Methods of accessing metalinguistic awareness: a question of quality?

Maureen Ehrensberger-Dow and Alexander KüNZLI

Abstract

Various methods have been developed to access the metalinguistic awareness that translators have of what they are doing, and why, when they translate. In this paper, concurrent verbalization and cue-based retrospection are compared in a small study designed to monitor the translation processes of novices working into their first language (German) or strongest second language (English). A number of measures were devised to address the question of whether and how these two methods affect the quality of the translation process, the metalinguistic data, and the target text itself. An attempt is made to operationalize translators’ metalinguistic awareness of their translation practices and strategies when translating from and into their first or second language. Finally, the possibilities and limitations of the combinations of methods employed in this type of translation process research are discussed.

1. Introduction

In the Capturing Translation Processes (CTP) project, we are monitoring student and professional translators at various points in their careers. The data allow comparisons between the same students at the beginning of their translation degree program, the end, and 2 years post graduation as junior professionals; between professionals with different levels of experience (junior and senior); between student translators and professionals; between different language combinations; and between translation into L1 or L2.

---


2 Although many professional associations recommend that translators only translate into their L1, the reality of commercial translation demands means that many
Since we are interested in what translators do during normal work routines, we have chosen a rather complex methodology that nonetheless has the advantages of being ecologically valid and relatively non-invasive for the translators involved. It is based on progression analysis, a multi-method approach developed by Perrin (2003). Progression analysis, which combines ethnographic observation, interviews, computer logging, graphical representations of writing processes, and cue-based retrospective verbalizations, has proven valuable in studies of the writing processes of journalists, communication professionals, and schoolchildren (e.g. Gnach et al. 2007; Perrin 2003, 2006; Perrin & Ehrensberger-Dow 2008) and has proven applicable to investigations of translation processes (Ehrensberger-Dow & Massey 2008; Ehrensberger-Dow & Perrin 2009).

One of the main goals of the CTP project is to release expert knowledge, so it is important to have a reliable method of accessing metalinguistic awareness of this knowledge and to be alert to the advantages and disadvantages of various methods. In this paper, we therefore examine the differences in the quality of the data yielded by two such methods: concurrent verbalization (also known as think-aloud) and cue-based retrospective verbalization.

2. Concurrent and retrospective verbalization

Concurrent verbal reports have been used in psychology and language studies for over a century (Camps 2003). They have not been without controversy, in particular regarding the effect of the instruction to think aloud on the process under investigation, and the completeness of the protocols. According to Ericsson and Simon (1984/1993), concurrent verbalization yields valid data if the participants are asked to verbalize only the contents of their working memory rather than to describe or justify their thought processes, and if there is no delay between task execution and thinking aloud. Ericsson and Simon’s review of a large number of studies in psychology led them to conclude that concurrent verbalization does not influence the sequence of thoughts during the task at hand, although it may increase the amount of time needed to complete the task.

translators are expected to be competent in more than one language combination and direction, often translating into their L2 (Schmitt 1999).
Translation researchers have generally relied on this evidence, despite the fact that Ericsson and Simon (1984/1993) reviewed studies which above all investigated well-defined problems with predetermined correct answers. Translating, however, involves a large variety of pathways and solutions. It is possible that concurrent verbalization may in some way or another change the translation process. Thus, Jääskeläinen (1999: 151–158) compared eight translations produced with and eight translations produced without the think-aloud method and found some evidence that thinking aloud may influence the translator’s lexical decisions. Jakobsen (2003) observed that thinking aloud reduces translation speed and forces translators to process text in smaller segments. Kring (2001) compared a think-aloud group with a non-think-aloud group and noted that thinking aloud led to more target-text revisions.

