Preisträger des Berichtsjahres 2022

Mit der Lichtenberg-Medaille 2022 wurde DOUGLAS HOFSTADTER aus den USA ausgezeichnet.

Douglas Hofstadter

Understanding versus Blunderstanding: Comparing Human Translation with Machine Translation

Douglas Hofstadter, zugeschaltet in der Öffentlichen Sommersitzung der Akademie 2022

Shortly after receiving my doctorate in physics, I spent the years 1975–77 at the Institute for Mathematical Studies in the Social Sciences at Stanford, where I had a lively German officemate named Wilfried Sieg, who was a doctoral student in philosophy. We became friends, and one day Wilfried told me about the eighteenth-century German physicist and aphorist Georg Christoph Lichtenberg. I was intrigued and looked Lichtenberg up in the *Encyclopaedia Britannica.* Among other things, I learned of so-called „Lichtenberg figures", which are generated by the passage of electricity through nonconducting substances. They are sometimes picturesquely called „captured lightning" or „electron trees". Here is a photograph of a typical marvelous Lichtenberg figure:

By chance, my father had such a figure made of lucite, which sat on a desk in his study and looked very similar to the one above. Over the years, I'd seen that object many times, but had never known what it was. It turned out that Lichtenberg had discovered this phenomenon in the year 1777 in Göttingen. By sheer coincidence, I was learning about it in the summer of 1977 – exactly 200 years later – and so, on a whim, I decided to throw a 200th birthday party for the Lichtenberg figure, and I scheduled it for the 1st of July, which was Lichtenberg's own birthdate. It was a tiny but jolly party, the only guests being my parents and Wilfried. I ordered a cake at a local bakery, and asked them to approximate a Lichtenberg figure on the icing. That was fun!

As an extra treat for my three guests, I read aloud a handful of witty aphorisms composed by Lichtenberg, which I'd found in the *Encylopaedia Britannica* article, including the following ones:

Die Fliege, die nicht geklappt sein will, setzt sich am sichersten auf die Klappe selbst.
A fly that doesn't want to be swatted is safest if it sits on the flyswatter.

Heutzutage machen drei Pointen und eine Lüge einen Schriftsteller.
Nowadays three witty turns of phrase and a lie make a writer.

In unsern Zeiten, wo Insekten Insekten sammeln, und Schmetterlinge von Schmetterlingen schwatzen.
In our day, insects collect insects, and butterflies gossip about butterflies.

Die gefährlichsten Unwahrheiten sind Wahrheiten mäßig entstellt.
The most dangerous of all falsehoods is a slightly distorted truth.

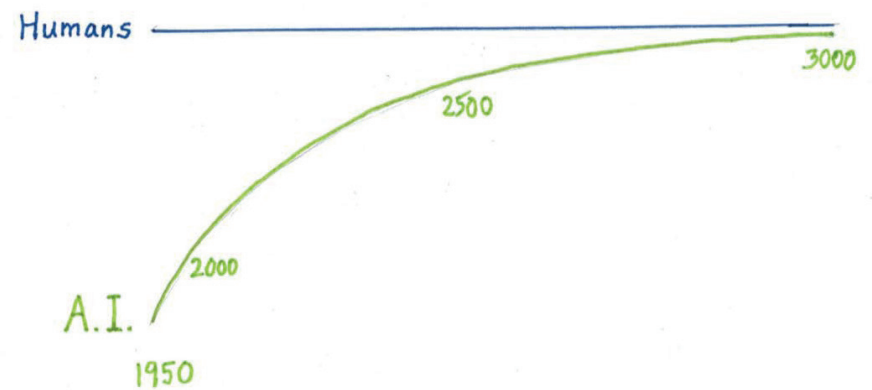
Ein Buch ist ein Spiegel, aus dem kein Apostel herausgucken kann, wenn ein Affe hineinguckt.
A book is like a mirror: if an ape peers into it, an apostle is hardly likely to peer out of it.

