Anne Schjoldager*

Translation for language purposes: Preliminary results of an experimental study of translation and picture verbalization

Abstract
The relative merits and drawbacks of translation for language purposes, i.e. translation as a teaching and testing tool in the foreign-language (L2) classroom, are largely unexplored empirically, and opinions differ widely. Within the research field, scholars often ignore this use of translation or tend to condemn it out of hand. Other authors may be more positive towards translation in the L2 classroom but tend to base their views on personal beliefs and individual teaching experiences. With this paper I hope to contribute to the ongoing debate with some useful empirical fuel. The point of departure is a claim implicit in much criticism, namely that translation leads learners to commit errors that they might not otherwise commit. The paper describes the background, subjects, assumptions, procedure, and preliminary results of a small-scale experimental study of L2 translation (Danish into English) and picture verbalization in L2 (English). According to some preliminary results, learners who translated made more and other errors than learners who wrote a comparable picture verbalization. This is probably due to enhanced interference from Danish, their first language (L1).

1. Introduction
In the Routledge Encyclopedia of Translation Studies, Guy Cook gives an interesting overview of the relation between translation and language teaching. According to this, there is an obvious discrepancy between the practice and theory of language teaching:

“Despite the widespread popular assumption that translation should play a major and necessary part in the study of a foreign language, twentieth-century theories of language teaching and learning have at best ignored the role of translation, and at worst vilified it. From the

* Anne Schjoldager
English Department
The Aarhus School of Business
Fuglesangs Allé 4
DK-8210 Aarhus V

The turn of the century onwards almost all influential theoretical works on language teaching have assumed without argument that a new language (L2) should be taught without reference to the student’s first language (L1)”. (Cook 1998: 117)

Thus, whereas many teachers and policy makers apparently believe in translation as a useful component of the L2 curriculum, most theorists are inclined either to ignore it or to regard it as an inadequate, even harmful, teaching tool and an unreliable testing tool (see also Pedersen 1974, Snell-Hornby 1985, Titford & Hieke 1985, Keith & Mason 1987, Sewell & Higgins 1996, and Malmkjær 1998). When the issue is discussed in the literature, arguments are often theoretical, based on subjective views, or derived from individual teaching experiences, with only a few empirical studies to either support or condemn translation for language purposes.

My own investigation is an attempt to contribute with some useful empirical fuel to the ongoing debate, which I find an important one. In the following pages I shall first give a brief overview of the field and then describe the background, subjects, assumptions, procedure, and preliminary results of a small-scale experimental study of L2 translation (Danish into English) and picture verbalization in L2 (English).

2. Translation for language purposes: Opinions and studies

As already suggested, most scholars who discuss the merits and drawbacks of translation for language purposes do this from the perspective of practical language teaching, and it is hardly surprising therefore that many arguments put forward in the literature in favour of or against translation may be categorized as opinions rather than well-documented facts. I shall now give a brief overview of some of these opinions.

The basic argument in favour of translation in the L2 classroom tends to be that translation, being a controlled task, is a time-saving way of teaching and testing L2 proficiency (e.g. Duff 1989/1992: 7, Sewell 1996: 142; see also Campbell 1998: 58, who sees L2 translation as a controlled variety of L2 writing). A related argument is that translation promotes students’ linguistic agility and accuracy, perhaps mainly at the advanced level (Snell-Hornby 1985: 21). It is also suggested that translation is useful as a tool in contrastive studies, making students aware of structural
differences between L1 and L2 (e.g. Snell-Hornby 1985), thereby helping them to develop strategies to avoid L1 interference in their L2 production (Sørensen 1988 and 1990; see also Sørensen 1991).

The basic arguments against translation in the L2 classroom tend to contradict the above-mentioned arguments in its favour: First of all, it is claimed that translation has very little to do with real language practice, the true aim of language teaching and learning. Rather than helping students to communicate in L2, translation merely gives them an opportunity to study the L2 language system (e.g. Coleman 1986 and Marsh 1987). Furthermore, because translation in the L2 classroom tends to be a word-for-word exercise, it is often accused of giving students the impression that there is also such a one-to-one relation between languages (e.g. Lado 1964: 53-54), which tends to lead to flawed and unidiomatic L2 production (e.g. Irons 1998: 26). (For a more detailed discussion of these pros and cons of translation for language purposes, see Schjoldager forthcoming.)

