Think-Aloud Protocol Analysis in Translation Studies

Paul Kussmaul and Sonja Tirkkonen-Condit

Article abstract
Think-Aloud Protocol Analysis in Translation Studies: This paper reports on think-aloud protocol (TAP) research in Finland and Germany. It discusses some methodological issues: choice of subjects, TAPs in a language-learning and in a professional context, monologue and dialogue protocols, the use of models provided by psycholinguistics. Two types of processes - successful and less successful ones - are distinguished and specified as to the subjects' comprehension and reverbalisation processes, their focus of attention, decision-making, monitoring, flexibility, creative thinking and attitude toward the task. Some tentative results of the research going on in Germany and Finland are presented. The results are expected to serve as hypotheses for the teaching of translation.
Think-Aloud Protocol Analysis in Translation Studies

Paul Kussmaul [sections 1-5] and Sonja Tirkkonen-Condit [section 6]

1. Aims and purposes of think-aloud protocol research

The analysis of think-aloud protocols (TAPs) in translation studies began in Europe in the late 1980s. It was felt that in order to complement the hitherto predominantly deductive and often also normative models of the translation process, empirical and inductive methods should be developed. The models presented until then usually described what ideally happened or rather – with a pedagogical aim – what should happen, in translating. It was people like Krings, Königs and Lörscher in Germany, Dechert and Sandrock in Britain, Jääskeläinen and Tirkkonen-Condit in Finland, to name but a few, who were no longer happy with this state of affairs. They began to ask what actually happens when people translate.

There has always been a kind of empirical research: translation criticism and error analysis, but this was by nature product- and not process-oriented. When comparing the target text with the source text or when looking at errors, one could at best speculate in retrospect about what had happened in the mind of the translator during the translation process. In cases such as interferences these speculations may have had a high degree of
probability, but other types of diagnoses were hard to arrive at. For instance, from the analysis of errors one may have inferred that a translator's foreign language competence was not good enough, but when discussing matters with him, one may have found that he had problems expressing himself in his mother tongue.

What was needed was a means to find out what goes on in the translator's mind, a means to get a glimpse into the "black box," as it were. In the field of psychology, a method had been developed by Claparède (1932) and Duncker (1935) in connection with introspection (Börsch, 1986, p. 198), which was taken up by Ericsson and Simon (1984) and applied to the translation process by the researchers mentioned above. In these think-aloud experiments, as they are called, subjects are asked to utter everything that goes on in their minds while they solve a task — in our case when they translate a text. These utterances are tape-recorded or videotaped and then further transcribed into think-aloud protocols, which are then analysed from a variety of viewpoints.

In addition to increasing our potential for describing and explaining the processes of translation, and thus our theoretical understanding, these analyses have at least two pedagogical purposes. (1) The strategies observed in the TAPs may serve as models for successful translating (Lörscher, 1992a, p. 146; Jääskeläinen, 1993, p. 101; Krings, 1988, p. 66; Kussmaul, 1993). This implies that the translators serving as subjects possess some degree of professionalism and expert behaviour. Naturally, one would not expect beginner students to exhibit this kind of behaviour. (2) If students training to become translators are used as subjects, TAPs may be used to find out where they have problems. The results of the analyses can then form a basis for translation pedagogy (Krings, 1988, p. 64; Kussmaul, 1989a passim; 1989b passim; Kussmaul, 1994 passim). One might argue that teachers of translation already know which strategies to recommend to their students. From years of experience they know what their students need. This may be true to some extent, but teaching experience shows that we sometimes draw the wrong conclusions from our students' translations. We may, for instance, have the impression
that students have problems with text-comprehension while, when we talk to them, we find that they actually have problems expressing what they had understood. TAPs can help us to see matters more clearly.