On the other hand, Leow and Morgan-Short (2004) and Bowles and Leow (2005) more recently showed in second-language acquisition research that the critique that concurrent verbalization results in invalid data is based on anecdotal rather than direct empirical evidence. They found evidence of reactivity (the act of thinking aloud triggering changes) only for latency (the response time). Thinking aloud thus significantly increased the amount of time on task, but had neither detrimental nor facilitative effects on internal processing. Moreover, this observation applied to both ill-defined and well-defined tasks. Leow and Morgan-Short (2004) nevertheless point out that the questions of reactivity and veridicality (accurate representation of the thought processes) are fruitful areas of investigation. They argue that the potential for reactivity to be an issue might be higher if the dominant strategy employed in task execution by the control group is not the same as in the think-aloud group. Bowles and Leow (2005) moreover underline the potential for reactivity according to individual differences. Surprisingly, they also found metalinguistic verbalizations (i.e., when participants were required to justify their thought processes) to be nonreactive, which is in contradiction to one of the basic assumptions of Ericsson and Simon (1984/1993).

In translation studies, several researchers have made a case for data collection based on multiple methods, as a result of the possible reactivity of concurrent verbalization on translation processes and products. Retrospective reports, performed immediately after task execution, have
been identified as a useful alternative to concurrent reports. A tool that has often been used in combination with retrospective verbalization is Translog. Replaying a recording of the participant’s translation process and asking him or her to comment allows the researcher to collect cue-based retrospective data that is less susceptible to memory decay, a concern often put forward regarding non-cued retrospective verbalization (cf. Hansen 2006; for potential disadvantages of retrospection, see Göpferich 2009: 27–28).

Another advantage of retrospective verbalization is that different modes of expression (writing and talking) do not have to be used simultaneously; the talking has no impact on the translation process because it happens afterwards. Quantitative data on global aspects of the process, such as the time spent on the different phases of the translation process, can therefore more easily be compared when using retrospective verbalization than when using concurrent protocols. Retrospective protocols are also more cost-effective than concurrent protocols; in some research designs, their collection may add very little time to overall data collection (cf. Camps 2003: 215).

The above-mentioned points may be some of the reasons why concurrent verbalization is gradually being used less frequently in translation process research. Another reason might actually be related to one of the inherent advantages of concurrent verbalization: the wealth of data it yields. It is possible that researchers are deterred by the often limited resources for data analysis and interpretation. Göpferich (2008: 95) rightly points out that one of the shortcomings in the use of concurrent verbal reports in translation studies has been that analysis has often been based on paradigmatic cases and examples rather than a systematic, thorough examination of the phenomena studied. In what follows, we will illustrate how we try to address this challenge within the framework of our project.

3. Using progression analysis to compare verbalization methods

In investigations of translation processes, we collect data for progression analysis at four levels: (1) the situation surrounding the translation activity; (2) the translation movements and practices that the translators engage in; (3) the translation strategies that can be inferred from the translators’
comments about their translation processes; and (4) the translation products themselves.

In order to address the question of whether verbalization conditions have an affect on the data obtained, we have devised various measures of the process, the metalinguistic information, and the product. These include: differences in the patterns of the processes; distribution of time in each phase of the processes; distribution of comments in each phase; types of comments; linguistic markers of uncertainty in each phase; and global measures of the target texts. These will be presented within the respective level of progression analysis in the sections below.

3.1 Participants, setting, task

The first level of progression analysis considers the translation situation, including the linguistic and educational background information of the translators as well as the setting and the task demands. The five women participating in the present study, all paid volunteers, had German as their first language (L1) and English as their second language (L2). All had recently completed a four-year undergraduate translation degree program at our institute (Zurich University of Applied Sciences) and were enrolled in the first semester of a post-graduate interpreting program. They had also all taken part in data collection for the CTP project in their previous program, so were accustomed to having their translation processes monitored. They were told that this study was designed to test new eye-tracking equipment for the purposes of translation process research.

The setting for the translation processes was the usability lab at our institute; the computer that they were working at had the same MS Office environment that they were familiar with from other institute computer workstations as well as unrestricted access to the internet, institute on-line dictionaries, and other library resources. In addition to a familiarization session in the usability lab with the eye-tracking equipment, the participants took part in three recording sessions.

