

Post print of *Target* 25:1 2013. pp. 140–154, DOI [10.1075/target.25.1.11mee](https://doi.org/10.1075/target.25.1.11mee)  
This article is under copyright and the publisher should be contacted for permission to re-use or reprint the material in any form.

Title: Sound effects in translation

Authors:

Inger M. Mees [im.ibc@cbs.dk](mailto:im.ibc@cbs.dk)

Barbara Dragsted [bd.ibc@cbs.dk](mailto:bd.ibc@cbs.dk)

Inge Gorm Hansen [igh.ibc@cbs.dk](mailto:igh.ibc@cbs.dk)

Arnt Lykke Jakobsen [alj.ibc@cbs.dk](mailto:alj.ibc@cbs.dk)

Postal address (all four authors): Department of International Business Communication,  
Copenhagen Business School, Dalgas Have 15, 2000 Frederiksberg, Denmark.

## Sound effects in translation

### *Abstract*

On the basis of a pilot study using speech recognition (SR) software, this paper attempts to illustrate the benefits of adopting an interdisciplinary approach in translation. It shows how the collaboration between phoneticians, translators and interpreters can (1) advance research (2) have implications for the curriculum (3) be pedagogically motivating and (4) prepare students for employing translation technology in their future practice as translators. In a two-phase study in which 14 MA students translated texts in three modalities (sight, written, and oral translation using an SR program), Translog was employed to measure task times. The quality of the products was assessed by three experienced translators, and the number and types of misrecognitions were identified by a phonetician. Results indicate that SR translation provides a potentially useful supplement to or alternative for written translation.

### **Keywords**

Interdisciplinarity, written translation, sight translation, oral translation, speech recognition, pronunciation, productivity

## **1. Introduction**

Consider the following three scenarios.

### **Scenario 1**

Peter has just broken his arm and has been told by the doctors that he can't use it for six weeks.

Nobody would be pleased to receive information of this kind, but it is particularly inconvenient for Peter, who makes his living as a translator and depends on being able to type. He could dictate the translation, but he would then have to hire a secretary to type it out, which is an expensive solution.

In addition, he hates not being able to make revisions as he goes along. It seems such bad luck, because Peter is working on a translation of an annual report, and he has several major projects in the pipeline. What can Peter do to get his translations done in time?

### **Scenario 2**

Emma is a talented language student specialising in translation and interpreting. Along with the actual courses dealing with these fields within Translation Studies, the curriculum also requires her to take a number of other subjects, including courses in linguistics, such as phonetics. "It's so boring", says Emma, "Why do we have to be taught all this useless stuff about the articulation of consonants, vowel length, weak and contracted forms, and the differences between Danish and English vowels? When I'm abroad I can communicate with English-speaking people without any problem whatsoever." Emma wants more "relevant" subjects which can make life at university more interesting and also assist her in her future career. What can the university do to motivate students like Emma, who likes the practical components of the translation and interpreting programme, but finds it difficult to see the point of a number of other language subjects such as phonetics?

### Scenario 3

David is a full professor of marketing and consumer behaviour. He works at a Danish university, where he has recently become director of a research centre with a global network. He is delighted because he has just received funding for a large-scale project involving researchers from many different countries. So, in future, David's working language will be English, and all his articles and reports will have to appear in English. David is a very proficient *speaker* of English, being fluent, having a wide vocabulary and native-like pronunciation, but *writing* poses a major obstacle. For despite his many talents, David has one great handicap: he is dyslexic. How can David be helped?

For these three scenarios, based on real life cases (names have been changed), one solution could be to introduce speech recognition (SR) software. Peter would be able to dictate his translation and have it converted into a written translation on the spot. Emma would suddenly see the relevance of phonetics. And David would overcome his difficulties with writing in English. Both Peter and David would perhaps save time, and all three might even deliver output of an equally high – or improved – standard, and experience more variation in their work.

