

Working Environment
changes induced by
translation memories

Evelyne Ross
University of technology of Compiègne

Working Environment changes induced by translation memories

The purpose of this proposed paper will be to show how the introduction of translation memories into the workflow-process of the translator can significantly change the working environment. This environment change will include the areas of collaborative work, organization of the work, and finally the cognitive process of translating. Our results are based upon empirical studies made by twenty-six translators in France, Switzerland and Germany. One group, in the public sector, includes international organizations, government, and the European Union, while a second sector included private entities like insurance companies and banks. Our empirical studies were computer-generated examples, while the individual case studies result from traditional paper-based evaluations generated from sector participant response. We will compare the proposed intellectual strategies with the cognitive processes that intervene with, or impact upon, the process –with and without the use of translation memories. The work and study data were completed between early 1998 and mid-2000. This analysis provides for the design of a new sense of translation methodologies, which can significantly impact effective communications across our global society.

Introduction

Translation memories are more and more integrated in the working environment of the translator. They should increase productivity and ensure greater terminological accuracy, and consistency in text types displaying a relatively high degree of repetition.

The introduction of electronic tools changes the translating process. In this sense, we can refer to the Ph.D. thesis "Was in den Koepfen von den Ueberseter vorgeht?" (Krings96) and ask the question "What is going on in the head of the translator when he uses a translation memory?" We should answer the following questions: How does the translator understand the source text? What kinds of strategies are used for the validation of a good translation? Which translation is a good one? Which criteria are in play?

Research on the task of translating by a translator and research on the task of translating with an information processing system begin with the same questions.

I will first define what translating is from contemporary cognitive psychology's point of view. Secondly, I will report on some remarks made by translators about the kind of research that has developed in and around the task of translating. Thirdly, I will present our methodology for task analysis and finally, I will introduce a cognitive model that could be used to study the task.

Empirical study

In software development, the definition of the user characterization is a key factor for the usability of the system. There must be an understanding of who the users are, together with an appreciation of their relevant characteristics. The analysis was guided by the following questions: Who are the translators? What are their educational backgrounds? How many years of practice do they have in translation with a translation memory? What are the main steps in the translation process? What are the translator's goals and activities? What problems do they have in the production of their translation with the translation memory? How do they solve these problems?

We interviewed and observed 26 translators working in France, Switzerland and Germany. One group, in the public sector, includes international organizations, government, and the European Union; while a second sector included private entities like insurance companies and banks.

The empirical route advocates a collection of information about the translator and his reaction of the system. For the knowledge acquisition activity, two kinds of empirical methods were used viz. interviewing techniques and where it was possible protocol analysis, where the interviewing techniques consisted of orientating interviews and structured interviews.

Translating as categorization

Comprehension is the foundation on which any translation will be built. Oftentimes, it is also its most difficult part, while reformulation poses no particular problems. In translating with the help of a translation memory, comprehension is made more difficult by a number of constraints, including time pressure and the simultaneousness of several tasks.

What is the source text to be interpreted? Like any text, it has a macrostructure that can correspond either to a minimal summary of the speech, or to some message the speaker wants to deliver. This macrostructure finds its way into sentences. Between this macrostructure and the words that are mentally pronounced, there are many intermediate levels that provide both structure and meaning. The translator can produce a very large number of different texts from the same macrostructure and that structure is relevant to the author's intention. A satisfactory translation would respect all the levels of the text's structure from the choice of words to the message. Secondly, and more crucially, note that a large number of texts can have similar introductory parts and elements without sharing the same message.

In his/her daily work, the translator has the opportunity to read the entire text before starting to translate; he or she has access to the whole structure. Unfortunately, the translation memory displays only parts of texts, sentences that

are to be translated, or sentences that should correspond to the sentences of the source text. In this case, the translator has access to the microstructure of the source text. Most of the time, the translator has to reconstruct the structure word by word. The problem at hand is thus one of grasping the structure through inductive inferencing: a problem of translation and evaluation.

In addition, words and sentences are polysemous. Even in word-to-word translation one must understand which meaning of a word is being used in order to choose the correct word in the target language.

How do polysemy and context affect our understanding? Our theory of understanding provides some explanation. It is a post-Piagetian Theory in the Cognitive Sciences that emphasizes Domains, Objects and Properties to model the Human Information Processing System within Semantic Networks of Knowledge. It is also an ecological theory (Anderson91). The mind has to provide an interpretation of the world in order to simplify it, complete it (most of the meaning is literally absent or implicit) and understand it. This understanding is Task-based, goal-oriented (Marr82). Information is processed through Language (Syntax/Task-procedure) to Categories (Semantic Networks).