In the cue-based retrospection condition (RVP), they translated a short German news text into their L2, English, and commented afterwards

---

3 The eye-tracking data will not be considered further for the purposes of the present paper.
on what they had done while they viewed a screen recording of their processes replayed in real time. In the concurrent verbalization condition (TAP), they translated a comparable English news text of similar length and level of difficulty into their L1, German, and commented on what they were doing as they translated. The RVP and TAP sessions were at least one month apart. A student research assistant was present in both verbalization conditions but sat where the computer screen could not be seen easily, not only to increase the ecological validity of using verbalization as a method but also to prompt the participants to continue talking whenever necessary.

Because of the potential of cross-linguistic influence on thinking aloud in one language while translating into another, no attempt was made to counterbalance the two conditions and language directions in this study. Instead, the data from this group of students was compared to those from groups of students who had translated the same texts without concurrent verbalization (but had commented on their translation processes with cue-based retrospection).

3.2 Translation movements and practices

In the second level of progression analysis of translation processes, the translation movements are recorded and practices determined. Screenshot software and keystroke logging record the development of the emerging translation, all of the revisions to it as well as the search terms and electronic resources that were accessed during the translation process. The keystroke logging software developed by Van Waes and Leijten (2006) was used to record the position and order of insertions and deletions during the translation process, which was compressed in S-notation (Kollberg & Severinson-Eklundh 2001).

S-notation simplifies the detail analysis of the translation process; the broader pattern, by contrast, is traced in a progression graph (Perrin 2003) that reflects the development of the translation over time. The progression graphs in Figure 1 were made from the keystroke logging records of each

---

4 The recordings of the screen events as well as all of the verbalizations were done with Camtasia Studio 4.0 software.

5 In the third session, which occurred between the RVP and TAP conditions, the participants did a sight translation of a similar news text, either from English or German, as part of the test of the eye-tracking equipment.
of the participants’ translation processes in the RVP and TAP conditions (on the left and right, respectively). They indicate how the translators move with the cursor through the developing translation. The temporal sequence of revisions is shown on the horizontal axis; the spatial sequence of revisions in the final target text is on the vertical axis. If a translator completed a target text by only moving forward and never deleting anything or jumping to previous parts of the translation, the graph would be a straight line from the upper left corner to the lower right.

In contrast to the idealized progression of a straight line, the jagged lines for CTPI01 (first pair of graphs at the top of Figure 1) indicate that she made many changes throughout her translation, both in the RVP and TAP conditions. The patterns are quite similar, but there are far more revisions during the concurrent verbalization process. In each condition, the participant works on the emerging text in a translation phase until the final sentence of the target text before going back to the beginning of the text (position 0) to begin a revision phase. The similarity in pattern and relative number of revisions for this pair of progression graphs (about 50% more in the TAP condition/translation into L1) held for all of the participants in this study.

The second pair of progression graphs (from participant CTPI02) shows much smoother translation phases, with few jumps back and forth, and then two revision phases in both conditions, with those of the TAP process more clearly delineated than those in the RVP. The third pair of progression graphs (CTPI03) also shows quite a smooth translation phase in both conditions and many more revisions in the TAP condition. The jagged progressions of the fourth pair of graphs (CTPI04) are very similar to each other, suggesting that translators might have “signatures” detectable by comparing their patterns of movements in various translation processes (see Perrin & Wildi 2009, for a proposal on how to analyze and compare such processes statistically). The two progression graphs in the last pair (CTPI05) are again very similar to each other, except that the one from the TAP condition has a more prominent revision phase.
Figure 1. Progression graphs of translation processes in each verbalization condition
Since the focus of this study was to assess the quality of the data obtained with each method, we compared not only the number of revisions but also the amount of time spent overall in each phase of the translation process. In an attempt to address the confounding variables of translation direction and verbalization condition (i.e. German-English with RVP; English-German with TAP), the processes from these participants were compared to those of equal-sized control groups from the main CTP corpus. The control groups were matched for translation ability (based on final-year grades in the respective version) and had the same background as the participants in this study. They were all unpaid volunteers who had translated the same texts and provided retrospective verbalizations about them as part of their CTP project involvement two years previously.