Gerade das Gegenteil tun, heißt auch nachahmen.
To do precisely the opposite is also a form of imitation.

I had avidly studied German in college and later on had lived in Germany for six months, so I felt it was only logical to carry out the English translations myself. And this brings me to the main topic of this Lichtenberg Lecture, which involves comparing *human* translation with *machine* translation (the latter sometimes being called „MT"). The key question I wish to explore here is this: *Is MT empty?*

In order to broach this issue, I feel it necessary to say a few words about artificial intelligence in general, and more specifically, about my personal involvement with it, which started during the early 1970s. In those bygone days, I romantically saw the human mind as quasi-magical, and AI as a noble but probably hopeless quest. AI was, after all, an attempt to force *rigidity* (that of any machine) to mimic *fluidity* (that of a mind). If you will permit me a caricature analogy, it was as if someone were trying to use a Meccano set to build an octopus that can swim.

In short, in the 1970s, I conceived of AI as a grand philosophical adventure – highly romantic, but also highly quixotic. Below I exhibit a graph that will give you a clear sense of how I personally imagined AI would develop over the course of time:

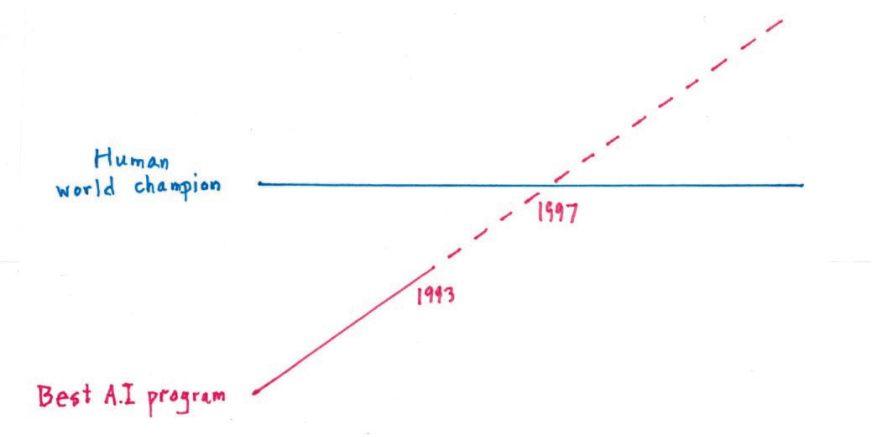


I can translate the graph into words. It expresses my early feeling that human-level intelligence was an unattainable goal; it was simply the asymptotic limit that would be slowly approached (from below) during hundreds (or perhaps even thousands) of years of attempts to force a „square peg" (mechanical rigidity) into a „round hole" (organic fluidity).

In my 1979 book *Gödel, Escher, Bach: an Eternal Golden Braid,* I devoted two chapters to AI, and several pages to the huge challenge of machine translation, using Lewis Carroll's famous nonsense poem *Jabberwocky* as translated by brilliant human beings into both French and German, in order to highlight the immense challenges that are posed by translation.

In the subsequent decade (the 1980s), an immense number of nonsensically exaggerated claims about AI f lourished, so in the 1990s, I gave two seminars at Indiana University called „Hype versus Hope in AI“ („hype“ being short for hyperbole, or in other words grotesque exaggeration), in which I mostly debunked the extreme claims.

