The Unbearable Elusiveness of Natural Language Translation

Global Affairs  JULY 22, 2015  08:00 GMT  Print  Text Size

By Jay Ogilvy

Every time you turn around and see a forecast of the technological wonders awaiting us, one of the favorites is the imminent availability of natural language translation. Recent advances like Siri, Google Translate and IBM's Websphere are pretty amazing. But if you think natural language technology is up to the task of high-level diplomacy
or reliable intelligence, don't hold your breath.

A few years ago I received an e-mail warning of the perils of global marketing using literal "translations." I don't know the ultimate source, the unsung genius who first collected these blunders, but I'll pass a few of them along anyway:

- In Taiwan, the translation of the Pepsi slogan, "Come alive with the Pepsi Generation" came out as "Pepsi will bring your ancestors back from the dead."

- The name "Coca-Cola" was first rendered in China as Ke-kou-ke-la. Unfortunately, the company did not discover until thousands of signs had been printed that the phrase means, "bite the wax tadpole," or, "female horse stuffed with wax," depending on the dialect.

- Also in China, the KFC slogan "finger-lickin' good" came out as, "eat your fingers off."

- Scandinavian vacuum manufacturer Electrolux used the following in an American ad campaign: "Nothing sucks like an Electrolux." Not bad as a literal translation ... but, as I'm here to argue, literalism has its limits.

- The American slogan for Salem cigarettes, "Salem-Feeling Free," got translated in Japanese into, "When smoking Salem, you feel so refreshed that your mind seems to be free and empty."

- Chicken-man Frank Perdue's slogan, "It takes a tough man to make a tender chicken," got terribly mangled in a Spanish translation. A photo of Perdue with one of his birds appeared on billboards all over Mexico with a caption that explained, "It takes a hard man to make a chicken aroused."
You laugh. You protest that the newer programs, or at least the programs soon to come, would not make such blunders that date, after all, from more than a decade ago. But do you know why your nervous laughter spells inevitable peril for natural language translation programs? Let me tell you why in four easy lessons.

1. **Natural Language Is Not Code**

People who get worked up about the prospects for natural language translation are victims of a very simple, very basic confusion. They think that translating one natural language into another — English into Japanese, or Portuguese into French — is like translating Morse code into English. Wrong!

Codes, as opposed to natural languages, permit very simple one-to-one equivalence. Dot-dash in Morse code is, always and everywhere, the one-to-one equivalent of the letter A. That's how codes work. Crack the correspondence rules, and you can find the simple synonymies.

Not so with natural languages. Colloquial English, French, German, etc., are highly complex products of eons of evolution. They don't work like codes. Rather than relying on simple synonymy, natural languages in their everyday use thrive on ambiguity, multiple meanings, plays on words and context as an aid to figuring out a speaker's intentions. "Bank" can refer in one context to the side of a river, in another to Chase Manhattan, or to the tilt of a curve in a road, a shot in billiards, and so on.

Once you crack the code of an encryption program, however complex it may be, you can establish one-to-one equivalencies for each symbol in one sign system to its corresponding symbol in another sign system. This is not the case in translating natural languages, as the opening salvo of literal translation disasters
What is Global Affairs?

2. Metaphors Are Essential to Natural Languages

Linguistics professor George Lakoff and philosopher Mark Johnson have built their well-deserved reputations around the insight that metaphor is not just second best to literalism: Metaphor is the very meat of natural language.

In a series of books that began with *Metaphors We Live By* and continued with *Women, Fire, and Dangerous Things* and *Philosophy in the Flesh*, these two pioneers have crossed the boundary between academic specialization and useful common sense. They have shown, I think definitively, that natural language is not just some degraded form of an ideal, crystal-clear code for representing the outside world into non-fiction documentary movies in the mind. Instead, virtually every corner of natural language is built around metaphors drawn from the verticality of the body, the nature of space, and the everyday geography of the human condition.

It's no accident that up is good, down is bad. It has something to do with our learning to walk vertically. The fact that we inhabit our human bodies leads us to say things that cannot be literally true: "I grasped his argument." "Hold that thought!" Despite the fact that we all know that minds don't have fingers or hands, we accept these turns of phrase as useful, easily understandable moves in natural language. Not until a linguist and a philosopher come along and point out such things do we become sufficiently aware of how thoroughly our everyday language is successfully built upon turns of phrase that have become so commonplace that we forget they are
not literal representations but metaphors.

When a natural language translation program attempts literal translations from one natural language to another, the lack of literalism *comes home to roost* (there goes another one!). The lack of literalism *rears its head* (and another). The lack of literalism ... becomes manifest (how dull!).

