

JUGGLING WITH WORDS: SOME INSIGHTS FROM THE HUMAN TRANSLATION PROCESS

by Claude Bédard

TRADUCTIX inc.

Abstract

An attempt is made to modelize the process of human translation, in order to illustrate the complexity of the process, and to better define successive quality levels: syntactic, idiomatic, and stylistic.

1. Introduction

Translation is... what translators do! This punch line is from none other than the prominent computational linguist Martin Kay, who goes on to explain that translation is essentially a communicational, rather than linguistic, process. Although translation involves a major linguistic component, communication is the driving force. If I may propose a culinary metaphor, linguistics is definitely the main ingredient, but communication is the recipe to follow... the translator being the cook.

Shouldn't MT research, then, be curious about how the cook operates? Some MT researchers would say that translation theory is not useful for MT, the proof being that human translators themselves can be successful without having to learn it. This statement is easy to turn around, however, because indeed MT has failed to make the same demonstration, and this in fact reinforces the point that MT would benefit from some form of translation theory. But it is also true that existing translation theory, i.e. for human consumption, is by no means adequate — just the same way that a regular Webster dictionary, be it CD-ROM, is of no use for an MT system. What we need is an MT-oriented translation theory.

In the scope of this short paper I will attempt describe the workings of the human translation process, with a dual aim in mind:

- putting more flesh on the concept of translation as a communication process;
- proposing a few reference concepts on which translators and researchers could base future discussions.

2. The human translation process

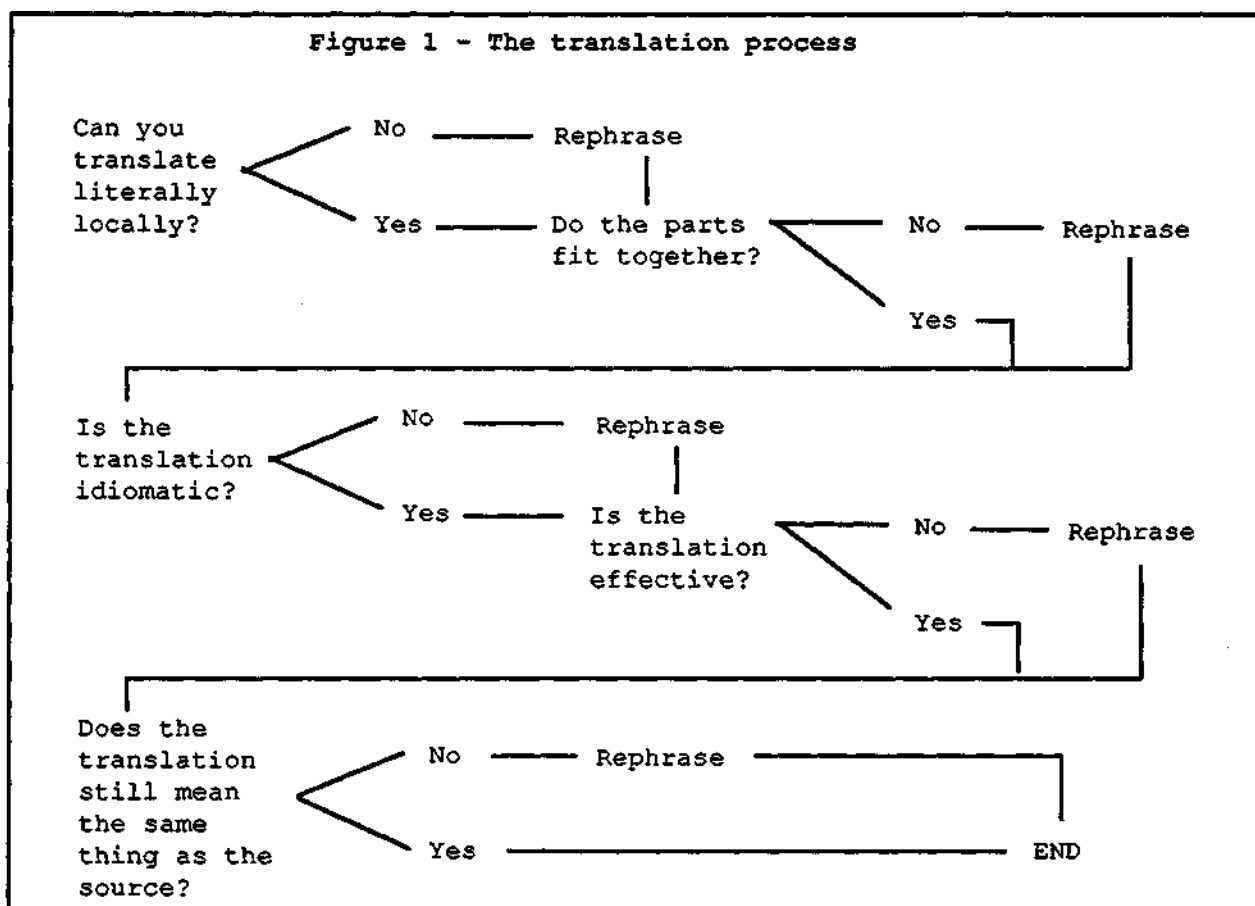
From a layman's viewpoint, it is difficult to understand what is so tricky about translation. After all, can't it be summarized by two basic commands:

- replace all SL terms by the appropriate TL equivalents;
- translate the surrounding verbiage so that there will be no grammar or syntax mistake.

Of course these two commands have to be carried out in a meaning-preserving way, but this doesn't sound like a problem. Well, things are not so simple, as we are going to see.

2.1 Overall description

I will now describe the basic process of the professional translator. Of course my linear description is a gross simplification, as the translator's mind does not have to work linearly. Figure 1 summarizes the translation steps, which will be explained hereafter.



2.2 Detailed discussion

Let's now examine each step more closely.

2.2.1 Translating locally

Translating at the local (phrase) level involves performing lexical transfer and applying syntactic rules on the result.

2.2.1.1 Lexical transfer

MT researchers are generally aware of the magnitude of the lexical transfer problem — choosing between the many possible TL equivalents for each SL word or term. What is less obvious is that this problem carries right into technical terminology. *The* poor limited set of words we use in technical domains has become so overworked that it is no longer an effective tool for communication.

One illusion we should give up is that connecting a huge term bank into an MT system would take care of the terminology issue. This would not be a satisfactory solution, for several reasons:

Even technical vocabulary is polysemous. Polysemy is not restricted to general language, and cannot be managed simply by tagging terms by domain.

- Polysemy exists even within a restricted field. For example, in electrical engineering the word *power* has to be translated in French by *puissance*, *énergie*, *électricité*, *alimentation*, or *courant* according to context.
- Within a given document, domain priority may well shift from sentence to sentence.

Technical writers often make up their own terms. In many cases the cause is not the need to create labels for new concepts, but by a lack of awareness of existing terms. In other cases the text introduces specific concepts which reflect a different viewpoint on existing technology, hence the need for text-specific terms.

Many writers assign different meanings to known terms. The translator will often see familiar terms used in unusual ways. This is due to the same causes described above.

Clashes or gaps between terms may occur in the translation. Since the mapping between languages does not occur along parallel lines, the translator must compensate for clashes or gaps which come up naturally in the TL. For example, one specific document will use the terms *apparatus* and *equipment* to designate different items. Both terms can normally be translated in French by *appareillage*, but in this specific document the translator has to make sure that they are translated by different terms.

There is no real coherence within technical terminology as such¹, but within each text the translator can and indeed should create coherence. He cannot simply passively assign TL equivalents to SL terms on a one-by-one basis, but rather deal with the text's terminology as a system.

2.2.1.2 Syntactic transfer

On the syntactic side, typically the translator will first go for the least effort — he will consider a straight (i.e. literal and syntactically correct) translation. Success depends much on whether lexical transfer is direct, i.e. a target word of the same syntactic category is available for each source word. If this condition is not met, as in the following example, rephrasing is called for.

Successful participants will receive a prize.