At every point along the line that goes from perception to understanding, the mind deals with a fuzzy, ambiguous and uncertain world by making inferences that are knowledge-based (pragmatics included) and goal-oriented. Knowledge and goals make inductive decisions possible by using categorization processes. Categorization is the main mechanism that underlies the choice between alternative interpretations, because it helps differentiate between homonyms and between meanings of polysemous terms (when the context activates or deactivates categories or connotation).

The important role of the context and the goal is to provide cue-properties that prime categories that make sense. This priming effect corresponds to the expectations one has about what is going to happen, the direction the text is taking.

A categorization is defined by properties, and recent studies have shown that three properties suffice to identify one category among six hundred (Mantyla86). This shows that categories can be activated in the memory without being named. Once a term has been assigned to a category, the category will in turn provide cues for further terms to be categorized. For instance, the sentence "Jack took the knife" provides the category of "instrument used to cut" which is a subcategory of "sharp instrument", when preceding cues have provided a "fight" or "dress making" context. The chosen category will in turn provide expectations about the patient object (the object to which something is done). The choice between alternatives is goal-oriented and is provided by the macrostructure of the text: understanding emerges from networks of categories and macrostructure from a limited number of more general categories. For instance, "Check the battery", in a technical document, is processed as words (lexicon) in a specific order (syntax) which correspond to the task:

Invites [A (an individual, x)],
[B: to check [C: motor]]

In order to understand this sentence, A, B, and C have to be assigned to categories that will differ according to context.

What are the underlying mechanisms of categorization?

1. *Translating occurs by categorization in Semantic Networks of categories* with inclusion of categories and inheritance of properties: "Ducks" or "Parrots" are kinds of "birds", "birds" are kind of "animal". "Ducks" or "Parrots" have specific properties, but they inherit "Birds" properties as well as "animal" properties (Collins, Quillian69, Gelman88). The same mechanisms are at work in determining the meaning of verbs.

2. *Categorizing is either generalizing* ("she likes flowers": "she likes plants") or specifying ("she likes flowers": "she likes roses"), depending on the task at hand (Hunt, McDaniel93).

3. *Categorizing is reducing ambiguity/complexity by goal intension* which means understanding objects from the point of view of the selected properties (a flower can be seen from many points of views such as: "needing water", "a present", etc.).

4. *Categorizing is inference making*: the information processed from the text is completed with unseen or unstated properties of cognitive attributes.

5. *Categorizing provides an internal representation of the situation at hand* by creating a temporary semantic network of categories that we construct and deconstruct until we get coherence. This can be seen as the understanding of the macrostructure of the text.

Research on Translating: Some remarks

In this part, we will express three remarks that will differentiate the presented approach from others.

1. In the research today, the translator is the object of the research. Translating is a problem solving task (Wills88, Hoenig95). The model stipulates that the goal exists which can be reached within an organization of the translation process. The model is characterized by:

- The current topic
- The audience
- The author's motivation and
- The translation under construction.

The translator is represented by his/her competence residing in memory (knowledge about the topic, the audience, and translating plans), and by the translating process he performs. A monitor steers the translating process, using different strategies. The organization of the translating process is divided into three phases: planning, translating and reviewing. Since a first draft is seldom acceptable, it is systematically improved during the revision phase. All main phases contain subunits that remain, however, rather global in their definition. The phases

do not follow each other in a linear sequence, but a monitor steers their succession in a reasonable fashion.

The translating process is a goal-directed activity in which the translator develops and generates his own goals. There is a hierarchical organization of processes and goals; attention to these processes is delegated to the monitor. It steers the translating process with respect to individual strategies and to the way translators handle specific functions.

2. The relevant domain of analysis seems to be ergonomics, or cognitive ergonomics, including task analysis, which means that the translator has a task, a goal, a starting and an end point, a context, as well as operators, procedures and strategies with which to perform the task, and that he or she has some experience, expertise, and some general knowledge to bring to the task, for instance general guidelines or rules of thumb.

Task analysis

In many professional domains, expertise is not explicitly transmitted to students but acquired by practice. This kind of "learning by doing" is implicit learning; we know how to do it. Task analysis forces us to make it explicit. Part of the modelling activity is to make explicit what usually remains implicit.

Since the development of the Newell and Simon (72) approach to problem solving, cognitive psychologists have provided methods for describing tasks in many domains. The problem-solving paradigm has been successfully applied.