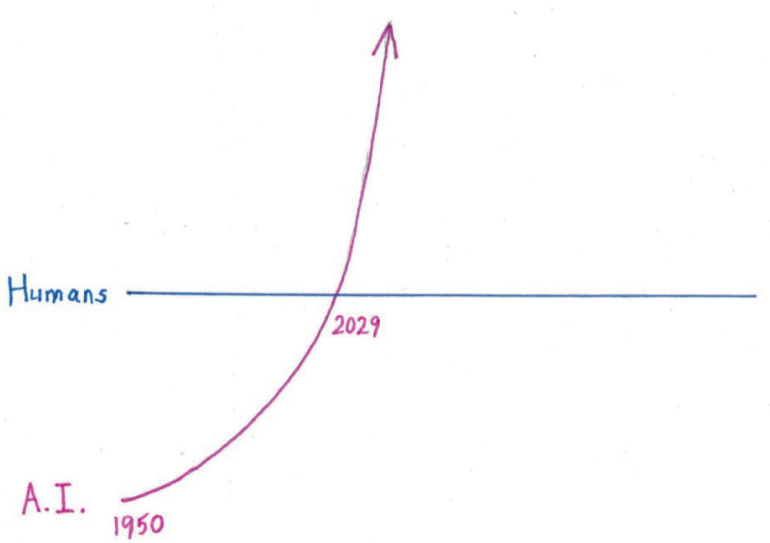
In 1993, I read an article about computer chess in *Scientific American*, and it featured a graph showing how computer chess had progressed over the decades, and also how it was likely to progress in coming years. The graph really shocked me. Below, I have drawn what I remember about that graph. The *solid* line shows the playing level (essentially the F.I.D.E. rating) of the best chess programs until 1993, plotted against time. The *dotted* line shows what the article's authors believed would transpire in the near future.



What actually happened in the ensuing years? Well, sure enough, in 1997, exactly as predicted, Garry Kasparov, the then-reigning world chess champion, was defeated in a tournament by IBM's massively fast chess program called „Deep Blue“. Although I knew that Deep Blue was not doing anything comparable to what great human chess players do in their minds, I was nonetheless surprised, shocked, and quite scared, because this did not match up at all with my own personal intuitions about how AI would evolve.

Of course, some people thought that Deep Blue's triumph over Kasparov was a *triumph for humanity*, since, after all, humanity itself had created Deep Blue. I well understood this point of view, but I didn't agree with it. Instinctively, I felt the human mind *should* be unattainable, and I rebelled against the idea that it could be defeated by a mechanical creation, even if that creation had come from human minds.

Then, two years later, in 1999, a pair of books came out, both of them chock-full of surreal predictions: *The Age of Spiritual Machines* by Ray Kurzweil, and *Robot: Mere Machine to Transcendent Mind* by Hans Moravec. These books excitedly spoke of the so-called „Singularity", which was the hypothetical result of exponentially accelerating AI progress, according to which AI would surpass the human mind's level in roughly 2029 A.D. This vision is roughly captured in the following graph:



The mood of excitement that these authors were attempting to transmit to their readers did not resonate at all with me. In fact, it *terrified* me, rather than exciting me. At the same time, however, I felt that the predictions, especially those made by Kurzweil, were based on very little evidence. He was quite a clever fellow, but I often felt that he was talking through his hat rather than making serious arguments. About his book, I once stated: „It is like what you would get if you were to dump equal amounts of a fancy gourmet meal and of dog excrement into a food blender, then were to turn it on and mix them intimately together."

So I was both skeptical and frightened – a very strange mixture of emotions. Kurzweil's and Moravec's visions of the „Singularity" evoked great worry on my part, and not long after reading their books, I was so worked up that I organized two large symposia dedicated to exploring the astonishing and wild-sounding idea of the Singularity – the first one at Indiana University in 1999, and the second one at Stanford in 2000. Both Kurzweil and Moravec participated in the Stanford symposium, but to my disappointment, neither of the symposia resulted in much clarity about what was going to transpire in the next few decades.

Around 2010, a new computational technology, commonly called „Deep neural nets“ (DNN) or „Deep learning“ (DL), started to come into its own. This technique used vast numbers of computers running in parallel, and it trained them on inconceivably large data bases.

I didn't pay much attention to these „deep“ developments, and in fact I was quite annoyed at their use of that adjective, since it was a kind of play on words. The neural networks employed were not just a *couple* of layers deep, as early neural networks had been – they were now *several* or *many* layers deep – and thus they were „deep“ only in that limited technical sense. But the mere act of calling the networks and the learning „deep“ automatically evoked the connotations of *profundity* that the word ‚deep' carries along with it. Exploiting this ambiguity of the word ‚deep' struck me as rather dishonest and cheap, but by then, of course, I was familiar with that kind of thing in the world of AI.