There is no convenient, all-around French equivalent for *successful* applied to a person. The translator will typically be forced to paraphrase, such as *qui auront réussi* or *qui ont répondu correctement*. Notice that paraphrasing often requires extra knowledge — in the first case the future tense (the event has yet to take place), in the second case the activity (answering a question). This is one example of the need to understand in order to fix a very low-level translation problem.

2.2.2 Fitting the parts together

If all local translations are syntactically compatible with each other, fine. Otherwise, the translator has to fit them together, which involves changing some perfectly acceptable local solutions for this purpose. The translator is then confronted with the open-ended problem of choosing which parts to unbuild, and how much to unbuild them, before refitting the whole thing.

Creating a syntactically correct translation is often similar to puzzle-work. Syntax has structural constraints as the component parts must be "in balance", somewhat like in an algebraic equation. In MT terms, the translator must switch from a "direct" approach, to a high-end "transfer" approach.

Here is an example of fitting together local solutions:

ABCD Co. is a leading supplier serving the commercial, industrial, contractor, and institutional markets.

Here we have four parallel components:

commercial market	=> marché commercial
industrial market	=> marché industriel
contractor market	=> marché des entrepreneurs
institutional market	=> marché des institutions

Fitting parallel components together is perhaps the most illustrative case of syntactic juggling. In the example, putting those parts together as they are would produce *les marchés commercial, industriel, des entrepreneurs et des institutions*, which is hardly acceptable because of paralleling adjectives and noun complements. We have to find something better.

One possible way to re-establish parallelism would be to repeat *marché* in every instance (*le marché commercial, le marché...*), but this would sound heavy and clumsy.

The solution finally adopted was to skip completely the word *market* and nominalizing the adjectives to create fresh, independent noun groups, which results in *les clients commerciaux, les industries, les entrepreneurs et les institutions*. Note that nominalizing *commercial* involves introducing the word *clients* which is absent in the source text. (This is an example of a "padding" word, as discussed later.)

If translation was strictly a matter of linguistic transcription, translators would often be quite embarrassed. But since translation is primarily about communicating meaning, translators can use this to back off from various syntactic dead-ends. Here are some of the processes used by translators for rewording a translation considered inadequate — either syntactically, idiomatically, or stylistically.

2.2.2.1 Reassigning items

In this example, the translator "unplugs" the component *powerful action* from *penetrates* and replugs it into *remove* (transforming the latter into a tensed verb).

SL sentence	Translation	Back-translation
Powerful steam action on every cycle penetrates every crevice to remove the toughest residues.	À chaque cycle, la vapeur pénètre dans tous les recoins et son action énergique déloge les résidus les plus tenaces.	On every cycle, the steam penetrates every crevice and its powerful action removes the toughest residues.

Nothing is lost, and the target sentence sounds much better than a direct translation.

2.2.2.2 Omitting or making implicit

Every sentence includes words which, although syntactically essential, don't contribute anything to the meaning. Language has built-in redundancy, and some individual styles are more redundant

than others. Words are often used as syntactic "padding" to make the sentence balanced and complete.

Thus in a sentence the translator may find three strata:

- essential words, earning the focus or intent;
- meaningful words (content words);
- empty or "padding" words (in terms of text grammar, these words don't bring any significant or new information).

In the event of a high bearing temperature *condition* of the *blower*, the *operator* should *immediately* investigate to *establish* the cause of the *problem*. Should a *blower* shutdown be *indicated*, the *operator* should *first* start up the spare *blower* then shut down the problem *blower* via the local *panel*.

In the above example, the focus words are underlined (none of the happens to be technical terms) and the empty words are in italics. This is demonstrated by a telegraphic version of the same sentence, which remains understandable.

If high bearing temperature, *immediately* find cause. If shutdown necessary, *first* start spare, *then* stop other locally.

For the translator, detecting empty words and using the resulting margin of freedom is akin to using cracks and crevices in solid rock for a mountain climber. Sometimes a whole idea can be skipped because it is already present in the context.