The first step in analysing the process of translating with the translation memory is to make an ergonomic description of the task, which entails, as for any task, identifying the translator's goal, available procedures and strategies (the most important of these strategies being the metacognitive one based on the knowledge of how to deal with specific difficulties) as well as identifying the limits of a translator's abilities. What we mean here is the ability to evaluate alternative procedures, to elaborate better strategies to improve performance, and to diagnose dead-ends as in problem solving, and all in short periods of time.

The second step is to model the internal construction of the macrostructure of a text, something that must be done, as early as possible, starting from microanalysis of sentences. Our hypothesized model is based on the construction and deconstruction of the overall intention of the translator. This is modelled within the constraints of the task, so that the translator is able to provide himself with a coherent representation of the content of the source text.

The first step is to analyse the advantages and the limits of the translation memory.

Lastly, we have to address the attentional aspect of translating. These are attentional costs both at the task level and at the representation level of the microstructure of the text, given the attentional limits of the human information processing system.

Let us see what has to be done to analyse the task of translating with the translation memory. We define any task as two states of a micro world (the initial state and the final state) plus a main goal, namely moving from the initial state into the final state, by means of "actions/operations" that become a procedure once the "operator" knows how to perform the task. There are many differences between tasks:

1. "static" versus "dynamic" micro world

A static micro world is one that does not change unless you perform some actions (e.g. solving a problem on paper); a dynamic micro world is one that changes even if you are not acting (translating a speech).

2. "well defined" goals versus "open-ended" goals

The goal can be either very precise, as when the end-state micro world can be described with exactitude and precision ("checkmate" in chess), or it can be openended, a goal for which the end-state can only be defined globally, within the limits of a certain number of constraints that define what the micro world must and must not be: e.g. the fact that the reader must be able to understand the text in another language. In the case of dynamic micro world and openended goals, we define two kinds of constraints: first-order constraints and second-order constraints.

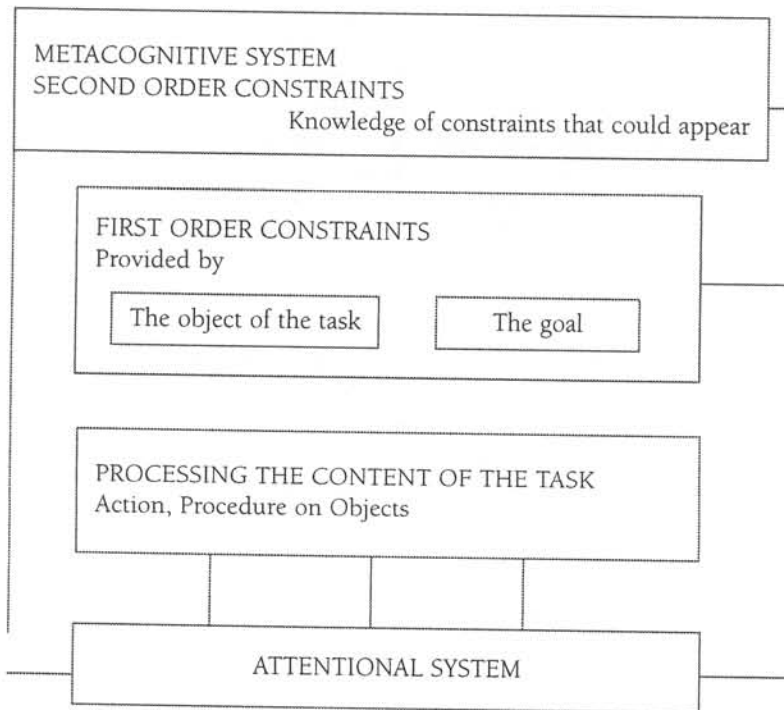


Fig1. Metacognitif system

First-order constraints are constraints that guide actions, or procedures, by providing variables with values that have to be respected (for instance, about the dynamic evolution of the text source; *variable*: part of the text to translate; *value*: give the sense of it). Second constraints are metacognitive constraints that define the values of the first-order constraints, when all the constraints are not known before starting the task.

Metacognitive constraints supervise the task by diagnosing the first-order constraints (e.g. to evaluate the resources that will be necessary; one variable could be: some problems with unusual professional terms; and another: I will have time to do my best) and by providing constraints values to them (value: I might have some problems with technical terms: I will not have time to do my best).

Task analysis requires defining first and second-order constraints, and making an inventory of all known procedures for solving the problem. This corresponds to a theoretical and formal analysis done at a macro-level where it is not necessary to take into account the human operator who will perform the task, because one might be able at this point to describe any kind of behaviour even if one has no idea of what will in fact be observed when a particular individual performs the task.