In 2016, in *The New York Times Magazine*, I read an article that described, using an incredible amount of hype, the brand-new *deep-learning* version of Google Translate. My reaction to this article was, once again, a strange and subtle blend of skepticism with fear.

As I stated above, high-quality translation had always struck me as being a shining case of the human mind's flexibility, and in fact, in my writings, I had often mocked machine translation, citing all sorts of MT failures that I myself had discovered. But now, I wondered, had MT crossed some kind of magical threshold, and was it now going to give even the very best of human translators a run for their money? I was very scared. I really didn't want to see the human mind humiliated.

The only way to confront my fear was, needless to say, to *test* the program. And so I dreamed up a set of tricky challenges for the new version of Google Translate.

Here is a simple but lovely example. The „his“/„her“ distinction in English does not have a natural, easy counterpart in French, since the French possessive adjectives „son“, „sa“, and „ses“ agree in gender (and number) with the *possessed item*, not with the *possessor*. So, I wondered, what if I were to feed the following English sentence to Google Translate and then ask it to render it in French?

> *In their house, everything comes in pairs. There's his car and her car, his towels and her towels, and his library and her library.*

What is the gist, the crux, the essence, of this sentence? Of course, to any human reader, it's self-evident. It is the contrast between *his* stuff and *her* stuff. Not to see that would be to miss the point entirely. My own French translation of my challenge sentence ran as follows:

> *Chez eux, ils ont tout en double. Il y a sa voiture à elle et sa voiture à lui, ses serviettes à elle et ses serviettes à lui, sa bibliothèque à elle et sa bibliothèque à lui.*

Are we seeing *understanding* here? Yes, of course. I used my lifelong intimacy with French – dare I call it „my *deep* intimacy with French"? – to find a subtle trick allowing me to express the intended meaning.

But actually, if you look carefully, you might note that there's something amusing in my translation, and I wasn't even aware of it when I came up with it. In particular, the original English sentence mentions first the *man's* possessions, then the *woman's*. But in my French version, the *woman* is mentioned first, then the man. Of course that „infidelity" makes no difference whatsoever to the sentence's *essence*, and only an obsessive nitpicker would say that the translation was flawed. In fact, this nearly-invisible reversal of genders beautifully underlines the enormous flexibility of the human mind. The order of „his" and „hers" was unintentionally flipped by the human translator – myself – but it resulted in no change of meaning. How ironically curious!

Now let's look at Google Translate's 2017 translation of my sentence.

> *Chez eux, tout va par paires. Il y a sa voiture et sa voiture, ses serviettes et ses serviettes, et sa bibliothèque et sa bibliothèque.*

Are we seeing *understanding* here, on the part of the „deep" translation engine? No; what we're seeing is *blunder*standing, in the sense that using „sa" and „sa" (or „ses" and „ses") together is a grotesque blunder. After all, in the mechanically-created French sentence, there's no indication that what's under discussion is the contrast between a man and a woman!

As the infamous Italian sound-bite says, *Traduttore, traditore*, meaning that *every translator is a traitor*, a notion with which I vehemently disagree – but in this particular case, it happens delightfully to hit the nail on the head. (I will return to the translator-bashing sound-bite a little later.)

That year, I wrote an article for *The Atlantic* magazine, in which I used this provocative example, along with several other egregious failures of Google Translate (involving French, German, and Mandarin Chinese), to demonstrate the inadequacy – the deep inadequacy, I daresay – of this mechanical approach to a challenge that inherently involves *thinking* and *understanding*. (The article came out early in 2018.)