Although the total times of the translation process vary somewhat among the four groups (see Table 1 below), the only significant differences are between the TAP condition and the RVP condition (Wilcoxon test, \( p<0.05 \)) and the TAP condition and the control groups (Mann Whitney tests, \( p<0.05 \)). This confirms the common observation that concurrent verbalization significantly extends time-on-task: the processes in the TAP condition were more than twice as long on average than the translation processes from the CTP corpus in the same direction, with the same source text, and retrospection (2433 vs. 1199 seconds).

Table 1. Average times for translation processes in various conditions

<table>
<thead>
<tr>
<th>Group</th>
<th>Verbalization type</th>
<th>Average time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTP RVP GE</td>
<td>retrospective</td>
<td>1463</td>
</tr>
<tr>
<td>CTP TAP EG</td>
<td>concurrent</td>
<td>2433*</td>
</tr>
<tr>
<td>CTP control GE</td>
<td>retrospective</td>
<td>1149</td>
</tr>
<tr>
<td>CTP control EG</td>
<td>retrospective</td>
<td>1199</td>
</tr>
</tbody>
</table>

\* \( p<0.05 \)

Because of the significant difference in time-on-task in the two verbalization conditions, percentage rather than absolute time has been calculated for the comparison of the phases of the processes: orientation, translation, and revision. The orientation phase was defined as everything until the first letter of the translation is typed,\(^6\) the translation phase until

\(^6\) The orientation phase cannot be captured effectively in the progression graphs, since they are based on movements within the emerging text, by definition after orientation is over.
the last character of the first draft, and the revision phase as everything after that. The graph in Figure 2 below shows that the relative proportion of time spent in the orientation phase was similar for all the groups (ranging from 4.7%–8.4% of total time). However, the percentage time spent in the translation phase was lower for the participants in this study than for the control groups because relatively more time was spent revising.

![Bar chart showing percentage of time spent in each phase of translation process](chart.png)

Figure 2. Percentage of time spent on each phase of translation process (n=5 per group)

The difference between the TAP condition and the RVP condition in time spent revising was significant (Wilcoxon test, p<0.05) as was the difference between the TAP group and the EG control group (Mann Whitney test, p<0.005). The hypothesis that this might be characteristic of translation into the L1 is not supported by the data from the control groups, since there was proportionately less time spent on revision in this direction.

The quantitative measures in the second level of progression analysis suggest that there is a difference in the quality of the translation process related to the verbalization method used. Concurrent verbalization appears to increase time-on-task not only overall but differentially as well, with proportionately more time being spent in the revision phase and many more revisions being made.

### 3.3 Metalinguistic information from comments

On the third level of progression analysis, we examined various characteristics of the metalinguistic data provided by the two verbalization methods. Sample extracts are provided in Table 2 from CTPI02 just after
she started typing the target texts in each condition. In the RVP condition, she talked about translating word-for-word, mentioning that she checked external resources for words she could not think of a translation for immediately, and stated where she looked for a particular source text word (indicated by italics in Table 2) and what she thought of the choices that were offered.

Table 2. Sample extracts from a retrospective and a concurrent verbalization (translated from the original German)

<table>
<thead>
<tr>
<th>Extract from a RVP (German ST, “English TT”)</th>
<th>Extract from a TAP (English ST, “German TT”)</th>
</tr>
</thead>
<tbody>
<tr>
<td>and I just translated, more or less word-for-word</td>
<td>“mehr gefährliche Produkte” banned has to be done in German differently somehow because you can’t, yeah there is no one word. That is, it somehow has to be several words</td>
</tr>
<tr>
<td>and at the same time I looked up the words that I, that I couldn’t think of right away</td>
<td>okay “die Europäische Kommission” is expected to … yeah expected but that can actually be left out, I’ll just put “wird am” … Thursday is, not Tuesday, Wednesday, Thursday right … Thursday</td>
</tr>
<tr>
<td>I wanted to look up Elektrogeräte in Pons afterwards, because I wasn’t quite convinced about the translations Leo offered. And I liked that one much better (CTP02_GE_2009)</td>
<td>a rapid rise in the number of dangerous goods</td>
</tr>
<tr>
<td>“wird am Donnerstag” a rapid rise, “einen schnellen Anstieg der Anzahl” but that “schnell” is no good somehow, but I’ll just write “einen schnellen Anstieg der Anzahl gefährlicher Produkte” (CTP02_EG_2009)</td>
<td></td>
</tr>
</tbody>
</table>