Table 1. Selected arguments

<table>
<thead>
<tr>
<th>In favour</th>
<th>Against</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation increases L2 proficiency.</td>
<td>Translation merely increases knowledge of the language system.</td>
</tr>
<tr>
<td>Translation increases linguistic awareness.</td>
<td>Translation gives learners wrong ideas about language.</td>
</tr>
<tr>
<td>Translation counteracts L1 interference.</td>
<td>Translation enhances L1 interference.</td>
</tr>
</tbody>
</table>

It would be unfair to say that the field is completely unexplored empirically, and I shall now draw attention to the results of a few important studies. Scherer & Wertheimer (1964) is the most extensive empirical
study that I have come across. In their book, the authors report on a two-year experiment involving 300 beginning students of German at the University of Colorado, in which they investigate the relative merits of the grammar-translation method, which includes translation as a teaching tool, and audiolingualism. Their hypothesis is that audiolingualism is a superior teaching method. The conclusion is somewhat surprising: after two years of language instruction, there is no significant difference between those who were taught by the audiolingual approach and those who were taught by a “traditional” approach as far as proficiency in reading and L2 translation (i.e. translation from English into German) is concerned; those who were taught by the traditional approach excel in L1 translation (from German into English), whereas those taught by the audiolingual method only excel as far as speaking is concerned.

In an exploratory study, Berggren (1972) compares two methods of teaching “fairly simple everyday language” (Berggren 1972: 2), namely L2 translation (Swedish into English) and what she refers to as intensive reading in L2 (English), an “accepted form of close reading with a view to language points only” (Berggren 1972: 10). Though appreciating that it may be difficult to ascertain the extent to which learners, in her case first-year university students, are influenced by the two methods, Berggren (1972:19) concludes that L2 translation does not have a negative effect.

---

1 It was only recently that I discovered this study myself. I am grateful to Viggo Hjørnager Pedersen for drawing my attention to it.

2 As far as I can see, Scherer & Wertheimer have no precise definition of the grammar-translation method, but I assume that the method investigated is similar to that which Richards & Rodgers (1986: 3) aptly describe as “a way of studying a language that approaches the language first through detailed analysis of its grammar rules, followed by application of this knowledge to the task of translating sentences and texts into and out of the target language”.

3 As a means of characterizing the audiolingual method, Scherer & Wertheimer (1964: 9) cite Pierre Delattre’s dictum from 1947: “[..] A language is first of all ‘speech’ – a system of sounds transmitted directly from mouth to ear and produced by automatic reactions of the speech organs. The functioning of those automatic reactions depends on the linguistic habits of the speaker, and it is the acquisition of those habits that must come first [..]”. See also Richards & Rodgers (1986: 44-63) for an analysis of the background, theory, and implementation of audiolingualism.

4 I am grateful to Marie Källkvist for sending me a copy of this work report.
Buck (1992) compares the efficiency of L1 translation (English into Japanese) as a test of L2 comprehension with other tests, and finds – much to his surprise and regret – that L1 translation works fairly well.

Other studies are more critical of translation in the L2 classroom. Klein-Braley (1987) investigates the objectivity, reliability, and validity of L2 translation (German into English) as a testing tool of L2 proficiency, comparing it with other tests. According to Klein-Braley’s (1987: 128) findings, the translation tests did measure language proficiency, but they were not the best possible procedure available; and as far as validity was concerned, the translation tests always scored lowest.

Larsen (1990) also focusses on translation as a testing tool and explores the possible discrepancy between university students’ performances in L2 translation (Danish into English) and other tasks, namely summarizing and essay writing. Her findings more or less confirm her initial impression that translation is no reliable test of L2 proficiency.

Källkvist (1997) supports Klein-Braley’s (1987) stance that L2 translation, which in Källkvist’s case is translation from Swedish into English, cannot stand alone as a test of general L2 proficiency (see also Källkvist 1998). In Källkvist (forthcoming), she compares the relative efficiency of L2 translation (Swedish into English) and non-contrastive monolingual exercises (in English) in a course of English grammar, finding that those who did translation performed slightly worse. Källkvist therefore concludes that translation may not be superior to monolingual teaching methods.