Introducing SR into translation (see Jurafsky and Martin 2000: 235–284 for an introduction to SR technology) presupposes the integration of multiple disciplines simultaneously. “Speaking” rather than “writing” a translation itself involves crossing the borders between translation and interpreting since the translation is produced orally, as in interpreting, but is visible on the screen, as in translation (see Gile 1995, 2004; Agrifoglio 2004; Lambert 2004 for discussions of interpreting and sight translation in comparison with written translation). Thus we are dealing with a hybrid. In addition, it is necessary to draw on the field of phonetics (pronunciation) in order to discover how

one can facilitate and improve the dictation of a translation, and specifically how one can reduce the number of misrecognitions to a minimum. To gauge the potential for collaboration, the authors of the present article, who represent different areas of research (translation, interpreting and phonetics), decided to embark on an interdisciplinary adventure.

As stated by Leijten and Van Waes (2005: 741) “most research in the field of speech technology has dealt with the technical improvement of speech recognition and the role of phonology”.

However, more recently there has been interesting work on the applications of speech recognition in a range of other fields, e.g. studies on how SR affects the writing process (Leijten and Van Waes 2005), on error correction strategies employed by professional SR users (Leijten, Janssen and Van Waes 2010), and on the use of SR technology in the domain of re-speaking in live subtitling (Luyckx, Delbeke, Van Waes, Leijten & Remael 2010). However, to our knowledge there are no studies such as the one below on the use of SR in translator training.

## **2. Methodology**

A two-phase pilot study was undertaken in which we examined the impact on the translation process and product of using an SR system, as compared with typing a translation and producing a normal sight translation.

Two batches of experiments were carried out, the first taking place in March 2010. Sixteen MA T&I students out of a class of 22 volunteered for the experiments and translated Danish texts into English, their L2, under three different conditions (sight, written, SR). The data for two participants subsequently had to be discarded owing to technological difficulties, leaving 14 students for analysis. After the first round, half of the participants (experimental group) worked with the SR

program at home while the other half did not (control group). In December of the same year, the same participants translated comparable texts under the same three conditions. The students did not have access to the Internet, dictionaries or other support. Although this restriction made the set-up less ecologically valid, it was considered a necessary constraint as students would probably have spent more time on information retrieval under the written than the oral conditions, which would have given a skewed picture of the task times in the different translation modes.

In the SR modality, the participants were instructed not to use the keyboard but only to avail themselves of the oral commands. The motivation for this was that three recording programs were running simultaneously during the SR task (keylogging, eye-tracking and SR), and pilot runs had shown that keyboard activity in the SR task caused Translog to crash. More importantly, if students were allowed to use the keyboard during the SR task, there was the risk that they would revert to typing whenever they encountered problems with the SR system,<sup>1</sup> and this would have defeated the purpose of the exercise. See Dragsted, Mees and Hansen (2011: 16) for more detail.

Asking the participants to dictate in their L2 is not as curious as it may seem at first sight. In Scandinavia there are “good grounds for referring to English as a *second* language rather than a foreign language” (Phillipson 2003: 96). Since English and Danish are both Germanic languages exhibiting close correspondences in their sound systems, Danes encounter fewer dictating and pronunciation difficulties than speakers from many other countries. In Denmark, English is taught from age 9 and pervades Danish society (films are subtitled, many companies use English as a lingua franca). Furthermore, since Danish is not a world language, most Danish translators are forced to work bi-directionally if they want to make a living. Translator training in Denmark thus focuses equally on translation in both directions (see Dragsted et al. 2011: 12). Finally, the shortage

of English mother tongue translators in the EU makes it increasingly necessary to educate English L2 translators.

The objective of this study was pedagogical, and we wish to state from the outset that the research design can be faulted in several ways. Notably, students' activity between the two data collection phases should have been monitored more carefully (see section 3.3). Our study should therefore be regarded as a pilot which can provide preliminary insights and serve as a basis for an improved experimental procedure in a larger study. Nevertheless, despite the flaws in the methodology, there were clear indications that an interdisciplinary approach can be effective in translator training.