The verbalization extract for the TAP condition is longer, reflecting the greater time-on-task reported in Section 3.2. There is also far more verbalization of the source text and the target text (in italics and within quotation marks, respectively, in Table 2) in the TAP condition than in the RVP condition. In the TAP extract, the translator talks about the difficulty of translating a particular source text word banned, how to rephrase things to produce a natural target text, including an example of how she prompts herself to recall the correct translation for one of the days of the week, and her uncertainty about one of her solutions. In general, verbalization in the TAP condition evinces more code-switching, since the source text is often read aloud.

To address the question of the quality of the metalinguistic data provided by the two methods, all of the retrospective and concurrent
verbalizations were first transcribed using the Text Encoding Initiative (TEI) conventions, as suggested by Göpferich (2008: Ch. 4.3), as were the verbalizations of the two control groups. As an initial measure of quality, the distribution of verbalization in each phase of the translation process was determined for the RVP and TAP conditions and compared to that of the control groups. The amount of verbalization in each phase showed the same pattern as the time distribution, with more verbalizations made during the revision phase in the TAP condition (as measured by percentage of characters in the transcriptions).

Because there is more reading of the source text and target text in the verbalizations in the TAP condition, the relative proportion of verbalization in the various phases was compared after editing out the reading aloud of the source text and target text. Although there seems to be relatively more verbalization during the revision phase in the TAP condition than in the RVP condition, the difference is not statistically significant (see Figure 3).

![Figure 3. Percentage of verbalization in each phase, excluding reading aloud of source text and target text (n=5 per group)](image)

Any verbalizations not specific to the translation process were collated separately (i.e. “situation in general”). Slightly more comments about the situation, interview, etc. were made by the participants in the RVP
condition than in the TAP condition and by the control groups, but again the differences were not significant.

All verbalizations were coded and the resulting nineteen codes were grouped into eight categories. As noted above, the largest difference between the two conditions was the proportion of verbalizations that refer to or actually involve reading the source text or the target text (relatively low in the RVP condition, largest category in the TAP condition). Excluding those verbalizations, the relative proportions of the remaining seven categories (word choice, formulations, strategic decisions, external support, situation, commission, comprehension) are relatively similar in both conditions (see Figure 4).

![Figure 4. Proportion of each category of verbalization by participant and condition (excluding reading of source text or target text)](image)

However, there was far more explicit mention of actions in the retrospective than in the concurrent verbalizations. For example, reliance on external support in the TAP condition was sometimes as elliptic as "hmm announce, announce, okay ankünden" (impossible to code accurately without referring to the screen recording) whereas in the RVP condition the same translator was more likely to say something like "And then I decided to look in Leo". The utterances in the RVPs tended to be

---

7 The coding was done with HyperResearch software.
8 Leo refers to the on-line dictionary Leo (http://www.leo.org).
relatively long and often in the form of complete sentences and “X because Y” propositions (reflected in the higher proportion of the “strategic decision” category in Figure 4). The utterances in the TAPs, by contrast, were often single words or phrases interrupted by pauses, with many short propositions and little explanation of actions offered.

As a final measure of the quality of the metalinguistic data provided by the two verbalization methods, we focused on the participants’ expressions of uncertainty during their translation processes. Language involves a high potential for ambiguity. Consequently, there are a large variety of pathways and solutions to carry out the same translation task. As a result, Tirkkonen-Condit (2000) suggested that tolerance of ambiguity and uncertainty is necessary in translation. If translators reveal identifiable patterns of dealing with uncertainty in the translation process, it is moreover possible to look at uncertainty management as a dimension of translation strategies.