### Table 2. Selected studies

<table>
<thead>
<tr>
<th>Supporting translation</th>
<th>Not supporting translation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teaching</strong></td>
<td></td>
</tr>
<tr>
<td>Scherer &amp; Wertheimer (1964)</td>
<td>Källkvist (forthcoming)</td>
</tr>
<tr>
<td>Berggren (1972)</td>
<td></td>
</tr>
<tr>
<td><strong>Testing</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Larsen (1990)</td>
</tr>
<tr>
<td></td>
<td>Källkvist (1997)</td>
</tr>
</tbody>
</table>
3. Background

What are we to believe? Is translation a useful component of language teaching and testing, or is it not? As we have seen, opinions differ widely, and we still lack necessary evidence to answer this question. For the time being, we must therefore be content to conclude that the issue is undecided: translation may have some value as a teaching tool and it is probably rather unreliable as a testing tool, except perhaps as a test of L2 comprehension (cf. Buck’s 1992 study).

Personally I am inclined to agree with this conclusion. As far as translation as a teaching tool is concerned, I strongly believe that this may indeed be quite valuable, especially if the grammar-translation type of approach is supplemented (if not replaced) by a functional-communicative one, using for instance Nord’s (1997) version of the skopos theory as a conceptual and analytical framework (as suggested in Schjoldager 2002). In this way, the negative aspects of translation, for instance enhanced L1 interference (see section 5. below), may be outbalanced by its positive aspects, especially the fact that translation is a controlled writing task forcing students to venture into new areas of L2. As far as translation as a testing tool is concerned, I am equally certain that it cannot be a reliable tool. I can see translation as a test of translation skills, but I cannot see it as an efficient test of general L2 proficiency. First of all, translation is bound to require more than is required for a monolingual L2 production task. Klein-Braley (1987) puts it like this about her own data:

“[..] the translation tests measure the construct less adequately than the other, far more economical, test procedures, and there is good evidence for suspecting that they are two-dimensional, measuring other skills, probably translation skills, in addition to general language proficiency.”
(Klein-Braley 1987: 129)

(For some interesting papers on how to define, develop and assess translation competence, see for instance Schäffner & Adab (eds) 2000.) Secondly, and this seems to be implicit in most criticism of translation for language purposes, the enhanced risk of L1 interference may possibly lead students to make more and other errors than they might otherwise commit.

In view of the general lack of evidence within the field, I decided to concentrate on one interesting factor, namely that of enhanced L1
interference. By means of a simple comparative analysis of errors occurring in L2 translations and picture verbalizations in L2, the aim of my investigation is to answer these research questions: Do learners who translate commit more and other errors than learners who write comparable picture verbalizations? If they do, may this be due to enhanced L1 interference?

4. Subjects

The subjects of the investigation were third-year secondary-level students (gymnasieelever) from Risskov Amtsgymnasium, specializing in languages (Sproglig linje højt niveau), and third-year university students taking the compulsory translation course in the department of English at the University of Aarhus, my employer at the time of data collection (the autumn of 2000).

Subjects were chosen out of larger groups for these characteristics: They had Danish L1 and English L2; they were brought up in Danish families in Denmark; and they had learned English as a foreign language in Denmark (as opposed to acquiring it abroad). Several had to be excluded from the corpus because they lacked these necessary characteristics. No streaming was carried out and the division into groups, i.e. those who translated and those who did the picture verbalization, was accidental.

Table 3. Distribution of subjects

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Translation</th>
<th>Picture verb.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary</td>
<td>10</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>University</td>
<td>16</td>
<td>16</td>
<td>32</td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>27</td>
<td>53</td>
</tr>
</tbody>
</table>
5. Assumptions

As already mentioned, I assume that learners who translate into L2 may commit errors that are caused by an enhanced influence of L1 interference in translation. There are at least two reasons for this: (1) Learners who translate from an L1 text are led to rely on processing via L1, and (2) learners tend to be ‘hypnotized’ by the source text (ST) wording, which prevents them from making full use of their L2 mastery (Larsen 1990: 97). (See also Wilss (1982: 207), who makes a similar point about L1 translation.) I shall refer to errors that are deemed to be the result of an enhanced influence of L1 interference in translation as translation-induced errors.

