Abstract:
Empirical research studies in Translation Studies have increased remarkably since the 1980s. In this article more than 50 studies on written translation are classified according to their objective, the sample and the measuring instruments used. The major problems shown by these studies are analysed and, finally, some steps that could help researchers in this sense are proposed.

1. Introduction: brief history of empirical research in Translation Studies

Until the 1980s, research carried out on written translation was mainly philosophical, abstract and inductive; in other words, it was based on philosophical or philological discussions, and theories were developed chiefly on the basis of experience. All the existing approaches or 'schools' give rise to a series of theories based on induction and observation which Neunzig refers to as "interpretative translation studies" or the "hermeneutic approach", in which the methodology consists mainly of analysis of a textual corpus or self-observation (Neunzig, 1999:4). Gile also refers to this type of research when he refers to “theoretical research which focuses on the intellectual processing of ideas” (Gile 1998:70).

The few examples of empirical research of which we are aware up to 1980 are sporadic, isolated and relatively unsystematic. It is only after that date that interest in empirical research in the field of translation begins to emerge, with a focus similar to that found in the cognitive sciences (experimental psychology, cognitive psychology, psycholinguistics, neurophysiology, etc.), in the search for methods and concepts that could be used to study the increasingly centre-stage translation process (See Danks et al., 1997).

In 1982 (Sandrock, 1982) and 1986 (Krings, 1986), the introspective method known as Thinking-Aloud Protocol, or TAP, was used for the first time in a study on written translation. The subject or subjects of the study were asked to verbalise their thoughts as they translated, that is, to think aloud, expressing the ideas that went through their minds. The subjects’ verbal accounts were taped and then transcribed as protocols for subsequent analysis. Since those early studies, and despite the criticism that they have attracted, to which we shall return later, TAPs have continued to be used as a technique for gathering empirical data on translation.

2. The proliferation of empirical studies in the field of Translation Studies

Since 1982, when the first doctoral thesis on translation to discuss the introspective data collection method known as Thinking-Aloud Protocol appeared (Sandrock, 1982)¹, and since the first study to use this technique in the analysis of the translation process (Krings, 1986), empirical research has
continued to use the method, as can be seen from the more than fifty research studies reviewed in
the present article.

Although this article is concerned with written rather than oral translation, we should also mention
the large number of empirical studies that have appeared in the field of interpreting, as shown by
the studies by Gran 1990; Lambert, 1992; Moser -Mercer 1994; Darò, 1994, 1997; Fabbro and
Darò, 1994; Kurtz, 1994; Padilla et al., 1994; Padilla, 1995; Padilla et al., 1995; Pöchhacker 1995a,
1995b; Schlessinger, 1995; Tommola, 1995; Braun and Clarici 1996; Gambier et al., 1997;
Jiménez, 1999. To these should be added the works by Gile (1990a, 1990b, 1991, 1995a, 1995b,
1995c, 1998)

which, although they do not refer to studies by the author himself, deal with many studies by other
authors and discuss research methodology.

To return to written translation, empirical studies have also been carried out to analyse the
translation process using techniques different from TAPs to collect data, such as the computer (see
and 1999) and translation diaries (Fox, 2000). There are also experiments the purpose of which was
to analyse the translation process, but rather other variables such as the use of dictionaries in
translation (Atkins and Varantola, 1997).

Almost simultaneously with the above-mentioned studies there have arisen critiques of the research
carried out, and there is currently a large number of publications, including research papers and
doctoral theses, devoted to discussing the methodology employed in research on Translation
Dancette and Ménard, 1996; Jääskeläinen, 1998; Kreutzer and Neunzig 1998; Neunzig, 1999,
PACTE 2000; Padilla et al., 1999; etc.

3. Research review

The large number of empirical research studies on the written translation process that have
appeared to date, together with their disparity in terms of the objectives, research methods and data
gathering techniques employed, mean that further concretion is required before those studies can be
discussed.

Our focus here will be those studies carried out using empirical methods since 1980, most of which
have attempted to analyse the translation process. Those dealing with research methodology in
general but which do not refer to any study in particular, are obviously excluded, as are those
studies discussing the research from a contrastive linguistics and cultural stand-point.

Notwithstanding our efforts to obtain information on as many research studies of this type as
possible, consulting the original publications, we have to admit that the huge proliferation of
bibliography in this field makes it very difficult to be exhaustive. The classification shown in table
1 brings together over fifty research works, arranged in chronological order. In the first column we
list the author and the bibliographical reference; in the second column we specify the research
objective; in the third column we give the sample, that is, the number of subjects participating in
the trials, and in the fourth column we specify the task carried out by the subjects and the technique
or techniques used to gather the data. Various abbreviations are used: N.A. means 'not available'
and indicates in the samples column that the authors do not specify how many subjects took part in
the study; with reference to the tasks, we have included, when specified in the study, whether the
translations were 'inverse', that is, into the foreign language (inv.) or 'direct', that is, into the

2
language of habitual use (dir.), and we have also abbreviated the word translation (trans.). With regard to TAPs, we distinguish between simultaneous (sim.) and retrospective (we only differentiate studies with retrospective in addition to simultaneous TAPs, since those using simultaneous TAPs are much more frequent), and we include, in those cases in which the authors provide the information, whether the protocols were recorded on audio tape or audio and video tape.

As can be seen from the classification in table 1, the research studies are extremely diverse, although a common denominator can be found: with regard to the data gathering techniques, TAPs are the most frequently used (75%). However, there are considerable differences as regards samples and objectives, as we shall now see.

