircraft designers, imaging systems for physicians, even (unfortunately) quantitative models for investment managers. Full-bore MT, however, is often experienced by translators as an impediment to be overcome rather than a positive asset: for many of us, "post-editing" has become code for "Find all the mistakes the MT system made and get

paid peanuts for correcting them, even though that takes much longer than translating it right in the first place." Only in the translation business is it considered acceptable for a new technology to introduce errors that professionals then have to spend their time fixing.

We want tools that are made for us, not for people without training in language. Translators would like greater access to tools developed with translators in mind, since we are uniquely qualified to understand how such tools work and how to use them safely and effectively. Only in the translation business is it assumed that new technology can and should be used largely by end users in order to bypass professionals. No responsible software developer would devise an online tool for reading your own X-rays, and then promote it as a way for patients to cut out those expensive and finicky radiologists. But translation mistakes can be just as deadly and costly and embarrassing, and can affect far more people.

We also want tools that we have helped to make. Translators will remain essential to MT development for the foreseeable future, because there are lots of tasks that can be done best by people who possess our skills: translation memories must be continuously refined and updated, glossaries specific to a particular industry or audience should be created by translators who know that industry or audience, and pre-editing and post-editing demand expertise not only in subject matter but in language. We can play a valuable role in the development of new tools, applying our linguistic expertise to ensure that humans and technology function most effectively together.

One particularly vital task, however, may be to help update and detoxify the bitext corpora that are the foundation of statistical machine translation. To quote a prominent member of AMTA, "the quality of a 'data-driven' MT system is highly dependent on the quality of the translations on which it is 'trained'." But if the nightmare comes true and all the human translators do in fact die out, MT systems will end up being able to "learn" to translate only from other MT systems, and in the absence of fresh input of high-quality examples translated by real translators, errors and misinterpretations will accumulate to the point where the results become unusable. As Mike Dillinger once vividly described it, without constant input from expert human translators, any statistical MT system will eventually be poisoned by its own waste products.

5 Where to now?

Leaving aside our differences of opinion, the relationship between translators and the MT community can become mutually beneficial, provided we both do our part:

- We need to keep one another honest. Translators will keep insisting that translation is a human activity that enables communication between real and very specific authors and audiences, who in turn are not generic or interchangeable, but each require individual attention to their individual natures and needs.
- We hope that you in turn will be straightforward and realistic about the narrow range of MT contexts in which human translators really cannot do anything useful, or in which we really would not want to participate: we believe in the special nature of what humans do, but we do not seek out drudgery.
- Translators already extensively use (and will continue to use) computer technology, up to and including interaction with fully automatic MT as essential participants. We hope that human researchers will continue to develop ways to adapt technology to assist human translators and expand our capabilities, and we welcome the opportunity to streamline those aspects of our work. You are the specialized manufacturers here, and we are the power users; an effective dialog between us will inevitably improve the product.
- ATA as an organization puts a great deal of time and effort into education—of our members, of translation users and of the general public. We must all continue the process of education within and between our respective communities, and extend that mission to translation buyers and to society at large.

Above all, we need to keep talking to each other. The encounter in Denver between our two associations was illuminating and mutually enriching, and I hope it will set the tone for a future in which we no longer work at cross-purposes.

6 Not The End

Despite dire predictions from some, this is not the beginning of the end for translators, not a slow slide into obscurity along with the manufacturers of horse-drawn carriages and typewriter ribbons. It is instead a particularly exciting time to enter this profession, because the prospect of fruitful and positive interaction with the MT community opens up a new and much wider range of opportunities to interact with language and acquire new capabilities.

In order to prosper in the future, we translators of course must continue to develop and maintain our unique linguistic skills, but we must also constantly refine our awareness of how technology can be used to enhance those skills even further. We look forward to developing new ways for both communities to serve the cause of human communication, since we are all in the same business.

Acknowledgments

This paper has benefited greatly from discussions with my ATA colleagues Lillian Clementi, Chris Durban, Grant Hamilton, Kevin Hendzel, and Dorothee Racette, and with Mike Dillinger and Laurie Gerber of AMTA. Any errors or misunderstandings that it may still contain are mine.