### *2.1 Research questions and methods of analysis*

The following research questions were addressed and are listed together with the methods of analysis employed in each case. For all three questions, we compared the results for the two phases.

1. What are the differences between task times in the three translation modalities (written, sight, SR<sup>2</sup>)? Translog Audio (Jakobsen and Schou 1999<sup>3</sup>) was used: (a) to record oral and written translation output; (b) to investigate transient versions of oral and written translations; (c) to time the activities.
2. Is there any difference in translation quality in the three modalities? Three experienced teachers/translators were asked to assess the translation quality of the products (see section 2.2.).
3. How many and what type of misrecognitions occur when students sight translate with SR? Phonetic analyses were performed to identify and categorise the misrecognitions.

In addition to addressing these issues, which were all measured quantitatively, we conducted retrospective interviews with the participants after the completion of the second experiment in order to gather their impressions of working with an SR system.

## 2.2 Procedure

As stated above, the participants were divided into two groups, an experimental group (N=7), who were allowed to borrow the SR software and work with it at home, and a control group (N=7), who did not use the equipment in the interim period. It should be pointed out that, initially, all 14 students trained the SR program as specified by the system. (For a discussion of SR training requirements, see Zong and Seligman 2005: 215-219.) The students in the experimental group were expected to further train the SR system at home and use the software for their translation course assignments and other text production tasks in the interim period, but unfortunately did so in a very limited manner. Consequently, the two groups are very similar in most respects. Therefore the results have been conflated in all the tables below except those where there was a difference between the groups.

Six different text excerpts were selected from the same Danish text (the chairman's statement at the annual general meeting of a bank), each consisting of approximately 110 words. All students translated three texts in the phase 1 experiment and three different texts in the second (total:  $14 \times 3 \times 2 = 84$ ). Every effort was made to ensure that any process/product differences that might emerge across the translation tasks were caused by the translation modality and not by the level of difficulty of the text. To achieve this, translation tasks were rotated in such a way that four translators produced a written translation of Text A, a sight translation of Text B and an SR translation of Text C; five produced a written translation of Text B, a sight translation of Text C and an SR translation of Text A; and five produced a written translation of Text C, a sight translation of Text A and an SR translation of Text B.



Three raters – all experienced translators/teachers – were asked to award the translations a score on a scale from 1 to 5, 1 being the poorest and 5 the highest quality. The inter-rater agreement was high, though one evaluator gave somewhat higher scores than the two others. The oral translations without SR were transcribed and the transcriptions served as a basis for the evaluations. The transcribers (research assistants) were instructed to ignore temporary solutions and only to write out the final version; punctuation was also added. Thus the raters could not know the modality in which the text had been produced. The task times in Table 2 refer solely to the time it took to sight translate and not to the subsequent transcription process.

### **3. Analyses**

#### *3.1 Results for phase 1*

The results for phase 1 of the experiment have been discussed in Dragsted et al. (2011). The conclusions are summarised below. (Details can be seen in section 3.2, where the results for phase 1 have been compared with the findings for phase 2.)

1. **Task times.** Sight fastest, written slowest. SR takes an intermediate position, but is closer to written than to sight (Table 2).
2. **Quality.** Highest in written translation (mean: 3.2), somewhat lower in SR (mean: 2.8) and sight (mean: 2.7) (Table 2).
3. **Misrecognitions in SR translations.** These are caused by (a) hesitations/word boundary problems, (b) homophones, (c) students' incorrect pronunciations, and (d) misrecognitions by SR system (cf. Derwing, Munro and Carbonaro 2000: 599). The different types are exemplified in Table 1.