**TABLE 1**

<table>
<thead>
<tr>
<th>RESEARCHER (REFERENCE)</th>
<th>RESEARCH OBJECTIVE</th>
<th>SAMPLE</th>
<th>DATA GATHERING TECHNIQUE</th>
</tr>
</thead>
</table>
| Dollerup, 1982          | Analysing translation techniques and strategies in translating from Danish to English. | 55 subjects | - Task: translations  
- Assessment and analysis of translations |
| Dechert and Sandrock, 1986 | Analysing the translation unit during translation process. | Case study | - Tasks: dir. translation  
- TAP (audio) |
| Krings, 1986 and 1987   | Analysing the strategies used by foreign language students to solve problems during the translation process. | 8 subjects | - Tasks: dir. & inv. Trans.  
- TAPs (audio)  
- Questionnaires |
| Jääskeläinen, 1987     | Analysing the translation process in translation students, comparing novices and experts. | 4 subjects | - Tasks: translations  
- TAPs (audio) |
| Krings, 1988           | Analysing the translation process in a professional and bilingual translator. | Case study | - Tasks: translations  
- TAP (audio) |
| Königs, 1987           | Analysing the syntactic and lexical problems during the translation process. | 5 subjects | - Task: dir. translations  
- TAPs (audio) |
| Gerloff, 1987          | Pilot study set to see what type of data can be obtained through TAPs and to create a coding system in order to ease the analysis of such data. | 5 subjects | - Task: dir. translations  
- TAPs (audio)  
- Assessment of translations |
| Gerloff, 1988          | Comparing the translation process and the quality of the product in translation students, professional translators and bilinguals. | 12 subjects | - Task: dir. translations  
- TAPs (audio and video) |
| Jääskeläinen, 1989    | Analysing the role of the brief in the translation process for novice and expert translators. | 4 subjects | - Tasks: translations  
- TAPs (audio) |
| Séguiot, 1989          | Analysing the role of the brief and the editing process in the translation process. | Case study | - Task: translations  
- TAP (audio & video) |
| Tirkkonen-Condit, 1989 | Analysing the criteria used in decision taking during the translation process. | 3 subjects | - Tasks: translations  
- TAPs (audio) |
- TAPs (audio) |
| Tirkkonen-Condit, 1990 | Analysing the differences of the decision taking process in translation students and professional translators. | 3 subjects | - Task: dir. translations  
- TAPs (audio) |
<table>
<thead>
<tr>
<th>Study</th>
<th>Methodology</th>
<th>Subjects</th>
<th>Task Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jääskeläinen and Tirkkonen-Condit, 1991</td>
<td>Analysing the automatic processes of novice and expert translators during the translation process.</td>
<td>7 subjects</td>
<td>Task: translations - TAPs (audio)</td>
</tr>
<tr>
<td>Kussmaul, 1991</td>
<td>Analysing creativity during the translation process.</td>
<td>2 subjects</td>
<td>Task: translations - TAPs (audio)</td>
</tr>
<tr>
<td>Lörscher, 1991, 1992 and 1996</td>
<td>Analysing strategies used to solve problems during the translation process.</td>
<td>52 subjects</td>
<td>Tasks: oral, dir. &amp; inv. translations - TAPs (audio)</td>
</tr>
<tr>
<td>Séguinot, 1991</td>
<td>Analysing learning strategies, whether conscious or automatic, in translation students.</td>
<td>195 subjects</td>
<td>Task: dir. &amp; inv. translations - TAPs (audio &amp; video)</td>
</tr>
<tr>
<td>Bélanger, 1992</td>
<td>Analysing cohesion in translation</td>
<td>Case study</td>
<td>Task: translation analysis (carried out by researcher)</td>
</tr>
<tr>
<td>Demers, 1992</td>
<td>Analysing several linguistic features in translations between English and French</td>
<td>Case study</td>
<td>Task: translation analysis (carried out by researcher)</td>
</tr>
<tr>
<td>Mondhal and Jensen, 1992</td>
<td>Analysing the representation of linguistic knowledge in the translation process.</td>
<td>N.A. (at least 4 subjects)</td>
<td>Task: inv. translations - TAPs (sim. &amp; retros., audio &amp; video)</td>
</tr>
<tr>
<td>Tirkkonen-Condit, 1992</td>
<td>Analysing the role of linguistic and cultural knowledge in the translation process of professional translators.</td>
<td>3 subjects</td>
<td>Task: inv. translations - TAPs (audio)</td>
</tr>
<tr>
<td>Englund, 1993</td>
<td>Analysing semantic changes in translation from Russian into Swedish due to the comprehension process.</td>
<td>N.A.</td>
<td>Task: dir. Translations, oral and written - TAPs (audio) - Computer</td>
</tr>
<tr>
<td>Fraser, 1993</td>
<td>Analysing the role of cultural transfer during the translation process.</td>
<td>12 subjects</td>
<td>Task: translations - Retrospective TAPs (audio)</td>
</tr>
<tr>
<td>Jääskeläinen, 1993</td>
<td>Analysing the possibilities of investigating translation strategies through empirical data.</td>
<td>12 subjects</td>
<td>Task: translations - TAPs (audio)</td>
</tr>
<tr>
<td>Laukkanen, 1993</td>
<td>Analysing the differences between routine and non-routine translations.</td>
<td>Case study</td>
<td>Task: inv. translations - TAP (audio)</td>
</tr>
<tr>
<td>Schäeffner, 1993</td>
<td>Analysing the influence of the translator knowledge in the comprehension process of the source text.</td>
<td>Case study</td>
<td>Analysis of a text and its translation (carried out by researcher)</td>
</tr>
<tr>
<td>Shreve; Schäeffner; Danks and Griffin, 1993</td>
<td>Analysing the role of the reading process in both the comprehension process and the translation process.</td>
<td>33 subjects</td>
<td>Task: text reading and problem identification - Computer</td>
</tr>
<tr>
<td>Tirkkonen-Condit, 1993</td>
<td>Analysing the translators’ sensitivity to cohesive structures in the source language.</td>
<td>3 subjects</td>
<td>Task: dir. &amp; inv. translations - TAPs (audio)</td>
</tr>
<tr>
<td>Dancette, 1994 and 1997</td>
<td>Analysing the comprehension process of the source text, paying special attention to extralinguistic and thematic knowledge.</td>
<td>5 subjects</td>
<td>Task: translations - TAPs (audio &amp; video) - Questionnaires</td>
</tr>
<tr>
<td>Fraser, 1994</td>
<td>Identifying features of professional translation practice (to be able to apply them in translator training)</td>