In 2018, I learned of DeepL, a Cologne-based MT program using (as its name suggests) deep neural networks and deep learning, which was reputed to be far superior to Google Translate. As you might guess, my reaction, on hearing of DeepL, was a new round of skepticism mixed with fear. Let's take a look at its response to my „his/her" challenge:

> *Dans leur maison, tout va par paire. Il y a sa voiture et sa voiture, ses serviettes et ses serviettes, sa bibliothèque et sa bibliothèque.*

Once again, I feel compelled to ask, „Are we witnessing *understanding* here?" Well, no; in fact, as before, we're seeing only *blunder*standing. (By the way, I'm delighted to report that, as of June 2022, both Google Translate and DeepL are *still* translating this sentence into French in exactly the same robotically blunderstanding fashion.)

If I ask either Google Translate or DeepL to *pronounce* the original English sentence (a feature that they both boast of), here's what I hear:

> *In their house, everything comes in pairs. There's his car and her car, his towels and her towels, and his library and her library.*

As you see, both translation engines, instead of putting the stress on the possessive pronouns „their", „his", and „her", put stress on the four different *nouns*. What could more convincingly demonstrate that there is not the least iota of understanding here? Rather, what we hear reveals *blunderstanding*. Although every single *word* is enunciated perfectly, the sentence as a *whole* sounds weirdly unhuman and robotic.

In 2017, I myself translated the last few pages of the 1848 novella *Der arme Spielmann* by the Austrian writer Franz Grillparzer. It took me roughly eight hours to do so. I then tested Google Translate on the same text. It took GT all of *15 seconds* to do it. That's a ratio of about 1920 to 1. Does this mean GT is *1920 times more skillful* a translator than I am? Alternatively, does it mean that for it to be worth it for a human to translate a passage, the human-produced text would need to be *1920 times more polished* than the machine-produced text?

Well, in order to give you a tiny sense for this question, let me first exhibit the final paragraph of *Der arme Spielmann* in the original German:

> *Ihr Gesicht war dabei von mir abgewandt, so daß ich nicht sehen konnte, was etwa darauf vorging. Da nun zu gleicher Zeit die Magd mit der Suppe eintrat und der Fleischer, ohne sich durch den Besuch stören zu lassen, mit lauter Stimme sein Tischgebet anhob, in das die Kinder gellend einstimmten, wünschte ich gesegnete Mahlzeit und ging zur Tür hinaus. Mein letzter Blick traf die Frau. Sie hatte sich umgewendet, und die Tränen liefen ihr stromweise über die Backen.*

And here is the final paragraph as rendered by myself in English:

> *At that moment her face was turned away from me, so I couldn't see her expression. Right then, the daughter brought out the soup, and the butcher, not wanting to let my visit interfere, started intoning the dinnertime grace, and the two children loudly joined in. Sensing I was out of place, I simply wished them a blessed meal and walked out the door. My very last glance was of the wife's face. She had just turned around, and I could see tears streaming down her cheeks.*

And here is the final paragraph as rendered by Google Translate:

> *Her face was turned away from me, so I could not see what was going on. When, at the same time, the maid entered the soup, and the Fleischer, without letting himself be disturbed by the visit, raised the banquet in a loud voice, into which the children agreed, I wished for a blessed meal and went out the door. My last glance met the woman. She had turned around and the tears ran down her cheeks.*

Although there are all sorts of details I could talk about here, I will focus down on just Grillparzer's German phrase „…*die Magd mit der Suppe eintrat*…" and the two rival translations of that phrase, because both of them involve errors, but different *types* of errors. (As Georg Lichtenberg once observed: „Wir irren allesamt, nur jeder irret anders.")

My translation said „*the daughter brought out the soup*", which is a perfectly plausible scenario, although it's almost certainly wrong. The thought process that led me to this vision was that „die Magd" struck me as an old-fashioned way of referring to a young girl, and the only girl in the scene was the daughter, so my conclusion was that she was serving the soup. Why not? To me it made pretty good sense. Also, in the most reliable of all my German–English dictionaries, among the translations of „Magd" was „maiden", which seemed to support this vision.