On this basis, I chose to compare translations into English (L2) with picture verbalizations in English. I assume that picture verbalization and translation are sufficiently comparable because both are controlled tasks forcing subjects to express a given series of ideas. The comparability of the two tasks is further enhanced by the fact that the ST of my investigation is itself a verbalization of the pictures in question - or, rather, a synthesized version of four native-speaker verbalizations, as explained below. On the other hand, it is also assumed that picture verbalization is sufficiently different from translation because it does not enhance L1 interference: (1) Learners who do a picture verbalization are not led to rely on processing via L1, and (2) there is no verbal ‘hypnosis’ in picture verbalization.

6. Procedure

First I had to find some suitable pictures. These should present a series of ideas that could be verbalized, and the material should be sufficient for me to devise a plausible and coherent ST. Various comic strips, which have been used in other investigations, especially within psycholinguistics, did not quite fit my requirements – mainly because they do not offer enough verbal material. After a rather long search, I found some large colour pictures by Pia Thavlov (1996), in a charming children’s book about a family outing. Thavlov’s pictures were ideal for my purpose, for two reasons: (1) the story was told exclusively by means of pictures, and (2) each picture offered many details about the family’s various activities, which could be verbalized into a coherent text.
Then a Danish ST (426 words) was constructed on the basis of four native-speaker verbalizations. After some experimentation in a pilot study, the data collection could begin (in the autumn of 2000). For various reasons of research methodology and manageability, subjects did either a translation into English (L2) or a picture verbalization in English. Both groups were given a maximum time limit, 75 minutes and 65 minutes respectively, so as to enhance the degree of comparability between the two tasks. To make the assignment as plausible as possible, subjects were instructed to pretend translating/writing a story for a children’s programme on the radio, and, to rule out external factors, I made my subjects work without dictionaries in an exam-like situation. Admittedly, most students were not used to be given a brief like this for a translation assignment and none of them had been trained to work without dictionaries, but I have no way of assessing the possible impact of this. (For more details about the background and procedure of this investigation, see Schjoldager forthcoming.)

7. Preliminary results
Before I mention some preliminary results, I should perhaps point out that some (but not all) picture verbalizations differ somewhat in length and contents from the translations, as a few students who did picture verbalizations seem to have written more personalized texts, adding for instance some imaginative details. However, as the initial analysis involves no direct comparison between the target texts (TTs), the relative comparability between the two kinds of texts will not be discussed further at this stage. Also, it is important to note that the analysis below is based on the occurrence of errors of language, as opposed to translational errors, which would involve a comparison with the ST or pictures. Finally, though at some later stage I intend to employ a panel of colleagues to help me assess the TTs, the analysis below comprises my own tentative marking of errors.

1) Are there more errors in the translations? Yes. I found more errors in the translations than in the picture verbalizations, as shown in tables

---

5 See Nord (1997: 30) for a discussion of the concept of brief within the skopos
4 and 5. Both groups share this tendency, though the university students make fewer mistakes in both tasks, which is hardly surprising.

Table 4. Total number of errors

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Translation</th>
<th>Picture verb.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary</td>
<td>418</td>
<td>241</td>
<td>659</td>
</tr>
<tr>
<td>University</td>
<td>291</td>
<td>184</td>
<td>475</td>
</tr>
<tr>
<td>Total</td>
<td>709</td>
<td>425</td>
<td>1134</td>
</tr>
</tbody>
</table>

Table 5. Average number of errors

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Translation</th>
<th>Picture verb.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary</td>
<td>41.8</td>
<td>21.9</td>
<td>31.4</td>
</tr>
<tr>
<td>University</td>
<td>18.2</td>
<td>11.5</td>
<td>14.8</td>
</tr>
<tr>
<td>Total</td>
<td>27.3</td>
<td>15.7</td>
<td>21.4</td>
</tr>
</tbody>
</table>

2) Is this overrepresentation of errors translation-induced? Maybe. Apart from the above-mentioned difference in absolute numbers, so far I have found no evidence that the two tasks produce significantly different patterns. However, though these are very preliminary results,
I would like to mention that two categories of errors seem to be more prevalent in the translations, especially in those by the secondary-level students, namely – what I refer to as – non-translation and unEnglish, as exemplified in tables 6 and 7.