<td>21 subjects</td>
<td>Task: translations - TAPs (audio)</td>
</tr>
<tr>
<td>Dancette, 1995</td>
<td>Analysing the comprehension process of the source text in relation with linguistic competence, extralinguistic competence and the quality of the target text.</td>
<td>22 subjects</td>
<td>Task: inv. translation and text reading - Questionnaires</td>
</tr>
<tr>
<td>Kiraly, 1995</td>
<td>Analysing the translation process in relation</td>
<td>18 subjects</td>
<td>Task: translations</td>
</tr>
<tr>
<td>Researcher, Year</td>
<td>Study Title and Methodology</td>
<td>Subjects/Details</td>
<td>Tasks and Instruments</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------------</td>
<td>-----------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| Kussmaul, 1995 | Analysing the role of creativity in the translation process, taking into account the strategies used to solve problems dictionary use. | N.A.(between 8 and 15 subjects) | - Task: translations  
- TAPs (audio)  
- Questionnaires  
- Assessment of translations |
| Alves, 1996 | Analysing linguistic and cultural aspects of translations to be able to improve translation training. | 24 subjects | - Task: dir. translations  
- TAPs  
- Questionnaires |
| Dancette and Menard, 1996 | Analysing the comprehension process. | 5 subjects | - Task: translations  
- TAPs  
- Questionnaires |
| Königs and Kauffmann, 1996 | Obtaining data of the translation process as it is (not as it should be). | 3 subjects | - Task: inv. literary translations  
- TAPs |
| Mondal and Jensen, 1996 | Analysing the lexical searching strategies used by adult students of English as a foreign language. | 10 subjects | - Task: inv. translations  
- TAPs (audio) |
| Tirkkonen-Condit and Laukkanen, 1996 | Analysing the emotional aspect of the decision taking process during the translation process. | 4 subjects | - Task: dir. & inv. translations  
- TAPs (audio) |
| Atkins and Varantola, 1997 | Analysing the use of dictionaries during the translation process. | 32 subjects | - Task: translations  
- Questionnaires |
| Hansen, 1997 | Analysing the influence of Danish culture in translator training at a specific translation training centre (CBS) | 40 subjects | - Task: dir. & inv. trans.  
- Computer  
- Free writing |
| Kiraly, 1997 | Identifying problems posed by traditional training and assessment methods in translation didactics and looking for possible solutions. | 18 subjects | - Task: translations  
- TAPs  
- Questionnaires  
- Assessment of translations |
| Kussmaul, 1997 | Analysing the mental processes involved in problem solving that result in creative translations. | Case study | - Task: translations  
- TAPs (audio) |
| Neunzig, 1997b | Validate the computer as a simulator of a translation trainer. | 30 subjects | - Task: translations  
- Computer  
- Questionnaires |
| Halskov, 1998 | Analysing the comprehension process (without the use of reference books) in recently graduated translators, comparing their results in general and specialised texts. | 54 subjects | - Task: oral inv. translations, general and technique  
- Computer  
- Retrospective TAPs  
- Questionnaires |
- TAPs (sim. & retrospective, audio)  
- Direct observation  
- Computer |
| Lorenzo (en Hansen, 1998) | Analysing the decision taking process in inverse translation (comparing professional translators and translation students) | 12 subjects | - Task: inv. translations  
- Sim. & retrospective TAPs (audio & video)  
- Direct observation  
- Computer |
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Summary</th>
<th>Subjects</th>
<th>Tasks</th>
</tr>
</thead>
</table>
| Neunzig, 1998 | Obtaining reliable data about acceptance of the computer as a virtual translation teacher in the translation class. | 33 subjects | - Task: translations  
- Computer  
- Questionnaires |
| Roiss, 1998 | Analysing the inverse translation process in translation students. | 4 subjects | - Task: inv. translations  
- TAPs (audio) |
| González, Rodríguez and Scott-Tennent, 2000 | Analysing the effects of teaching translation strategies to students in their use of these strategies during the translation process. | 24 subjects | - Task: translations  
- Questionnaires  
- Written protocols  
- Translation diaries |
| Halskov, 1999 | Analysing how the manipulation of the source text to simplify it (without compromising its authenticity or its discursive integrity) can ease the comprehension and translation processes, comparing the translations of general and specialised texts. | 6 subjects | - Task: inv. translations, oral and written, general and technique (4 tasks)  
- Computer  
- Sim. & retrospective TAPs  
- Questionnaires |
| Hansen, 1999 | Analysing (through different techniques) the possibility of knowing the strategies used by the translator to solve problems as well as the degree to which the translator is conscious of the translation process. | 5 subjects | - Task: translations  
- Computer  
- Sim. & retrospective TAPs  
- Questionnaires |
| Livbjerg and Mees, 1999 | Analysing the influence of dictionary use in the quality of the target text in direct translation. | 5 subjects | - Task: dir. translations  
- Computer  
- TAPs |
| Jensen, 1999 | Identifying the strategies used, and analysing the process of translations made by professional and non-professional translators when working with time limit. | 6 subjects | - Task: dir. translations  
- Computer  
- TAPs |
| Lorenzo, 1999a | Comparing the translation process and the problem distribution in two different texts. | 3 subjects | - Task: inv. translations  
- Computer  
- Sim. & retrospective TAPs |
| Lorenzo, 1999b | Investigating the differences between direct and inverse professional translation. | 2 subjects | - Task: inv. translations  
- Computer  
- Sim. & retrospective TAPs |
| Waddington, 1999 | Analysing different assessment methods in translator training form Spanish into English. | 64 subjects | - Task: inv. translations  
- Assessment by different assessors, using different methods |
| Fox, 2000 | Analysing the use of translation diaries in the translation class. | 35 subjects | - Task: inv. translations  
- Translation diaries |