Later, though, when I saw Google Translate's phrase „the maid", I did a very painful double-take. Suddenly filled with shame, I thought to myself, „Of course! There must have been a *maid* in the household! How silly of me not to have thought of that!" And to add insult to injury, it had been Google Translate, of all „people", that had shown me the error of my ways. *That* was really a tough pill for me to swallow.

In my defense, I can argue that no housemaid or servant of any sort had appeared earlier in Grillparzer's story; moreover, servants are a long-gone part of the culture in which I grew up and live. But those are weak excuses. I should simply have remembered that the main meaning of ‚Magd' is ‚housemaid' (or just ‚maid'), and not ‚maiden' or ‚girl'. I had made a dumb error, and I just had to accept that as a fact.

Okay, but what about Google Translate's error in the same spot? Well, let's look. The highly touted MT system came out with the phrase „*the maid entered the soup*". Unlike me, it got the *maid* right, but now it has her *entering the soup*. What on earth does that mean? If you attempt to envision it, you have to imagine either a soup bowl as big as a swimming pool, or a maid as minuscule as a saltshaker – and if you do visualize either one of those scenarios, you can only laugh, because what in the world would have *motivated* this maid, or *any* maid, to „enter the soup"? How could Grillparzer have wanted us to imagine such a crazy scene?

In contrast with my error, which conjured up a *plausible* scenario, Google Translate's error conjured up an utterly implausible, *nonsensical* scenario. There's no comparison between the magnitude of these two errors. Google Translate did its job dazzlingly much faster than I did mine, but if it is going to make outrageous blunders like this, then you can't trust its output further than you can throw it. It just makes a fool of itself in a fantastically short time.

This is just one of literally hundreds of examples of blunderstanding committed by Google Translate in the several pages of *Der arme Spielmann* that I fed to it. Some of the examples were far more ridiculous – and far more serious – than this amusing blunder, but I don't have the time to exhibit and discuss them here. *Ars longa, vita brevis.*

By the way, since I just quoted Lichtenberg's aphorism „*Wir irren allesamt, nur jeder irret anders*", I think it would be interesting to show the difference between a human translation thereof (my own) and Google Translate's. In my version, the aphorism runs thus: *We all blunder, but each of us does so in a unique fashion.* By contrast, the 2022 version of Google Translate renders it as follows: *We are all wrong, only everyone is wrong differently.* Well, that's a nice try, but it's still just blunderstanding. The first half, in particular, is way off base. Lichtenberg didn't mean that all of us are *always* wrong; he merely meant that we all err *now and then.* And Google Translate kindly (although by accident) provided an example!

I would now like to proceed to a final case study – the challenge of translating into English the poem *Nähe des Geliebten*, penned in 1795 by Johann Wolfgang von Goethe. Here is the poem:

Ich denke dein, wenn mir der Sonne Schimmer
Vom Meere strahlt;
Ich denke dein, wenn sich des Mondes Flimmer
In Quellen malt.
Ich sehe dich, wenn auf dem fernen Wege
Der Staub sich hebt;
In tiefer Nacht, wenn auf dem schmalem Stege
Der Wandrer bebt.
Ich höre dich, wenn dort mit dumpfem Rauschen
Die Welle steigt;
Im stillen Haine geh ich oft zu lauschen,
Wenn alles schweigt.
Ich bin bei dir, du seist auch noch so ferne,
Du bist mir nah!
Die Sonne sinkt, bald leuchten mir die Sterne.
O wärst du da!

And now, here are two machine translations from 2022 (treating the poem as prose, since that's all they can do). First, by Google Translate:

I think of you when the sun shines from the sea; I think thine when the moon's flicker paints itself in fountains.
I see you when the dust rises on the distant road; in the deep of night, when the wanderer trembles on the narrow footbridge.
I hear you when the wave rises there with a dull roar; I often go to listen in the quiet grove when all is silent.
I am with you, no matter how far away you are, you are close to me! The sun is sinking, soon the stars will shine on me. Oh, were you there!