2. Choice of subjects and translation briefs

In the first studies (for instance, Gerloff, 1986 and 1988; Krings, 1986a and 1986b; Lörscher, 1986; Königs, 1986) the subjects were foreign language learners, and texts were chosen that were expected to entail a large number of difficulties. A translation assignment was not given, and a real-life translation situation was not taken into account. These conditions were, of course, completely different from those of professional translator training. When foreign language students are asked to translate, they are expected to show that they know the foreign language. For this purpose they often translate literally to please their teachers. In foreign language teaching the notion of a well functioning target text is normally non-existent or irrelevant. More recent studies (for instance, Hönig, 1988; Jääskeläinen, 1989 and 1993; Kiraly, 1990; Krings, 1987; Kussmaul, 1989a and 1989b; Kussmaul, forthcoming; Lörscher, 1992a and 1992b; Tirkkonen-Condit, 1989 and 1992), therefore, use semi-professional subjects, i.e. advanced students at translator training institutions, or indeed practising professionals. Usually a translation assignment is given, and texts are chosen that are translated in real life, or professional reality is at least simulated.

3. TAPs versus dialogue protocols

The method most frequently used consists of monologue protocols, i.e. experiments where one subject talks to himself or herself while translating a text. Since talking to oneself is normally not a natural thing to do, the subjects are given a chance to get used to this behaviour in a "warming-up phase." In order to make the situation less artificial, the experimenter may be present and listen "quietly but attentively" (Krings, 1986a, p. 56). In addition, the atmosphere in which the experiment takes place should be stress-free and the subjects should not have the feeling that they are being criticised for
their translations (Krings, 1986a, p. 56). As Krings (1986a and 1986b) points out, the think-aloud monologue method is very well suited for translation process research because there is a close affinity between translating and thinking aloud.

Since translation is by its very nature a linguistic process, the verbalisations externalise linguistically structured information available in short-term memory. Most criticism levelled against thinking-aloud data relates, in contrast, to those cognitive operations where non-verbal processes have to be verbalized (e.g. abstract problem-solving tasks; see Ericsson and Simon, 1980) (Krings, 1986b, p. 265f).

Monologue protocols are still predominantly the main tool for gaining access to the translation process. The artificiality that still remains has led some researchers (House, 1988; Höning, 1990 and 1991; Kussmaul, 1989a, 1989b, 1993 and 1994; Schmid, 1994) to get subjects to talk to each other. In a small-scale experiment, House compared monologue and dialogue protocols and found that monologue protocols contained a large amount of trivial data and that the process of selecting target language items, weighing alternatives one against the other, and deciding in favour of one particular translation equivalent remained unverbalized (House, 1988, p. 89).

In contrast, when talking in pairs, solutions to translation problems were negotiated and all partners in the pair thinking aloud sessions benefited in terms of incidental clarification of their own thoughts, and each individual's thoughts appeared to have been consistently shaped through the necessity of having to verbalize them (House, 1988, p. 93).

House concluded that the dialogue situation provided richer data than talking to oneself (p. 96). Later TAP experiments have shown, however, that the richness of data depends on the type of subjects and the translation brief, and, above all, on the priorities of the researcher. The data that one researcher find poor may seem rich to another researcher.
Another argument has been raised against monologue protocols. With increasing cognitive load, that is, when the subjects are deep in thought "they tend to stop verbalising or they provide less complete verbalisations" (Ericsson and Simon, 1980, p. 242). In TAPs this is reflected in pauses. On the other hand, subjects seem to stop verbalising when they have to do little thinking, i.e. with decreasing cognitive load. This is the case when they perform routine tasks. Since problem solving often has become routine for professional translators, protocols produced by them tend to contain few verbalisations (Jääskeläinen and Tirkkonen-Condit, 1991, p. 91). They do not talk about translating, they just translate, even without much pausing. If we think of cognitive load as a quantitative notion, it seems that at either end of the scale (much verbalisable thinking and little verbalisable thinking) there is no access to the translation process. It may, however, be important to see exactly which parts of the translation process are less verbalized by professional translators than others, in order to find out more about the nature of the process as such (Jääskeläinen, forthcoming).

Dialogue protocols may provide a way out of this dilemma. Thought may be brought to light by the questions subjects ask, by explanations, by arguing for or against solutions, by criticising and defending solutions etc., in short by the very nature of a normal communicative situation. The same is true of group protocols used by Höning (1990 and 1991) and Schmid (1994).

There are, however, problems with dialogue and group protocols as well. One may argue that while our aim is to observe what goes on in a translator's mind we are now not observing one mind at work but two or more, and that we record thoughts that would never have occurred to a single translator. This is true, but even if we use monologue protocols, we eventually may not want to find out what went on in one mind, but rather to draw conclusions from our observations of a sample of minds.