### 3.1 Objectives

Despite the diversity of objectives, various trends can be observed which allow the research to be assigned to different groups. Although some could be classified in various sections, we have selected the salient feature of each study:

1. Relevance of certain elements during the translation process:
   - Dechert and Sandrock 1986 (translation unit); Séguintot 1989 (translation brief and editing); Jääskeläinen 1990 (conscious attention); Englund 1993 (comprehension process); Shreve,
Schäffner et al. 1993 (reading in the comprehension process); Tirkkonen-Condit 1993 (cohesion structures in the SL); Dancette and Ménard 1996 (comprehension process); Königs and Kauffmann 1996; Halskov 1998, 1999 (comprehension process); Jensen 1999 (time limit); Lorenzo 1999a (comprehension process).

(2). Translation problems and strategies:

(3). Translation competence elements:
Kussmaul 1991 (creativity); Mondhal and Jensen 1992 (linguistic knowledge); Tirkkonen-Condit 1992 (linguistic and encyclopaedic knowledge); Fraser 1993 (cultural transfer); Laukkanen 1993 (routine); Schäffner 1993 (translator knowledge); Dancette 1994, 1997 (extralinguistic and specialised knowledge); Dancette 1995 (linguistic and extralinguistic competence); Kussmaul 1995 (creativity); Alves 1996 (linguistic and cultural competence); Tirkkonen-Condit and Laukkanen 1996 (affectivity); Atkins and Varantola 1997 (documentation); Livbjerg and Mees 1998, 1999 (documentation).

(4). Translation competence of the professional translator:

(5). Translation training:
Séguinot 1991 (learning strategies); Kiraly 1995; Hansen 1997; Kiraly 1997 (assessment); Neunzig 1997a, 1997b, 1998 (the computer as virtual translation teacher); Roiss 1998; Waddington 1999 (assessment); Fox 2000 (translation diaries).

(6). TAPs as a measuring instrument:

This classification according to objectives reveals that many of the research studies, in fact more than those devoted to the translation process, of which more than half deal with the comprehension process of the OT, focus on subcompetencies of translation competence. Another interesting fact is that only two studies attempt to probe the usefulness of TAPs as a data gathering technique, even though the vast majority of studies actually use them.

### 3.2. Measuring instruments

One of the greatest difficulties involved in conducting empirical studies in our discipline is the lack of measuring instruments specific to translation studies, since the majority of instruments which have been used in this field have been taken from other disciplines, as Neunzig points out:

Given the lack of tradition of an experimental approach in the field of translation studies, we face the additional difficulty, as we shall see, of not having a wide range of research instruments at our disposal. We will therefore have to design instruments specific to the experiment we wish to carry out and we will need to carry out either before or during each individual experiment studies to assess the validity of the measuring instrument which, at best, could prove to be normative studies; in other words, they might be used in subsequent experiments by other researchers for their own specific purposes (Neunzig, 1999: 25-26).
The measuring instruments which have been used so far in translation studies can be divided into two groups: instruments specific to the discipline and extraneous instruments taken from other fields of study:

a) Instruments specially designed for research in Translation Studies: translations and the computer (particularly, two programmes, Translog and Proxy)

b) Extraneous instruments: introspective techniques (TAPs of all kinds, interviews), questionnaires and psychophysiological measurements.

Regarding the computer as an instrument, recent years have seen the development of computer programmes such as Translog (created by Jakobsen, see Jakobsen 1998 and 1999) and Proxy, which act as data collection instruments in studies on written translation, and these programmes are increasingly being used in research.