Table 6. Example of non-translation

<table>
<thead>
<tr>
<th>TT</th>
<th>fishing</th>
<th>tang</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST(L1)</td>
<td>fishing</td>
<td>tang</td>
</tr>
<tr>
<td>L2</td>
<td>fishing</td>
<td>rod</td>
</tr>
<tr>
<td>Explanation?</td>
<td>L1 processing</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Example of unEnglish

<table>
<thead>
<tr>
<th>TT</th>
<th>childwagon</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST(L1)</td>
<td>barmegno</td>
</tr>
<tr>
<td>L2</td>
<td>pram(UK)</td>
</tr>
<tr>
<td></td>
<td>baby carriage (US)</td>
</tr>
<tr>
<td>Explanation?</td>
<td>negative transfer</td>
</tr>
<tr>
<td></td>
<td>‘direct’ translation</td>
</tr>
</tbody>
</table>
Non-translation occurs when a ST item (L1) has been transferred unchanged into the TT, i.e. a translation or a picture verbalization, and this is deemed to be unwarranted. UnEnglish occurs when a TT item is deemed to be non-standard in the target language (L2). Admittedly, these categories are somewhat fuzzy, just as the explanations in tables 6 and 7 must be regarded as intuitive. However, the fact that errors categorized as non-translation and unEnglish seem to be more prevalent in the translations indicates that (some of) these errors may be due to enhanced L1 interference, i.e. translation-induced.

8. Conclusion and discussion

In this paper, I have summarized some general opinions in favour of and against translation for language purposes and I have presented the results of a few important empirical studies. Both opinions and studies form the background of my own study of translation and picture verbalization, which is the focus of this paper. My preliminary results indicate that the task of translation may invite more errors than picture verbalization, that certain errors seem to be more prevalent in translation, and that these errors may be translation-induced, i.e. due to enhanced L1 interference in translation. This tendency is more marked in the secondary-level group than in the university group.

The bearing on translation as a testing tool of general L2 proficiency is quite clear: my preliminary results support that translation is rather unreliable. However, since my subjects were asked to do either a translation or a picture verbalization, my investigation cannot show how individual students would have performed in both tasks. It would therefore be interesting to devise a study in which subjects had to perform in both translation and picture verbalization in comparable ways, so as to ascertain whether the same students show a similar tendency of making more and other errors in the translations. It would also be interesting to study some inherent differences between the two tasks. My data show that picture verbalizations vary in degree of comparability with the translations in terms of length and contents, but my study cannot show why this is. Above I have suggested that the overrepresentation of errors in the translations may be translation induced, but there may well be other important factors, for instance the fact that picture verbalization invites more freedom.
My findings are less clear regarding the bearing on translation as a teaching tool. On the one hand, inviting students to commit errors that they might not otherwise commit may be construed as ‘dangerous’ because such errors may spill over into other L2 tasks. On the other, as many teachers will tell you, students, at least advanced ones, need to be made aware of systematic differences between their L1 and L2, and the marking of language errors in L2 translations is one method of doing this, perhaps an efficient one.

My own investigation draws specific attention to one important drawback of translation for language purposes, namely that of translation-induced errors, while other studies document other fragments of the issue. Some studies show that translation may be quite valuable, and some conclude that it is not (on its own), with rather strong evidence against translation as a testing tool. However, whatever we decide we still lack conclusive evidence: if we decide to abandon this kind of translation without some definite proof that it is really as harmful as many scholars claim, we run the risk of throwing out the baby with the bath water; if we decide not to abandon it and condone a long-standing tradition without reflection, we risk wasting our time and effort and we risk the exclusion of more efficient teaching and testing methods (cf. Källkvist forthcoming); and, similarly, if we decide to recommend translation for language purposes, say for advanced students and especially within a functional-communicative framework, as I myself am inclined to do, this is risky too because we still do not know if this is appropriate or not. In other words, more extensive, long-term studies are certainly required for us to know the true merits and drawbacks of translation as a teaching and testing tool of L2 proficiency.

References


Thavlov, Pia 1996: Familien Bartoldy. Frokost i det grønne. Published for Libero by Gyldendal, Denmark.