There is another argument which concerns group-dynamic processes; they may distort the results. One subject may assume a leading role not because he or she is a better translator, but because
of personality traits. Thus, solutions may be accepted not because they are better but because they are proposed by the more dynamic person. Also, one of the subjects may hold back his or her ideas for reasons of politeness — or even chivalry. In one case a male student accepted his female partner’s translation with the words: "O.K. I won’t quibble. You are the lady on our team." When analysing the dialogue protocols one should therefore take care to observe only those processes where both subjects take an equal part in solution-finding, or where the process is brought to an end without being impeded by the non-topical arguments and behaviour of one of the subjects.

One way of minimising these problems would be to choose "matching" subjects, that is, subjects where there is no psychological or social superiority of one over the other and where temperaments are fairly similar (Kussmaul, 1995). We should be aware of the fact, however, that we cannot completely control our variables. Furthermore, as can be concluded from the problems involved in both the monologue and the dialogue protocols, choosing the data collection method very much depends on what one wants to investigate. For instance, if the aim is to find out what are routine tasks for a professional translator and what tasks pose problems, a monologue protocol with long stretches of silence while routine tasks are performed and verbalisations when problems arise may give us just the clues we need.

Apart from these simultaneous types of data production there are also retrospective interviews. These are reports about actions that have just been performed. Subjects are asked immediately after having completed a translation task about the way they solved translation problems. Usually this method is used in combination with TAPs (Hönig, 1988; Kiraly, 1990). Kalina (1991) makes use of it in investigating simultaneous interpreting, since TAPs or dialogue protocols, naturally, do not lend themselves to this activity. What has been done in interpretation research is to interrupt the performance and prompt the interpreter to verbalize her thoughts at that point (Kalina, private communication).
The validity of all these methods has been questioned, basically for the reason that, in spite of their seeming closeness to the translation process, they nevertheless do not get close enough and still leave out too much. Thus, Hönig (1992) argues that only well-ordered thoughts are reflected in the protocols, and the often rather chaotic mental activities cannot be put into words, and that sometimes after-the-event rationalisations are produced, which do not mirror what actually took place (p. 10). Nevertheless, at the moment there is no better method available, and with more elaborated designs for experiments one might take better account of the drawbacks and arrive at ever more valid results (Krings, 1986a p. 522; Wilss, 1992, p. 209).

4. Models and heuristics

TAP data provide rich material on which a variety of hypotheses can be tested and research with a variety of aims can be pursued. In such rich data, however, we cannot find anything unless we at least vaguely know what we are looking for.

Königs (1991), for instance, mentions a number of types of actions that can be identified, such as macro-planning, identification of problems, solution of problems, associations, corrections, the use of dictionaries, and Jääskeläinen (1993) is interested in the translator's focus of attention, comprising both problematic and unproblematic processing (p. 101ff.).

In our search for relevant information, however, we ought to be careful not to mix up categories. Identification of problems, focus of attention, pauses, corrections and use of dictionaries are what might be called, surface phenomena, which can be noticed more or less easily, whereas macro-planning, solution of problems, associations and cultural transfer are more elusive. These can be called hidden phenomena. They can only be observed and classified after some interpretation and analysis has taken place. For instance, we can categorise a feature as a solution to a problem only after we have looked through the TAP and found that no further suggestion for a translation was made. Moreover, solutions may be good or
bad, and if we want to describe successful and unsuccessful processes, we must evaluate the solutions, which is another analytic procedure.

Such questions can only be answered if we have another type of heuristics available. In addition to categories like surface and hidden phenomena we need models for interpreting these phenomena. For example, to explain particular associations the notions of *bottom-up and top-down processes*, *scenes-and-frames* as used in psycholinguistics, and *convergent and divergent thinking* and *fluency of thinking* as used in creativity research (Guilford, 1975, p. 40) may prove to be helpful.

The TAP studies undertaken so far have not explicitly resorted to linguistic or psycholinguistic models in the interpretation of their results. In the Finnish studies, however, hypotheses derived from translation theories have normally constituted a starting point for TAP research projects. For instance, Jääskeläinen (1993) takes Höök and Kussmaul’s functional notion of *translation strategy* as a starting point and tries to find out if it can be observed in successful professional translating (p. 112f.).