With the exception of computer programmes and translations, there is a general trend within the field of translation studies to use measuring instruments taken from other fields, such as Psychology and Education Sciences. We shall not here go into any great detail concerning the introspective techniques, of which the most widely used in Translation Studies is undoubtedly that of TAPs, since we shall later analyse the criticism that it has attracted. However, there are other introspective techniques which have been used in empirical research in Translation Studies, such as translation diaries (see Fox, 2000) and *a posteriori* interviews attempting to obtain information on the process with the 'help' of the researcher, who questions the translator as the latter looks at his or her translation, or as translator and researcher observe together the translation process by means of a computer programme which reproduces exactly, key by key, the translator's process in creating the target text (see Hansen, 1998 and 1999).

The questionnaires or tests, included by some authors in the analysis of translations or of the translator's skills, are undoubtedly the oldest and most 'traditional' instruments used in research on written translation, since on the one hand, the product, that is to say, the written translation, is readily available and permits a large number of data to be accumulated relatively easily, as observed by Neunzig (2000). We understand the translation (act) to be the task given to the study subjects, and the correction or analysis of the product by means of a given method as, at best, an instrument, but in our opinion the latter should not be termed either questionnaire or test, since these two concepts have been clearly defined by psychometric definitions which do not include the senses of 'analysis' or 'correction'.

As for questionnaires, as understood in psychometrics, these have also been in use for some time (see Nida, 1964; Nida and Taber 1969, Snell-Hornby 1983, etc.), particularly in the analysis of the subject's comprehension. Although it may be supposed that the questionnaires used have evolved considerably, it is difficult to analyse them, since in the majority of studies in which questionnaires have been used, they are not actually included in the authors' published work, a fact which is a great obstacle to our understanding of the study carried out and which moreover precludes replicability. One exception to this observation is found in the studies by Halskov (1999) and Neunzig (1997b; 1998).

Studies have also been carried out using various psychological and physiological aspects as indicators, although the majority of such research has focussed on oral translation. From the field of cognitive psychology, measurements such as eye movements, positron emission tomography (PET)
and evoked potentials (EP) have been used, in addition to processes such as reading times, reaction and production times and accuracy measurements (for example, percentage of recall)\(^4\).

Finally, we should like to mention that various research projects have been carried out using multiple data-collection techniques, that is to say, using a combination of several instruments, which interpret as a whole the data obtained by the various means, such as video-recorded TAPs which provide visual and auditory information, \textit{a posteriori} interviews and a computer programme which records all the subject's actions. For specific examples of this type of research, see Hansen, 1998 and 1999, which includes studies by various authors from the TRAP research group of the University of Copenhagen.

### 3.3. Samples

We should like to touch on a point which is not usually included in the literature on empirical research, which is the differentiation between research using subjects who know that they are taking part in a study, that is, those studies in which a specific sample is taken to carry out the research, even if the research is conducted on a single subject (see, for instance, Krings1987; Laukkanen, 1993; Kussmaul, 1997), and research in which for the most part published translations are analysed and compared with the also published OT, in which the 'subject', or translator, is not aware that he or she is taking part in a study (see, for instance, Bélanger, 1992; Démers, 1992; Schäeffner, 1993). Both types of research are perfectly valid, but in our opinion these particular circumstances, which are sometimes not stated in the reviews(as, for instance, in Dancette and Ménard, 1996), should be made explicit, since the design will be completely different, depending on whether or not the researcher intervenes in the process, and also depending on whether or not the subject is participating in a previously designed experiment.

As can be seen from the research classification, there is great variety regarding the number of subjects taking part in the trials, since the latter range from case studies involving one or two subjects to experiments involving samples of up to 54, 64 and even 195 subjects (see Halskov, 1998; Waddington, 1999 and Séguinot, 1991, respectively), as well as a single study looking at an experimental group and a control group with a total sample of 24 subjects (see González, Rodríguez and Scott-Tennent, 2000). It may be concluded from the foregoing that the general criticism concerning the size of the samples in specialist publications may very soon cease to be an issue.

There is one interesting aspect regarding samples which is not reflected in the classification, and that is the previous training, or the background and characteristics, of the subjects taking part in the studies. As an illustration of this point, we shall cite three samples composed of supposedly professional translators.

Tirkkonen-Condit (1990) and Jääkseläinen (1989) compare the translation of experts or professionals with that of untrained beginners, but the subjects included as being representative of professional translators are in fact fifth-year University translation students. Although the translation competence of the latter should admittedly be more highly developed than that of their first-year colleagues, in our opinion they do not truly reflect the characteristics of a professional translator. Moreover, the number of subjects used in these two studies is very small: only one in that of Tirkkonen-Condit and two in that of Jääkseläinen. The samples in question are therefore unrepresentative and small.
Another case in point is that of Gerloff (1988), who chooses four subjects who are indeed representative of professional translators, since they have at least ten years' experience working in the field; however, the total sample, which might appear to be appropriate in terms of size (12 subjects), is in fact divided into three groups which represent untrained beginners, bilingual subjects and professional translators, respectively. The real sample within each group, therefore, consists of four subjects, making the sample representative but small.

The third example is the study by Fraser (1994), in which the sample is not only representative of professional translators but is also based on an adequate number of subjects, twenty-one, all of whom have proven professional experience.

These three examples illustrate the great diversity of samples found in the empirical research on Translation Studies, the studies ranging from those using heterogeneous and unrepresentative samples to others based on adequate samples. Although these characteristics are fundamental to any analysis of the results of the research, in many cases the analysis of the data obtained does not take into account the previous training of the ample subjects. This fact is all the more alarming in that, in many cases, the results are considered as observed 'general trends'.