But how can we know how a given operator will perform a task? To simulate task performance, we must obtain precise knowledge about (1) the operator's representation of the task, which is the way he describes to himself the first and second-order constraints and the available procedures, (2) the content of the task, which is the way the objects of the micro world change through actions or procedures, both from the point of view of the operator and from the point of view of an objective description of the domain, (3) the way the attentional system (input/output) will deal with processing the constant changes that occur in the micro world, given the results of the operator's actions, the dynamic nature of the domain, and the operator's representation of the task.

In summary, this defines three levels of analysis. These three levels are not levels of independent processes, but levels of cognitive processes that interact and must, in the end, be integrated.

Ergonomic analysis of the translating task with the translation memory

Now, let us see how we might analyse the task of translating by listing all the constraints that could influence the process as a whole.

Controlling the task: the metacognitive system

As mentioned, metacognitive constraints are second-order constraints stemming from first order constraints. They provide the kind of knowledge about a task that allows people to get ready to meet its demands.

The first of these metacognitive constraints, besides not getting the two languages mixed-up, is the knowledge that the translation cannot be perfect, that the translator might come to some impasses, that some errors will be made along the way and that remedying these errors will require making further decisions. We are assuming that the kind of reader, customer, domain, type of documents, are variables that will define the demands of the task.

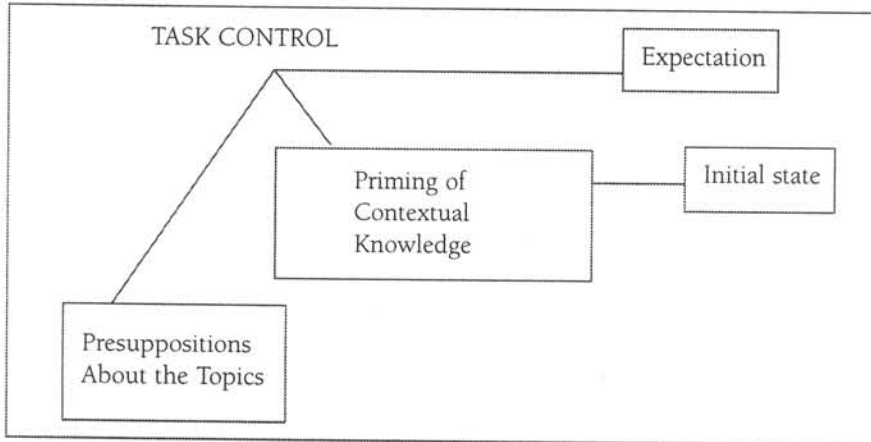


Figure 2. What is to expect, and being ready for

Solving the task: The working system

Constraints associated with the goal and to sub goals

Choosing a goal allows one to proceed on to choosing the strategies and procedures that are best suited to its pursuit. The main goal is to arrive at a state where the text has been translated. Is this sufficient? Here we have to list not only the usual goals of translation experts, but also all the possible constraints that could be used to define a given goal in a given context. When this is done, we have a list of sub goals, as they are the product of the constraints that stem from the main goal. Note that some of the sub goals' constraints will be static, which means that they are like steps to be taken in the course of the translation, and some will be dynamic, which means that they provide parameters throughout the translation.

Static sub goal constraints might include, for example: "Start translating as soon as possible", "find the structure of the text", "find the intention of the author", and so on.

Dynamic constraints might include "make the structure of the text clear".

Now, supposing we have listed (1) the metacognitive constraints, knowing what to expect from different kinds of texts with variable parts to be translated with the machine, and how these define specific goal constraints for specific kinds

of customers and domains, (2) the usual constraints, and (3) that we have separated these constraints into static and dynamic ones, this list can be considered as a model for translators. We hypothesise that intra-individual differences (quality differences between translations of the same level of difficulty made by the same translator), as well as inter-individual differences (quality differences in translations made by different translators with the same background in translating), can be predicted from the order in which the translator(s) has listed and established a hierarchy among constraints. The list will be more or less functional.

Constraints provided by Strategies and Procedures

From here on, the actual task of translation can begin. Procedures are what will be done to achieve the sub goals, and can be seen as co-ordinating sub-sub goals applied to the specific objects of the task: here the content of the text. They are content-based and cannot be analysed independently of the basic objects of the task. If the main object is the sentence, and not the word, such procedure should be hypothesised in Table 1.

For the purposes of ergonomic description, the procedures are presented here as rules to be applied by object-sentences. This is not the proper description. As described later, in the next section, we will see that these rules are part of the semantic construction of the representation of the translation –and also of the speech. We will show that the encoding process determines categories of sentences (sentences to be summarized, sentences to be corrected, etc.).