Tirkkonen-Condit (2000) proposed a number of linguistic uncertainty markers that we have investigated in our own work on translation and translation revision (e.g. Künzli 2006). Uncertainty phenomena contain expressions of epistemic and deontic modality, hedges on quality, questions, hypothetical statements as well as references to ignorance and uncertainty (see Table 3 for examples, translated into English, from the protocols from the present study).

<table>
<thead>
<tr>
<th>Examples (uncertainty markers, source text excerpts)</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) dangerous fungus that is a mushroom I think</td>
<td>epistemic modality</td>
</tr>
<tr>
<td>(2) it must be commissioner and it must be consumer protection too</td>
<td>deontic modality</td>
</tr>
<tr>
<td>(3) maybe it is a specific expression in the EU</td>
<td>adverbial hedge</td>
</tr>
<tr>
<td>(4) is she from Surrey or her trading office?</td>
<td>explicit question</td>
</tr>
<tr>
<td>(5) It should actually work if I do it like this</td>
<td>hypothesis</td>
</tr>
<tr>
<td>(6) it is not that difficult a text but appearances could be deceiving</td>
<td>conditional</td>
</tr>
<tr>
<td>(7) I wasn’t sure how to arrange the sentence elements</td>
<td>direct expression</td>
</tr>
</tbody>
</table>

We have investigated how uncertainty is expressed linguistically in the retrospective and concurrent verbalizations by the same translators, and to what extent uncertainty manifests itself in the different process phases.
Table 4 shows the number of uncertainty markers in the RVP and TAP conditions, along with their distribution in the different phases of the translation process and comparison figures from the control groups’ verbalizations. Because the rich and context-sensitive data of verbal reports sometimes make precise categorization difficult, only straightforward uncertainty expressions were included in the analysis.

Table 4. Distribution of linguistic manifestations of uncertainty phenomena in each phase of the translation process (in percentages and absolute numbers)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Retrospective verbalization (RVP GE)</th>
<th>Concurrent verbalization (TAP EG)</th>
<th>Control group (GE)</th>
<th>Control group (EG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientation</td>
<td>4.6% (5)</td>
<td>4.1% (8)</td>
<td>0% (0)</td>
<td>0% (0)</td>
</tr>
<tr>
<td>Translation</td>
<td>79.6% (86)</td>
<td>66.5% (129)</td>
<td>90.7% (49)</td>
<td>94.9% (37)</td>
</tr>
<tr>
<td>Revision</td>
<td>15.8% (17)</td>
<td>29.4% (57)</td>
<td>9.3% (5)</td>
<td>5.1% (2)</td>
</tr>
<tr>
<td>Total</td>
<td>100% (108)</td>
<td>100% (194)</td>
<td>100% (54)</td>
<td>100% (39)</td>
</tr>
</tbody>
</table>

Table 4 shows that whereas the share of uncertainty markers is comparable in the orientation phase in the RVP and TAP conditions, there are noticeable differences in the other phases. The concurrent verbalization protocols contain a higher proportion of uncertainty markers in the revision phase, whereas retrospective verbalizations are characterized by a higher proportion of uncertainty markers in the translation phase. This difference could be due to the fact that the translation phase is marked by a large proportion of local decisions (e.g. decisions regarding specific source- and target-text items that may be difficult to verbalize or be automated and therefore not leave any trace in the concurrent protocols). Uncertainties attached to these local decisions may, however, come to the surface and manifest themselves linguistically when translators view their processes retrospectively or by cues – from a distance, so to speak. Another explanation relates to the design of this particular study: the TAP condition was translation into the L1 and the RVP condition was translation into the L2.