And now, I give the floor to DeepL:

> *I think of you when the sun shines from the sea; I think of you when the moon's flicker paints itself in springs.*
> *I see you when the dust rises on the distant road; in the deep night when the wanderer trembles on the narrow path.*
> *I hear you when the wave rises there with a muff led roar; in the quiet grove I often go to listen when everything is silent.*
> *I am with you, however distant you may be, you are close to me! The sun sinks, soon the stars shine for me. Oh, if you were here!*

Hmm… These are both semi-decent, albeit slightly clunky, *literal* translations – quite in the style of contemporary American academics, who, alas, long ago abandoned the art of translating poetry into *poetry*.

By contrast, I would like now to show the translation that I myself crafted (very slowly and very carefully), in which I treated the poem *as a poem*. What does that mean? It means that I respected both the poem's *content* and its *form*.

In terms of form, *Nähe des Geliebten* consists of four *quatrains*, each of which features a long line followed by a short line, and then another long line followed by a short line. The eight *long* lines are all written using *iambic pentameter*, and they end in *feminine* (bisyllabic) rhymes (*e.g.*, „ferne"/„Sterne"), while the eight *short* lines, written using *iambic dimeter*, all end in *masculine* (monosyllabic) rhymes (*e.g.*, „strahlt"/„malt").

To my mind, the act of translation involves preserving the *spirit* rather than the *letter* of the original text. (That's why I wasn't in the least bothered by my own unintended reversal of genders in the „Everything comes in pairs" sentence.) More specifically, the act of *poetry* translation involves respecting form just as much as content. Doing so necessarily involves *molasses-slow, thoughtful, mental flexibility* as opposed to *lightning-fast, thoughtless, elemental reflexivity*. The name of the game is *compromise*. So, without further ado, here is my rendition of *Ich denke dein*:

I think of you, whene'er the sunlight's glimmer
On ripples breaks;
I think of you, whene'er the moonlight's shimmer
Reflects off lakes.
I see your face when, o'er a trail through ridges,
A dust cloud forms;
Or when some pilgrim, crossing moonlit bridges,
Fears far-off storms.
I hear your voice so clearly in the surging
Of waves that rush;
I sense, in silent groves, your voice emerging
From midst the hush.

Despite the miles, your soul with mine's entwining;
You seem so near!
The sun's now low, the stars will soon be shining.
Were you but here!

So many changes – some very big and some very small – but big or small, changes lurk everywhere! And this fact brings me back to the tempting Italian sound-bite *Traduttore, traditore.* As I said earlier, this cute little phrase greatly irks me, because it suggests that all translators are constantly betraying the authors who they are translating. I am so much in disagreement with this that I once invented a *counter*-sound-bite – one that *sounds* the same, but whose meaning is the total *opposite* of the Italian one. My counter-sound-bite runs like this: „Translator, trader".

The meaning of this quip is: *A translator is a skilled artist who makes careful tradeoffs all over the place, losing here while gaining there.* I used „Translator, Trader" as the title of a short book that I wrote in 2009.

„Translator, trader" expresses the opposite of „Translator, traitor" (which, by the way, is an essentially *ideal* translation of the cynical sound-bite *Traduttore traditore*, and thus it undermines its own claim). The funny thing is that whereas the phrase „Translator, traitor" is profoundly *false*, the nearly homonymous phrase „Translator, trader" is profoundly *true.*

Well, it's about time for me to draw this lecture to a close. I've shared with you the fruit of my long labors in converting *Nähe des Geliebten* into English, and I am quite content with it, although a nitpicker would surely try to pick nits with it. That's okay – *de gustibus non est disputandum.* The main thing I wish to say in conclusion is this: When an MT system translates poetry in as thoughtful a way as this, I'll know that the time has come for me to tip my hat and quietly bow out. The jig'll be up; my goose'll be cooked. But that won't happen for a while… or so I hope.