2.2.2.3 Breaking up syntactic ties

The translator decides to undo some awkward syntactic links, while keeping the semantic relationship obvious to the reader. This is achieved mostly by punctuation devices, as in this example:

SL sentence	Translation	Back-translation
The maintenance depends largely on whether or not the unit is run continuously or intermittently.	L'entretien dépend largement de la nature du service (continu ou intermittent).	The maintenance depend largely on the type of duty (continuous or intermittent).

2.2.2.4 Using a different keyword or viewpoint

The translator may also, as already indicated, try to bypass completely the syntactic puzzlework and come up with a fresh rendition. One useful technique consists in bringing to mind spoken language utterances, which often are based on a different starting point. In MT terms, the translator would switch from a "transfer" approach to a "meaning-based" approach.

SL sentence	Translation	Back-translation
Losing the tool half-way into the duct means starting all over again.	Si l'outil venait à se détacher au beau milieu du conduit, il faudrait tout recommencer.	Should the tool come off right in the middle of the conduit, one would have to start all over again.

This example illustrates one basic truth about translation: SL and TL may not offer the same wording facilities. It often happens that a symmetric facility is lacking in the TL. The English example is a nominalized sentence based on gerund forms, which is not possible in French, where a nominalized French translation (*La perte de... signifierait le recommencement...*) would sound heavy and unnatural.

2.2.3 Making the translation idiomatic

Different languages not only have different dictionaries and grammars; within the possibilities provided by their dictionaries and grammars, *they also choose not to say things the same way*. Indeed, each language superimposes on its grammar a set of stylistic preferences (well know, in French, as "le génie de la langue") which are so stringent that they probably become a criterion to determine whether a sentence, in Chomskyan terms, "belongs to the language".

Let's stay out of there.	=> Restons hors de là.
Ne mettons pas les pieds là.	=> Let's not put our feet there.

In this example, the two sentences on the left are idiomatic translations of each other, whereas on the right we have literal (although here "les" has been softened to "our" rather than the literal "the") translations, which sound quite odd. To compound the problem, the literal English translation would be idiomatic in a situation where it refers to specific locations on a floor, such as traps or fresh paint.

Let's consider this other example:

Si l'on y regarde de près, ce n'est que dans les textes de nature juridique qu'on bannit le non-dit.	=> If one looks at it closely, it is only in the texts of legal nature that one banishes the unsaid.
--	--

The English translation is understandable, and would be considered acceptable for purely informative purposes, but it would be unacceptable for publication.

As illustrated by the above example, idiomaticity is by no means restricted to the rather finite reservoir of "canned" idiomatic expressions ("where all is said and done", "it goes without saying", "it's not my cup of tea", etc.). It covers the omnipresent reality of casual situations where a native speaker "would (or would not) say it that way". The acute consciousness of this fact about the languages being handled (keeping each different "style" to the language it belongs to) is one of the main differences between a translator and a bilingual person.

Idiomaticity assessments vary from individual to individual (because of the reading material each person has been exposed to). They also follow a continuum, as there is no clear-cut point where idiomaticity begins or ends:

- When a sentence is grossly unidiomatic, one is tempted to say that the sentence is not grammatical, or does not belong to the language.
- At the far end of the idiomaticity continuum, a writer or translator may go further than a wording that does not hurt the habits of the native speaker— he may strive for stylistic quality, which is discussed next.

2.2.4 Making the translation effective

Now that we have an idiomatic translation, can't we stop here? It all depends on the circumstances. The standards for professional translators require that the translation should sound not as a "translation", but as an original text by a competent writer. Thus for publication-quality translation, the translator has to further ensure that the target text:

- is *stylistically appealing*, if the purpose is to favorably impress the reader (to bring him to buy the product — or simply to read on);
- is *readily understandable* and gives the reader the information he needs, in the case of instructional material.

2.2.4.1 Stylistic quality

The first sentence represents a level of style we normally expect from an official publication. The second sentence is a "dry" transcription.

Elegant sentence	"No-frills" sentence
The communication of information by means of binary signals has gained such an important place in American society that it is fair to say that successful operation of the U.S. economy would be impossible without it.	Binary data communication has become very important in the U.S., and now the U.S. economy absolutely needs it.