The circumstances described above may have an adverse effect on subsequent research. Let us take the hypothetical case of researcher 'a', who decides to conduct a study and, on examining the existing literature on the topic that he wishes to investigate, discovers that there is a previous study carried out by researcher 'b', who concludes that the subjects, professional translators, have an observable tendency to behave in a particular way: they only consult bilingual dictionaries. When researcher 'a' comes to carry out his study, he bases it on the results of that previous study, taking it for granted that they are accurate and may be held for the whole population represented by the sample, that is professional translators. On the basis of this premise, researcher 'a' designs an experiment which aims to create as natural an environment as possible, and, since professional translators tend to use the bilingual dictionary, he provides only a bilingual dictionary on the table, next to the computer. When the experiment begins, five of the six subjects get up and ask for encyclopaedias, dictionaries of synonyms and other reference works that they need for documentation purposes.

It is obvious that in our hypothetical case, the experiment would have failed because the researcher would have had to interrupt it in order to go and look for the reference works requested by the subjects. If researcher 'a' were carefully to examine the study carried out by 'b', he would probably find that the subjects in the sample were not what he would regard as professional translators (but rather recent graduates of Translation with no professional experience), and that the sample was very small, with the consequence that the observed tendencies, although accurate for the subjects concerned, were not necessarily so for other professional translators.

This may be a crude and exaggerated example, but it illustrates the need to base results objectively on the data, without making rash generalisations, and also to use samples which are appropriate to the objectives of the research, since in many cases subjects are selected simply because they are easy to obtain, irrespective of whether or not they are the most suitable subjects for the design of the experiment.

4. Criticism of the existing empirical research studies

Empirical research on written translation has been the object of criticism from various points of view in the field of Translation Studies, due to the application of the introspective technique, the
introspective technique itself, the lack of objectives in the research studies concerned, the
deficiencies in the data analysis methods, the characteristics and the small size of the samples used
and also the unjustified generalisations made by the studies. However, before reviewing all these
criticisms, we should stress the great value of the TAP studies, since they represented virtually the
only line of research on the translation process to be carried out in the field of Translation Studies
in recent years and have opened up a new field of research which would otherwise have been
impossible.

4.1. TAPs as a data-gathering technique
Within the introspective method, the specific application of the TAP technique, which had already
been at the centre of a widespread controversy concerning its application in the field of psychology
(see, for example, Jääkseläinen 1998:266-267), has been criticised by several authors in the field of
Translation Studies. The latter argue that the verbalisation of automatic processes (in the case of
professional translators) and of processes unfamiliar to the subject (in the case of trainee
translators) is highly problematic. In this context, we might quote Presas: "Another aspect which
should be borne in mind is the fact that in the case of professional translators, the process may be
entirely automatic and therefore inaccessible by means of think aloud protocols." (Presas, 1996:24,
our translation). The same view is held by Hurtado: "Although these studies represent a step in the
direction of analysing translation strategies, in our opinion the question remains unsolved as a
result of the areas of confusion arising from the analysis carried out using TAPs: first of all, the
intrinsic difficulty of the method of analysis employed; the difficulty experienced by professional
translators in verbalising highly automatic activities, and the additional difficulty experienced by
trainee translators, in whom learning strategies and translation strategies coincide and need to be

Dancette and Ménard (1996:142) also consider that “Il y a des limites théoriques et
méthodologiques à ces approches d’enregistrement du sujet sur le vif: nature très fragmentaire de la
verbalisation et probabilité de grandes distorsions entre le processus réel (boîte noire) et la
verbalisation.”. In this connection, we might also cite the criticism raised concerning the
introspective method by the members of the TRAP research group (Hansen, 1998:62-63), who
doubt that it is possible to carry out two complex tasks simultaneously (translating and thinking
aloud) without the one influencing and modifying the other. Fraser (1996a:67), who in turn refers
to opinions expressed by Ericsson and Simon (1980: 218), Mann (1982:95) and Zimmermann and
Schneider (1987), also adduces reasons for doubting the efficacy both of the introspective and the
retrospective methods, as does Jiménez (1999:118-120). Finally, Bell (1998:189) also refers to the
difficulties involved in attempting to observe a mental activity, a view shared by Dancette:

On a practical level, however, we must address the difficulties inherent in the methodology
of ‘observation of processes.’ Processes are not visible; only clues to such processes are
visible. But these signs are not an exact reflection of what is going on in the translator’s
mind because it is likely that many, if not most, processes remain unconscious. (Dancette,
1997:85).

Some authors who have used TAPs have attempted to solve these problems. For example,
Tirkkonen-Condit and Jääkseläinen try to overcome the problem of thinking aloud by carrying out
studies with professional translator subjects (three free-lance translators and one teacher of translation) who, according to the researchers, are unaffected by the problem of what the researcher might or might not wish to hear, since three of the four study subjects are used to thinking aloud as part of their professional translating activity (Tirkkonen-Condit and Laukkanen, 1996:47). However, the problem of automatic processes remaining unverbalised is still not solved, since automatic processes, by definition, occur at a level beneath the 'conscious' mind and therefore cannot be verbalised.

Toury raises other problems relating to TAPs, such as the relevance to Translation Studies of the results of the studies carried out using this data gathering technique:

The validity of introspective data for the study of cognitive processes has often been questioned, but most of the objections seem to have been disproved in an admirable way. In fact, it has been so much as claimed that, of all mental processes, it is translating which is most suitable for verbal reporting. As Hans P. Krings (1987:166) put it, ‘thinking aloud while translating is an almost natural type of activity to which most of the criticism leveled at verbal report data does not apply’. Be that as it may, my own concern here is not with psycholinguistic validity at all, but with the relevancy of the technique from the point of view of translation studies. (Toury, 1991:63)

There are numerous publications which are critical of TAPs, but as we have already dealt with the points that are criticised, we shall merely mention other reference works in this connection: Shuttleworth and Cowie, 1997:171-172; Jääskeläinen, 1998: 265-269; Bell, 1998:189; Kiraly, 1995:39-51.