Nevertheless, in more recent studies some models can be detected. Lörscher (1992a) uses the very basic distinction between *sign* and *sense* when comparing professional and non-professional translating (p. 152ff.), and when explaining non-professional behaviour he makes use of the concept of *interlanguage* as used in foreign language pedagogy (Lörscher, 1992a, p. 155). The interlanguage concept is also implied in Königs’ observation of learner-induced one-to-one correspondencies (Königs, 1987, p. 168f.), a notion which he uses to explain unsuccessful processes. Kalina (1991) uses psycholinguistic notions of the comprehension process, namely *frames*, *scripts* and *schemes* as well as *bottom-up and top-down processes* when analysing consecutive and simultaneous interpretation. Kiraly (1990, forthcoming) and Kussmaul (1989a, 1989b and forthcoming) also use psycholinguistic notions such as *bottom-up and top-down processes*, *schemata*, *scenes-and-frames* for their analysis of TAPs of semi-professionals.
When studying creativity in translation Kussmaul uses the standard four-phase model of creativity research. One might argue that the researchers' views are limited by the use of specific models and that they only investigate the TAPs from the perspective of the model. However, it may just be the purpose of the analysis to limit oneself to a particular point of view. On the other hand, if one wants to get a broad view of what goes on during translation, one must combine various models.

5. Some tentative results: Germany

5.1. General translation processes

In the early studies by Königs (1987), Krings (1986a and 1986b) and Lörscher (1986), who used foreign language students as their subjects, general comprehension, reverbalisation and monitoring strategies were recognized. Inferencing as a type of comprehension strategy was used when reference books turned out to be of no help. It appeared predominantly when the source text was in the foreign language. Spontaneous and fixed interlingual associations as a type of retrieval, i.e. reverbalisation strategy, appeared when no problems occurred. Königs (1986, p. 268ff.) found that these associations often resulted from learner-induced one-to-one correspondencies and could lead to erroneous translations as they did not allow for the context to be sufficiently taken into account. If interlingual associations could not be recalled, subjects used semantically related items such as both target and source language rephrases (Lörscher, 1986, p. 282f.), paraphrases, superordinate terms, archilexemes etc., or they made use of bilingual dictionaries in order to find equivalents. Their choice of dictionary equivalents, however, was often rather superficial; for instance, if there was more than one equivalent, they usually took the first one (Krings, 1986b, p. 273).

As far as monitoring is concerned, perhaps the most interesting form of behaviour observed by Krings is the "spot-the-difference

---

1. Preparation, incubation, illumination and evaluation; see Kussmaul (1993 and forthcoming).
strategy." The subjects compared individual target-language items with source language items, and if they found they did not match, they rejected the item, usually without taking the context into account.

The problem with these early studies was, as has been mentioned above, that the subjects were not training to become professional translators, and that in their translation courses such pragmatic matters as purpose of translation, target readership, type of medium in which the translation should appear were not discussed. Furthermore, the researchers usually refrained from evaluating what their subjects had produced. They, thus, did not observe the relationship between translation strategies and successful or unsuccessful results. Still, we may draw some pedagogical conclusions from their observations. It seems that their subjects' behaviour was often governed by lack of reflection (fixed associations and naive use of dictionaries) and that they would have benefited from an awareness of what they were doing, so that in the training phase translating would have become a more conscious and rational activity.

It was precisely the relationships between process and result which were focused on by Höning (1988 and 1990), Kiraly (1990) and Kussmaul (1989a and 1989b). They used semi-professionals as subjects and observed a number of partly successful and unsuccessful forms of behaviour. As far as top-down and bottom-up processes are concerned, Kiraly, who used monologue protocols, hypothesises that they take place in a subcontrol workspace which is not easily accessible to observation (Kiraly, 1990, p. 146ff.). In Kussmaul's dialogue protocols, some of these processes could be observed. To put it briefly and rather generally, a proper balance of top-down and bottom-up processes resulted in correct comprehension, whereas an imbalance between these processes resulted in miscomprehension of the source text and consequently in mistranslation.