The need for stylistic elegance is quite variable. It does not add much clarity to the message (in fact it makes for longer wording), but it enhances the appeal of reading and projects a better image of the originator (either an individual or an organization). If this is important in the situation, then elegance is vital.

2.2.4.2 Making the TL sentence understood

Here is another reason why the translator cannot simply sit on a so-called "accurate" translation. I believe the reader is entitled to a text he can understand at first reading, whereas a merely "accurate" translation may require two or three readings to deliver its message; the words do contain the message, but they obscure it rather than reveal it. This is another purpose of style, apart from elegance: making the sentence easy to understand.

The translation should make the thread of thought of the text as obvious as possible for the reader to follow. This requires a fresh effort for the translator, because text grammar devices cannot be expected to be parallel in SL and TL texts, especially with the lexical and syntactic shuffling involved in translating.

What is the sentence trying to say? Sometimes the real intent of the sentence is hidden, and translating "what meets the eye" will obscure it even more in the target sentence.

Rather than simply accounting for the *content*, the translation must also account for the *intent* of the source sentence.

In virtually all power lines, skywires are present for lightning protection.

This sentence may seem absolutely straightforward, but while attempting a translation one discovers that at least three intents are possible, depending of the focus word (which in spoken language would probably be stressed):

- The protection devices are *skywires* (rather than some other device...);
- Skywires are *present* (rather than absent);
- The purpose of skywires (which are necessarily present) is to *protect*...

Here the correct focus is determined by analyzing the context to see what new material is added by this sentence (rheme) about the already known material (theme). The translator should follow the thread of the message through successive sentences and paragraphs, and use this "text grammar" knowledge to focus his translation.

Does the sentence have one meaning only? One of the hallmarks of human language is ambiguity. Often the sentence can be interpreted in two (or more) ways, and the translator has to develop a sensitivity for this. Very often the first interpretation that comes to mind is the right one, but every once in a while this is not the case. In fact, one of the main qualities of a reliable translator is the ability to consider all possible meanings, even if the interpretation seems obvious. Failure to do that creates a fertile ground for misinterpretations.

Put the motor shaft down in a vise.

This sentence has a tricky double meaning: "put down the motor shaft in a vise", or "put the motor in a shaft-down position in a vise". In this particular case the second meaning was the right one.

2.2.5 Checking the meaning

Here the question is: does the translation say the same thing as the source sentence? This sounds like an easy question, and this is precisely the danger about it. In fact, as a translation teacher I have been struck by the fact that many mistranslations were due not to misunderstanding the source sentence, but to failure to recognize that the translation led the reader to understand something else than the source message.

Such failures at meaning preservation may be due to a "drift" brought about by lexical and syntactic shifting in the translation process. Because of the transposition into a different system of lexical and syntactic relationships, some new interpretations may be created and existing ones closed.

2.2.5.1 Syntactic shifting

In the following example, the normal translation [1] creates a double meaning, because *voisines* (= *neighboring*) is located so that it will probably be read as *voisines de la terre* (= located near the earth). To back out from this situation, the translator has to create an alternate wording [2].

The neighboring surfaces of the earth...

[1] Les surfaces voisines de la terre...

[2] Les surfaces de la terre situées à proximité...

In the next example, the translation of the preposition *from* into a more ambiguous *de* (which can mean either *from* or *of*) and the swap involving *exploitant* may well induce an erroneous reading *l'exploitant du câblage*. Only a careful rewording [2] will prevent this problem.

It separates the operator's facilities from local wiring.

[1] Il sépare les installations de l'exploitant du câblage local. (Back translation: the facilities of the operator of/from the local wiring)

[2] Il assure la séparation entre les installations de l'exploitant et le câblage local.

2.2.5.2 Lexical shifting

The next example is more subtle, but striking.

All colours may be changed.