Our own position with regard to TAPs is that they have serious methodological shortcomings. If they are to be used, therefore, they should be combined with other information gathering techniques and the data thus obtained should be used with the necessary caution. We should like to conclude this section with a quotation from Neunzig (2000:97), which sums up both the criticisms and the advantages of using introspective methods:

The Think-aloud method, which has become so fashionable (...) is appropriate to inductive research in clarifying matters relating to the translation process: large quantities of data are obtained concerning influence factors, underlying translation strategies, decision-taking, possible regularities in tackling a problem, etc. which help us to formulate or refine hypotheses (...). However, as a method for validating hypotheses they pose serious problems of environmental validity and extrapulpability. (...) In our opinion, another instrument is needed in order to ensure, at least from a theoretical point of view, the objectivity and environmental validity of the experiment, since these criteria are central to empirical research in the field of translation studies in general and the didactics of translation in particular.

4.2. Study samples
Another of the problems associated with the vast majority of research studies carried out to date in the field of translation lies in the samples used in carrying out the studies. To begin with, the samples are criticised as being inappropriate because they are not representative of the population
being studied. In this connection, Hurtado (1996) argues as follows: "Another problem lies in the subjects analysed, since in some cases (Lörscher, Krings) they are students not of translation but of foreign languages (...)" (Hurtado, 1996:12). In fact, the authors of the studies themselves are often aware of this problem and openly admit that it exists, as is the case of Jääskeläinen:

The pronounced differences observed in the professional translator’s behaviour could have been explained by the ‘occupational hazards’ involved in experimentation: in such a small sample of subjects, individual, even idiosyncratic, behaviour, may assume a more dominant role than in larger samples; it could thus be possible that in choosing the subjects we simply came across a group of exceptional personalities. Consequently, differences in personality or different cognitive styles, for instance, could have explained the lack of shared features in the professional processes (Jääskeläinen, 1993:100).

In referring to a sample of four professional translators, the above-mentioned author also touches on another weakness of the samples, in this case their representativeness: the small number of subjects in the samples, which is insufficient for the results of the research to be generalised. Toury's opinion on the matter is expressed thus: “To be sure, samples of insufficient size seem to be a common weakness of all experiments carried out so far. Nor has this fact gone unnoticed by the researchers themselves” (Toury, 1991: 52)

Such a view is widespread and well-founded, as can be seen from the following examples of research using samples which are either inappropriate to the research objective or are very small: (1) Krings, 1986: the subjects are foreign-language students carrying out translations into the foreign language; (2) Königs, 1987: the subjects are a professional translator and four foreign-language students; (3) Krings, 1987: there is only one subject; (4) Tirkkonen-Condit, 1992 and 1993: the subjects are three teachers of translation, of whom two carry out a translation into the foreign language while the third translates from the foreign language into his own language; (5) Laukkanen, 1993: there is only one subject; (6) Tirkkonen-Condit and Laukkanen, 1996: a new analysis is carried out on the data already obtained using Tirkkonen-Condit's subjects (1992) and Laukkanen's subject (1993), that is to say, a total of four subjects, of whom one is a professional translator, two are teachers of translation as well as being free-lance translators, and one is a translation teacher; (7) Jääskeläinen and Tirkkonen-Condit, 1991: the study is based on trials carried out by these two authors in 1989, using three subjects (see Tirkkonen-Condit, 1989) and four subjects (see Jääskeläinen, 1989), that is to say, using two different groups of students, each of whom translated a different text, and whose translations are directly compared; (8) Dancette and Ménard, 1996: this study is based on the TAPs produced by five students, of whom two are also professional translators, who are the same subjects whose protocols were analysed by Dancette in 1994 and 1997; (9) Königs and Kauffmann, 1996: the sample consists of three students of French who carry out a literary translation into the foreign language.

4.3. Experimental design
Another common criticism is that studies lack an experimental design, which is indispensable in order to be able to draw well-founded conclusions. Fraser, for example, argues that this is the key to the future of research using introspective methods:
To conclude, it seems clear that, if properly designed and differentiated, the introspective method can deliver valuable and interesting insights into a variety of linguistic activities at a number of levels. Design and differentiation do, however, seem to be the keys to making the findings capable of really relevant exploitation. (Fraser, 1996:77, our underlining).