It is also worthwhile remembering that the verbalizations took 39.1% less time in the RVP condition than in the TAP condition. However, the retrospective protocols contain 45.7% fewer uncertainty markers than the concurrent protocols. In our data, at least, uncertainty therefore manifests itself more prominently in the TAP condition. In other
words, the translators seem more certain about what they do in retrospective verbalization. The data for the control groups strengthen the hypothesis that the verbalization method used (RVP vs. TAP) plays a more important role regarding the degree of uncertainty expressed than translation direction (i.e. into L1 or L2). Thus, there are fewer instances of uncertainty markers in the retrospective than the concurrent verbal protocols. Other possible explanations for this difference may be related to a lack of motivation to comment on everything retrospectively or simply to forgetting uncertain episodes.

Finally, we observed that concurrent and retrospective verbalizations yield different types of uncertainty phenomena. In retrospective protocols, there are two types of uncertainty phenomena:

(1) regarding the translation process itself (e.g. “here, with this *beanstandeten* I have, I am not really sure what that means”)

(2) regarding the recollection of what happened in the translation process (e.g. “and then I don’t know what I did next”).

The data for the retrospective verbalizations in Table 4 only include the uncertainty phenomena attached to specific aspects of the translation process itself. In a future phase of our project, we will try to differentiate the object of uncertainty expressions in concurrent versus retrospective verbalizations in more detail.

### 3.4 Quality of the products

What we have referred to as the fourth level of progression analysis relates to the translation products. The question to be examined here is whether the increased concentration on revision in the TAP condition noted in the other levels of progression analysis is reflected in a difference in quality of the target texts. In order to assess whether the verbalization conditions have an affect on the product and its quality, two translation teachers for each version (G-E or E-G) evaluated the translations produced by the participants in this study. The translations were anonymized and presented to the teachers in random order: they were told that the translations had been collected as part of the CTP project but were not provided with any further details. The inclusion of an equal number of translations of the same
source texts by the control groups was meant to ensure an extra measure of objectivity in the evaluation process.

The teachers evaluated the translations independently, assigning a score between 0 and 4 in each of five broad categories in the grading scheme that is used in our institute. The two teachers of each version met to discuss their scores and agreed on an assigned grade for each translation, using the Swiss grading convention they were accustomed to (i.e. 1.0–6.0, where 6.0 is excellent and 4.0 is acceptable). Figure 5 shows the average grades assigned for the translations in each verbalization condition and for the control groups compared with the average final grades in their last translation courses.

![Graph showing average grades and final course grades](image)

Figure 5. Average assigned grades compared with final grades in translation courses

The average assigned grade for the translations carried out in the TAP condition is very close to the average final course grade whereas those performed in the RVP condition and by the control groups received lower grades on average. However, the relative difference between the RVP and TAP conditions is not significant: a closer examination of the participants’ assigned grades reveals that the lower average for the RVP condition was

---

9 The scheme groups errors into broad categories related to lexis, surface grammar features, sentence structure, completion, and strategic choices.
due to a single outlier with a very poor assigned grade. The four other participants received grades for their translations in the RVP condition that were very similar to their final course grades. Despite the additional focus on revision, we have found little evidence to suggest that translation products are better with concurrent verbalization. On the other hand, there is nothing in our results that suggests that the technique has a detrimental effect on the quality of the final products.

4. Conclusion

In attempting to operationalize translators’ metalinguistic awareness of their translation practices and strategies when translating from and into their L1 or L2, we have addressed the possibilities and limitations of two verbalization methods. Our results suggest that more information about revision might be accessible with TAPs, as indicated by translators repeating a source text element or reading a target text formulation aloud, whereas explicit information about the use of resources and reflections about strategies and dealing with translation problems might be more accessible with RVPs. In a workplace setting (i.e. with professionals who are being recorded in their normal workplace while they are performing their normal translation routines), RVPs of selected translation processes seem to be justified. In any case, several sources of data (e.g. participant observation, screenshot recordings, eye-tracking) are essential to be able to accurately interpret and categorize any type of verbalization.

A promising future direction for process research in translation competence and translation competence acquisition is to examine the information we obtain with concurrent and retrospective verbalizations in more detail and combine it with the information obtained with other techniques. We are convinced that such a multi-method approach to investigating translation processes will provide us with the differentiated understanding of translation competence that we need to make the right choices for the curriculum of translation training programs for students and professionals.
References