Furthermore, the subjects often did not recognize that paraphrasing was a valuable technique of translating in cases where
word-to-word correspondencies were impossible. Although they used paraphrasing in the comprehension phase where it led to a correct understanding of the source text, they nevertheless did not accept it as a proper translation method for the reverbalization phase.

Another phenomenon typical of semi-professionals was a strange fear of interferences. Even if formally identical or similar words such as German "Desillusionierung" for English "disillusionment" could properly be used in a text, the subjects did not venture to do so, because they seemed to fear that such translations might result in severe blunders. As far as attitudinal factors are concerned, the semi-professionals' ambivalent attitude toward paraphrasing and their fear of false friends can be interpreted as a lack of self-confidence. As teachers we should do everything to strengthen their self-confidence.

Translation processes of professional translators have been investigated by Krings (1987) and Lörscher (forthcoming). Their common observations, some of which are corroborated by the studies conducted in Finland, are that the units of translation are larger among professionals than among foreign language students. This is in line with the professionals' global way of handling problems, i.e. their observance of larger sections of a text, whereas the language-learners solve problems in a "linear" (Krings) way and are concerned with problems of a local kind. Furthermore, professionals take a mainly "sense-oriented" approach rather than the "form-oriented" approach preferred by learners. The professional translators mainly, though not exclusively, check their production with regard to stylistic and text-type adequacy. This can be related to the observation that they have a larger number of variants at their disposal than do the learners. It, therefore, seems that the processes observed in professional translators are very much in line with the strategies recommended by many teachers in translator training institutions.
5.2 A special field: creative processes

For the investigation of translational creativity so far only the products, i.e. translations, have been analysed (Wilss, 1988, p. 114ff.). The evaluation of the products is no doubt important for observing creativity. If the product is not new and unusual, the processes leading up to it have not been creative. What remains to be investigated is how the creative process itself works. This can be done through TAP analyses.

On a small scale, Kussmaul (1993) has tried to observe how the four phases commonly regarded as constituting the creative process (Poincaré quoted by Preiser, 1976, p. 42f. and Ulmann, 1968, p. 21ff.) are reflected in the TAPs. The results are of a tentative nature and can best be formulated as hypotheses. The preparatory phase is closely related to text analysis, interpretation, considerations about function and purpose of the translation. All these are conscious mental activities. The incubation phase, which directly leads to illumination, is marked by psychological and physical relaxation, which is often achieved by some parallel activity, that is, the subjects’ doing things other than translating (such as going to the kitchen, eating a bar of chocolate, turning the cassette), which help to overcome mental blocks. As far as thinking processes are concerned, "fluency" (for the term, see Preiser, 1976, p. 60) manifests itself in the subjects mentioning a large number of target language synonyms or semantically related words for a given source language word. The logic of thinking leading up to illumination is largely "divergent" as opposed to "convergent" (for the terms, see Guilford, 1975, p. 40), which shows up in the fact that the subjects abstract meaning from linguistic form and put concepts into categories via "chaining" (for the term, see Lakoff, 1987, p. 91ff.). For example, in a text, a Mediterranean holiday situation was described from the point of view of a single woman. When the subjects translated the phrase "fanned by the flattery of murmuring machos" the joined "machos" together with

2. See footnote 1.
"bewundernden Blicken" (admiring looks). They here put together into a category things which, according to the structuralist theory of semantic fields, do not form categories at all. Nevertheless, here again there is some link that could be described as "men's admiration of women." It is especially this kind of linking of notions which constitutes divergent thinking.

As far as evaluation is concerned it becomes obvious in the TAPs that a strict division into phases is problematic. A critical attitude toward one's ideas is a prerequisite for recognizing good solutions and discarding unsatisfactory ones.

The results of these types of investigations can be used in translation pedagogy. The processes leading to creative solutions may be used as models of successful translating.

6. Some tentative results: Finland

The overall aim of the projects underway in Finland has been to identify the cognitive and affective factors that underlie the translator's decision-making and motivation. Think-aloud has been used as the main method of data elicitation. The major applications of this research are thought to be in translation pedagogy. There is published work, e.g. on problem-solving and decision-making in professional versus non-professional translation (Tirkkonen-Condit, 1988, 1990, 1992a and 1992b; Pöntinen and Romanov, 1989), automatization of processes (Jääskeläinen and Tirkkonen-Condit, 1991), features of successful processes and strategic behaviour (Jääskeläinen, 1990 and 1993) and routine versus non-routine processes (Laukkanen, 1993). Work in progress focuses on the attitudinal and motivational factors that account for quality in the human translator's performance (Jääskeläinen, forthcoming; Tirkkonen-Condit and Laukkanen, in press).