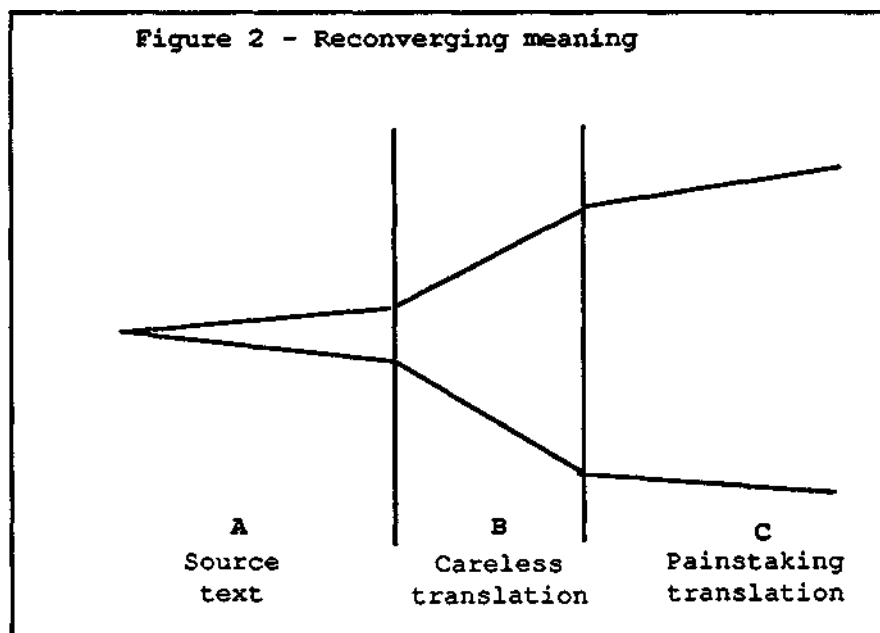
[1] Toutes les couleurs sont sujettes à changement.

[2] Toutes les couleurs sont modifiables.

Translating *may be changed* as in [1] switches the meaning around. In context, the intent is that "the user is free to choose other colours", whereas the word *sujettes* in the translation suggests "the management can choose other colours, therefore the user has no control over colour choice".

2.2.5.3 The need to reconverge meaning

The previous examples point to a basic truth: we simply cannot trust words to carry a message (either within one language or between two languages). Words are like birds, they fly in all directions. This explains why translation cannot be a passive process, as it requires a reconverging effort to retain the value of the text as a communication medium. Figure 2 illustrates the process. In the source text, meaning may be already divergent (i.e. liable to misinterpretation) to some degree (A); a careless translation will often aggravate this divergence or create new divergence, sometimes to the point of intelligibility (B); an active (sometimes painstaking) effort is required by the translator to "reconverge" the translation, ensuring that its message-carrying ability remains comparable to that of the source text, if not better sometimes.



3. Conclusion

This analysis tends to implicitly describe the qualitative gap between MT and professional-quality translation, by defining successive levels of quality : syntactic, idiomatic, and stylistic. Straightforward lexical and syntactic transfer, along with some limited idiomatic reworking in the better ones, is what we get from current commercial systems, although it may well happen (especially in rather simple sentences) that the syntactic level coincides with the upper levels.

If we view MT as an aid for the translator, we have to realize that this is the easy part of his work. Anything more difficult is left for him to do. This problem-solving activity is time-consuming (as the saying goes, the last 10 % in quality requires 50 % more effort), and puts a ceiling to the productivity gains translators should expect from MT.

If this gap between MT quality and professional quality is too large, MT loses its appeal because translators feel that "it's just as fast to translate from scratch". To narrow this gap, the choice is straightforward: either we drive up MT output quality, or we drive down the quality standards.

Maybe we could consider the second alternative, at least for the purpose of discussion, asking the question: Do translators put too much quality (making the translation sound like an original source text, and very well written) into the market, unnecessarily driving up both prices and turnaround times? It would be an interesting exercise for MT developers to challenge translators to justify their quality standards. This could be a way to create dialogue between the two communities, which could lead to a better mutual understanding. The classification proposed here could be used as a starting point.

¹ For a detailed discussion on terminology and its use in translation, see Bédard, Claude, *La traduction technique : principes et pratique*, Linguatex, Montréal, 1986, ISBN 2-920342-17-7.