Table 1: Example of misrecognitions identified for one student (S5)

Error types	Intended	Misrecognised as
<b>Word boundaries or hesitations</b>	...offer [ɒfə:] capital ...a [ə] credit stimulus ...downward spiral [daʊnwərd#s]	...offer a capital ...credit stimulus ...downwards by real
<b>Homophones</b>	...i.e.	...IE
<b>Incorrect Pronunciations</b>	...aid [eɪd] ...of [ɒv] a	...eight ...on a
<b>Inadequacy of SR software</b>	...and thus [dʌs] ...risk of a [ə] ...and [ən] (several)	...and bus ...risk of it ...in (several)

The examples in Table 1 have been extracted from the SR process of one of the students (S5), who produced all error types within a single translation. The first example is a word boundary problem. The participant pronounces final **r** in *offer (capital)*, which is unacceptable for non-rhotic Standard British English, that being the option he chose in the SR program.<sup>4</sup> More importantly perhaps, he prolongs the **r** sound somewhat. This we have categorised as a type of hesitation, since it is a manifestation of his trying to gain time while considering the next word. The lengthening of the sound leads the software into believing that he has said an additional word, namely the indefinite article *a*. Interestingly, there is also evidence that the program may interpret the indefinite article as a hesitation (*a* and *uh* sound identical) and therefore delete it (for example, *a credit stimulus* was registered as *credit stimulus*). The software sometimes finds it difficult to delineate the boundaries between words. S5 says *downward spiral*, but the equipment picks this up as *downwards by real*. (When /p/ is preceded by /s/ in syllable-initial sequences, it is pronounced similarly to /b/, Cruttenden 1994: 46, 140; Collins and Mees 2008: 72). A second error category is constituted by homophones. The participant dictates *i.e.*, but the program does not realise that he intends the abbreviation for “that is” and registers the utterance as *IE*.

Both of the categories we have mentioned above would presumably also be a potential source of error for native speakers though this was not tested. The third category is restricted to non-native speakers, namely incorrect pronunciations. The speaker pronounces the word *aid* with a final /d/ which is slightly too devoiced and which therefore sounds like /t/, and also with a vowel that is slightly too short. In English, vowels are shortened before voiceless consonants (e.g. /t/) and have full length before voiced consonants (e.g. /d/) (Cruttenden 1994: 141), and consequently the program hears his pronunciation of *aid* as *eight*. (The student subsequently rectifies this, after which the word is identified correctly.) The word *of* is pronounced with a full vowel rather than a reduced vowel (that is, as a strong form rather than a weak form, Wells 2008: 891) and as a result the software misinterprets it as *on* (which has no weak form). In addition to these three categories, there are a number of misrecognitions for which we cannot easily find an explanation, and which we have therefore categorised as “inadequacy of the SR software”. When the participant says *and thus*, it is identified as *and bus*, *risk of a* is interpreted as *risk of it*, while *and* is perceived as *in*.

Our findings in phase 1 led to the following hypothesis: *With more practice, SR task times will approach those of sight translation, and SR quality will approach that of written translation. In addition, it is expected that the number of misrecognitions will decrease.*

### 3.2 Comparison of results for phases 1 and 2 (March 2010 vs. December 2010)

3.2.1 Time and quality (Research questions 1 and 2). Table 2 shows means of the task times and quality ratings.

Table 2: Time and quality (means of 14 students). Phases 1 and 2 compared.

<b>Phase 1</b>	<b>Sight translation</b>	<b>SR translation</b>	<b>Written translation</b>
Task time (min:sec)	03:44	08:28	11:07
Quality rating (from 1-5)	2.7	2.8	3.2

<b>Phase 2</b>	<b>Sight translation</b>	<b>SR translation</b>	<b>Written translation</b>
Task time (min:sec)	03:40	11:02	11:55
Quality rating (from 1-5)	2.5	2.8	2.7

In phase 2, sight translation remains the quickest. The mean task times for the written translation are also more or less unchanged. However, our prediction that SR translation would become faster did not come out, since the SR translation in phase 2 on average took longer than that in phase 1; in fact, it took almost as long as the written translation.

The quality ratings for written translation and SR became more similar, as predicted, although it should be noted that the convergence was owing to written translation ratings becoming lower rather than those for SR translation becoming higher. The quality of the sight translation output was still the poorest. As the sight translations were transcribed by research assistants after the experiment, these texts may have been more correct with respect to spelling and punctuation than the translations produced by the same students in the other modalities. The quality ratings may therefore have been slightly better than would otherwise have been the case. However, as the sight translations received the lowest scores, this factor would only have further increased the differences between the modalities.