Strategies are meta-rules that guide and supervise the choice between alternative procedures in the course of translation. For instance, under time pressure, procedures will be adapted accordingly.

- | |
|--|
| <ul style="list-style-type: none">• Detect specific grammatical difficulty• Detect errors as they occur• If there are errors, correct. <p>If you cannot understand:</p> <ul style="list-style-type: none">• guess, aiming at consistency• change the level of text segmentation (parts of sentences, word-to-word) <p>Activate the function “Analysis” of the translation memory:</p> <p>If the tool proposes you several alternatives</p> <ul style="list-style-type: none">• read them all• evaluate them• select one of them• if none correspond, translate by yourself <p>When under time pressure:</p> <ul style="list-style-type: none">• do not correct small errors• and so on |
|--|

Table 1: Part of the list of constraints

Symbolic level: Understanding

We noticed that the translator only has the microstructure of the source text when working with the translation memory. He cannot access the macro-structure of the source text. He will have to translate without the whole context.

For example, the French sentence: *Eating raw food*

“Le manger cru pourrait avoir des vertus thérapeutiques”

“Eating raw food could have therapeutic virtues” (Cori, Marandin 94).

“Le manger cru” can either be interpreted as a nominal group (i.e. “raw food”), introduced by the requisite “Le” which we do not need to translate into English, or as a Verbal Sentence in which “Le” designates the thing (i.e. “it”) that could be eaten raw (i.e. “eating it raw”). As the sentence occurs in isolation and there is no possibility of anaphora with a pre-existing reference (i.e. That is a nice apple. Eating it raw could...), the common syntactic interpretation is to consider “Le manger cru” as a nominal group (i.e. “raw food”).

In this case, the translator can take at hand the original text that he has printed before.

Another example: *The Berlin Wall*

“The situation of Berlin when West Germany and East Germany were two different countries: Berlin, a city in East Germany, was divided in two ; into West and East Berlin. The Berlin Wall created an enclave for westerners in West Berlin”

In order to understand “the Berlin Wall situation”, one has to choose between different kinds of walls (straight walls, enclosing walls, etc.). But at the same time, one must be able to access the super ordinate category which “wall” belongs to. This is the category of all things that separate, isolate (borders, natural boundaries...). This helps to understand that the Berlin Wall (a specific wall) is not just a kind of wall (subcategory of “wall”) but also something that has a core meaning (to separate, isolate..., a super ordinate category). Understanding is capturing the general sense that corresponds to the macrostructure and the specified sense that corresponds to the specific situation at hand.

Conclusion

In the translation field, translating is seen as a problem solving process, text understanding and cognitive communication process.

We have noticed that:

- The translator is not the one who translates the text which is in one language into another. The machine does it for him. He becomes an “evaluator” (Ross2001).
- All translators we observed still have a paper version of the source text. They consult it very often.
- One third of the translators did not trust the proposed translation.

- They spend one third of the time evaluating the proposed translation.
- 60% of them have been using the translation memory for 3 years and for repetitive texts. They think that the tool is useful for them. They save time.
- Strategies for evaluating the proposed translation and their own by using the machine are not tolerated and these strategies have not been taught at the university.

What should we do? Train the translator? Make some changes to the program at the University? Should we develop a system to promote the use of these strategies?

Since the translator has a good memory, the customer must remember it, or as we say in French: "Alors que le traducteur a une bonne mémoire, il faut que le client le garde en souvenir".

References

- ANDERSON, J. R., The Adaptive Nature of Human Categorization. *Psychological Review*, 3, 1991, 409-429.
- COLLINS, A., QUILLIAN, M., Retrieval Time from Semantic Memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 1969, 240-247.
- CORI, M., MARANDIN, J., Preference in Natural Language Syntax. *Vème conférence internationale "Information Processing And Management of Uncertainty In Knowledge-Based Systems*. Paris, 4-8 juillet 1994.
- HOENIG, H. G., *Konstruktives Uebersetzen*. 1995.
- KRINGS, H. P., *Was in den Koepfen den Uebersetzern vorgeht*. Tuebingen, 1986.
- Mantyla, T., Optimizing cue effectiveness: recall of 500 and 600 incidentally learned words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 1996, 303-312.
- MARR, D., *Vision*. New York: W.H. Freeman, 1982.
- ROSS, E., Usages des mémoires de traduction. *Traduire*, 186, 2001, 7-18.
- WILLS, W., *Kognition und Uebersetzen*. Tuebingen, 1988.
- WILLS, W., *Uebersetzungsfertigkeit*. Tuebingen, 1992.