From the research carried out so far some tentative hypotheses have been formulated about the features of processes that lead to success in translation performance. Knowledge of successful processes is valuable in translator training and in any efforts aimed
at reforming existing professional practices. Good translation performance can be achieved in many ways, but certain features seem to characterize successful processes as well as the personalities behind success.

Successful translators seem to work more strategically, which shows in decision-making in that local decisions are subordinate to global ones. Successful translators also have relatively articulate subjective theories of translation which guide their performance. By virtue of their strategic touch they focus their attention, their conscious decision-making and their use of translation aids so that their investment in effort results in sufficient communicational gains. Since time is a limited resource, a good translator does not necessarily aim at an optimal performance but at a performance which is sufficient in a given communicative situation. The conclusion to be drawn from this is that translation pedagogy should pay more attention to translation briefs and marking. There should be more variety and realism in both. The student should be given an opportunity to develop a wide variety of strategies applicable to different translation briefs.

Less successful processes, especially those of novices and lay subjects, tend to be governed by local decision-making; they are thus linear by nature and proceed problem by problem, word by word or sentence by sentence. This often means a wasteful use of resources, since decision-making effort is not prioritized and strategically invested. These processes also tend to be linguistically rather than communicationally oriented. Good translators in turn are more prepared to use their world knowledge and their own inferences about text and text type as a basis for their decisions.

Less successful professional processes are characterized by lack of flexibility, which shows, e.g. in an attempt to use old routines in tasks which would require a fresh orientation. Thus, one translator in Jääskeläinen's experiment produced a poorly translated rough draft in twenty minutes, while the assignment called for a finalized, publishable text and did not spell out a time limit. In this translator's work place, translation speed was all that mattered, and
this routine was applied to the experimental task as well. Our think-aloud data in fact reveals a wide range of value judgements and ideas about professionalism in translation. These individual and subjective theories of translation constitute another area of our current research.

Our research on "non-translators," i.e. novices, non-professional and lay subjects has yielded some hypotheses which should be investigated. For example, it might be worthwhile recruiting specialists in particular fields to participate in intensive translator training on a master-disciple basis. What matters is the trainee's talent for linguistic and stylistic sophistication. Expert knowledge combined with bilingualism does not guarantee translating skills, but neither does training or professional experience in translation, if linguistic and stylistic talent is lacking. Some of our lay subjects seemed to possess such talent.

One method used in the investigation of the translators' values and individual theories as well as their motivation has been the analysis of the evaluative statements identified in the protocols. Among the attitudinal factors that seem to go together with success are personal involvement with the task, the translator's confidence in his own ability as well as professional pride and willingness to take responsibility for the communicative event. Personal characteristics such as tolerance of ambiguity, flexibility, realism and intellectual curiosity seem to contribute to success of performance.

Paul Kussmaul: Johannes Gutenberg-Universität Mainz, Fachbereich Angewandte Sprach- und Kulturwissenschaft, D-76711 Germersheim (Germany)

Sonja Tirkkonen-Condit: Kansainvälisen viestinnän laitos [Institut de communication internationale], Joensuun Yliopisto [Université de Joensuu] – PL 48.57 101 Savonlinna (Finland)
References


ABSTRACT — Think-Aloud Protocol Analysis in Translation Studies: This paper reports on think-aloud protocol (TAP) research in Finland and Germany. It discusses some methodological issues: choice of subjects, TAPs in a language-learning and in a professional context, monologue and dialogue protocols, the use of models provided by psycholinguistics. Two types of processes – successful and less successful ones – are distinguished and specified as to the subjects' comprehension and reverbalisation processes, their focus of attention, decision-making, monitoring, flexibility, creative thinking and attitude toward the task. Some tentative results of the research going on in Germany and Finland are presented. The results are expected to serve as hypotheses for the teaching of translation.