3. Literalism Is Just One Among Many Language Games

Ludwig Wittgenstein is probably the foremost philosopher of language from the 20th century. He opened our eyes to the ways that natural language works. His *Philosophical Investigations* is one of the great books of the last century, puzzled over by legions of perplexed students. Part of its greatness lies in the fact that it represents a 180-degree turn, a complete about-face, from Wittgenstein's first great work, *Tractatus Logico-Philosophicus*.

In the earlier work of his arrogant youth, Wittgenstein laid out the requirements for what he considered an "ideal language." Working in the tradition of Leibniz's logical calculus, Wittgenstein aspired to develop a language so "perspicuous" that one could tell from the mere form of a proposition (independent of its content) whether that proposition was true or false.

Wittgenstein's *Tractatus* put forth a "picture theory of truth." Young Ludwig thought it possible to purge natural language of every vestige of metaphorical ambiguity, every distortion that might bend a proposition away from a literal picturing of the facts.

After years of valiant attempts to reduce colorful counterfactuals like, "If wishes were horses, then beggars could ride," down to simple, declarative propositions, he and his followers admitted utter defeat. Moreover, Wittgenstein showed why the attempt to reduce all language to simple, literal declarative propositions was a
preposterous attempt in the first place.

In his later *Philosophical Investigations*, Wittgenstein blew the whistle on the entire ideal language program. He argued that there are *many different language games*, not just literal representation using declarative sentences. In addition to simple description, there's exhortation, joking, inquiring, inviting, or what philosophers J.L. Austin and John Searle later called a range of different "speech acts."

Example: *Watch out!* If you take those words as a literal description, they might mean something like, "Time piece out of pocket." But any five-year-old knows that those words are more likely to be an exhortation, an imperative to beware of some danger.

This thumbnail account of the later Wittgenstein only begins to tap the profundity of his breakthrough regarding natural language and its differences from literal description. If you still think we can iron the kinks out of our natural language translation programs and get them to function "perfectly," then you could do worse than study Wittgenstein's devastating critique of perfect correspondence and exact picturing.

The main point I am making is that natural language is precisely *not* a complex encryption of reality. As the later Wittgenstein showed us, natural language is a congeries of very different language games used in different ways, for different purposes that go far beyond literal description. As Lakoff and Johnson show us, the distorting lenses of metaphor are inexpungible from our use of natural language. If we think we can eliminate all of the "distortions" of metaphor (actually they are *enhancements*), then, like the early Wittgenstein, we're on a fool's errand.

4. **Beware the Holy Grail**

Now for my fourth and last argument against pursuing the holy grail
of natural language translation. Let me state it as concisely as I can, then unpack it.

The closer we get to workable, usable natural language translation programs — and they are getting better, month by month, year by year — the higher our confidence in them will become; the higher our confidence, the greater our dependence; and the greater our dependence, the bigger the screw-ups when they fail. And for the first three reasons I've given, fail they will, sooner or later.

Twenty years ago, we placed relatively little faith in the ability of our translation programs, whether from Japanese to English, or from handwriting to the English alphabet via Apple's Newton. Some may recall the joke, "How many Newtons does it take to change a light bulb?" The punch line, "three elephants and a rhinoceros!" pointed out that Newton failed to register, that is, translate, the question, much less come up with the right answer.

These days, the programming of hand-held touch screens has significantly improved. So have natural language translation programs, not only from spoken English to printed English via voice input systems, but also from many languages to English and vice versa via natural language translation programs. Translation devices are available on mobile apps, and once you've adjusted your settings appropriately, Google will automatically translate any foreign language site you pull up. The progress is real, no doubt about it. And it will continue as enough epicycles are recursively added to correct past mistakes.

But my argument is that the progress is intrinsically and necessarily asymptotic to perfection — a great word from mathematics meaning something like, "approaching ever more closely, but never quite getting there." It sounds almost like the unbearable elusiveness of orgasm in Tantra.
Natural language translation will never be perfect, any more than a map can be a perfect representation of a territory. But the closer we get to perfection, the more we will expect from our natural translation programs. And the more we expect, the greater our disappointments will be when they do eventually fail, as they must.

I fully expect natural language translation programs to be available for ordering meals in French restaurants where those arrogant French refuse to speak English, even if they know how.

I fully expect natural language translation programs to be adequate for many simple tasks, from asking for directions to the nearest hospital to translating anatomical descriptions of animals.

But if we ever get so confident in our natural language translation programs that we make the mistake of depending on them to conduct high-level diplomacy, or conduct a courtship, or translate poetry, or conduct business, or provide intelligence, then we are in for some big disappointments.

We know that various three-letter agencies are using Google Translate. We fear that unfortunate errors could occur. On a less cosmic scale, a friend writes just the other day: "And lo — the mistakes do slip past us because our dependency and 'techno-trust' set in so rapidly. How many times have we hit the 'send' button, only to realize nanoseconds too late that we just told Mom to do something obscene?"