In this connection, there are some authors who use the data collected in order to carry out one analysis after another, with different objectives. This shows that there has been no previous experimental design, since first of all all data are collected by means of a translation of a text, and then, as new hypotheses arise, new data are added by means of new analyses of the same, previously conducted studies. We might refer to authors such as Lörscher, Krings, Tirkkonen-Condit, Dancette, etc., but we shall merely cite the example of Jääskeläinen, who in 1987 carried out a study using TAPs as a data gathering technique, in which four students from the first and fifth years of a University degree course in Translation translated a text from English. First, the protocols were qualitatively analysed in order to compare professional with non-professional translation, and assuming that the translations of the fifth-year students were professional in quality (Jääskeläinen, 1987). Two years later, the same protocols were analysed, this time focusing on the differences in the role of the translation brief (Jääskeläinen, 1989); one year later, the protocols of the same four students were used yet again, this time adding the Thinking-Aloud Protocols of another eight subjects, of whom four were professional translators with 10-15 years' experience and four were University graduates with a good knowledge of English. The twelve protocols were analysed again, focusing particularly on the data which reflect the use of translation strategies, to ascertain whether the strategies could be analysed by means of empirical data or whether the processes were too highly automatic to permit such analysis (Jääskeläinen, 1993). The development of all these research studies shows that no prior research design existed, since they used protocols obtained at different times, using the translations of two different texts, according the a posteriori interests of the researchers. In the context of the problem arising from a lack of an adequate experimental design, one of the weaknesses of the empirical studies carried out using introspective techniques is the object of study. On the one hand, there is a general tendency on the part of researchers not to give a clear definition of their object or of the hypothesis they are attempting to test:

For many of the experiments which were so far applied to translation are characterized precisely by a certain uncertainty as to what they had been designed to do. The questions underlying them were often very general, even vague, and the investigators were all too ready to settle for general ‘insights’ rather than insist on answers which would bear directly on either theory or ordered application.” (Toury, 1991:63)

On the other hand, criticism has also been made of the fact that the objectives so far chosen for research have not been sufficiently stringent, since the studies have usually focussed on gathering data at specific times during the translation process of potential translators (including language students) or of professional translators, but without showing the progression involved between the two levels of competence, and without studying the necessary steps involved in the trainee translator's transition to the status of professional translator:
Thus, differences of performance on the ‘inexperienced-experienced’ axis have now been substantiated on the basis of TAPs too (see, e.g. Krings, 1988; Tirkkonen-Condit, 1989; Jääskeläinen, 1989). However, it is not enough to simply take heed of those differences, not even by establishing distinct variants of the psycholinguistic model for inexperienced vs. experienced translators. It is also vital to give some thought to the process whereby one type of ‘translator competence’ evolves into, or maybe is replaced by, the other. What we need to know, in other words, is not only what it takes to perform translation, but also what it takes to become a translator. (Toury, 1991:62)

This view is shared by Fraser: “Yet the transition from one to the other [from translation students to professionals] is not tackled systematically in any of the studies, although it is a major area of interest in both Séguinot’s and Tirkkonen-Condit’s work.” (Fraser, 1996:75).

Finally, criticism has also been made of the lack of objectivity in both the analysis of the data gathered in the course of the research and the subsequent generalisation of the results. Hurtado, for example, states the following: "In any case, the fundamental problem lies in the analysis of the results obtained. Thus, Lörscher's analysis (...) is distorted by the belief that the translation problems (...) are merely lexical, syntactic and lexicosyntactic. For his part, Kiraly (...) arrives at the paradoxical result that there are no differences between professional translators and trainee translators.” (Hurtado, 1996:12).

Gile observes that researchers are rashly prone to generalising the results of their studies, without taking into account the other aspects of experimental design and internal validity:

A second major weakness in Translation and Interpretation research lies with extrapolation. The flaws referred to here are not the technical problems caused by ‘convenience sampling’ in statistical inference (...) nor are we referring to the cases in which non-professionals are selected for experiments on professional practice. The problem is that even when professionals are given tasks that can be considered valid as Translation or Interpreting tasks, even in observational studies, which deal by definition with real I/T tasks, researchers tend to extrapolate somewhat imprudently.” (Gile, 1991:165)

Conclusions
In studying the research that we have classified, we may conclude that empirical studies carried out to date present a number of shortcomings which are summarised in table 2.

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TAPS AS AN INSTRUMENT</strong></td>
</tr>
<tr>
<td>-thinking aloud while translating is an unnatural activity which interferes in one or other of the two tasks. -if the subject is aware that he is being observed, he usually modifies his conduct, which means that the researcher is observing an unreal situation.</td>
</tr>
</tbody>
</table>
-introspective techniques, which are used to develop or refine hypotheses, should not be used to test them.

Criticisms made of empirical research.

We feel that it is important to bear these shortcomings in mind if we are to avoid repeating them in future and so contribute to further our knowledge of the complex phenomenon of translation. We believe that there are various steps which could assist us in this endeavour. First, the creation and validation of measuring instruments and adequate experimental designs; second, the use of other, complementary measuring instruments when TAPs are used as the principal instrument in a study; third, the selection of representative samples; finally, a rigorous approach to the use of the measuring instruments and to the interpretation and analysis of the resulting data.

1 Although Sandrock's thesis appeared in 1982, the results were not published until 1986 (see Dechert & Sandrock, 1986).
2 Programme adapted by the PACTE research group to study the translation process.
4 For more detailed information on the use of this type of instrument in Translation Studies research, see Padilla, 1995; Padilla, Bajo, Cañas and Padilla 1994; Padilla, Bajo, Cañas and Padilla, 1995; Padilla, Bajo, and Padilla, 1999.
5 See Orozco 1999, 2000a and 2000b, a doctoral thesis in which three instruments to measure the acquisition of translator competence are constructed and validated and an experiment is designed to compare the effect of different methodologies in first-year Translation and Interpreting students.
REFERENCES


PADILLA, Presentación; BAJO, Maria Teresa; PADILLA, Francisca CAÑAS, J.J (1994). “Procesos de memoria y atención: Hacia una teoría cognitiva de la interpretación”, in Sendebar, 5.