There was remarkable inter-rater agreement: in phase 1, all three evaluators rated the written translations higher than those produced with SR. In the second phase, two assessed the SR translations as being better than the written translations; one thought they were equally good.

Figures 1 and 2 show the task time patterns for the *individual* participants. (S1, S5, S6, S8, S10, S11, S12 were the students who had used the program at home.) It can be seen that, in phase 1, only two students (S2 and S11) were slower at producing the SR translations than the written translations. Contrary to our predictions, this increased to six of the 14 translators (S2, S3, S7, S8, S10, S11) in phase 2. On closer examination of the data, it appeared that particularly one of the students (S10) in the experimental group was responsible for a fair amount of the additional time required. Learning to use the commands in her case resulted in her spending much time correcting a few isolated mistakes. If this participant were to be excluded, the mean task times would then become: Written 12:04; SR 10:02; Sight 3:42.

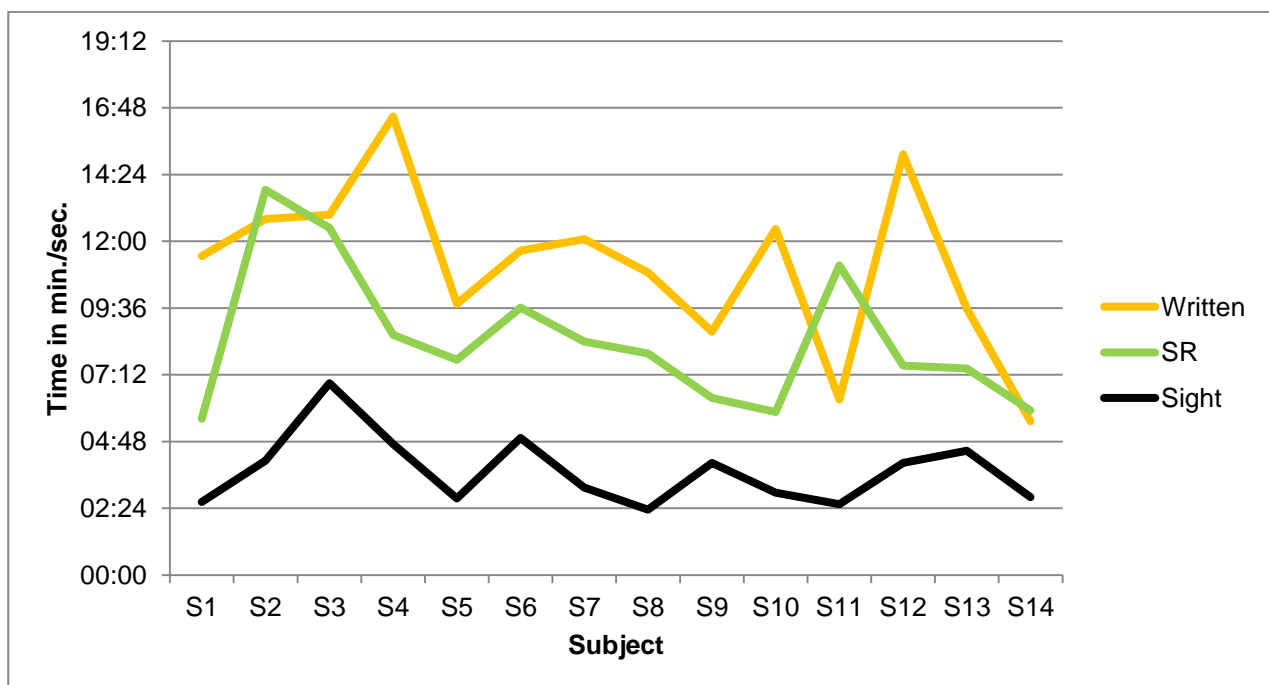


Figure 1: Individual task times in phase 1

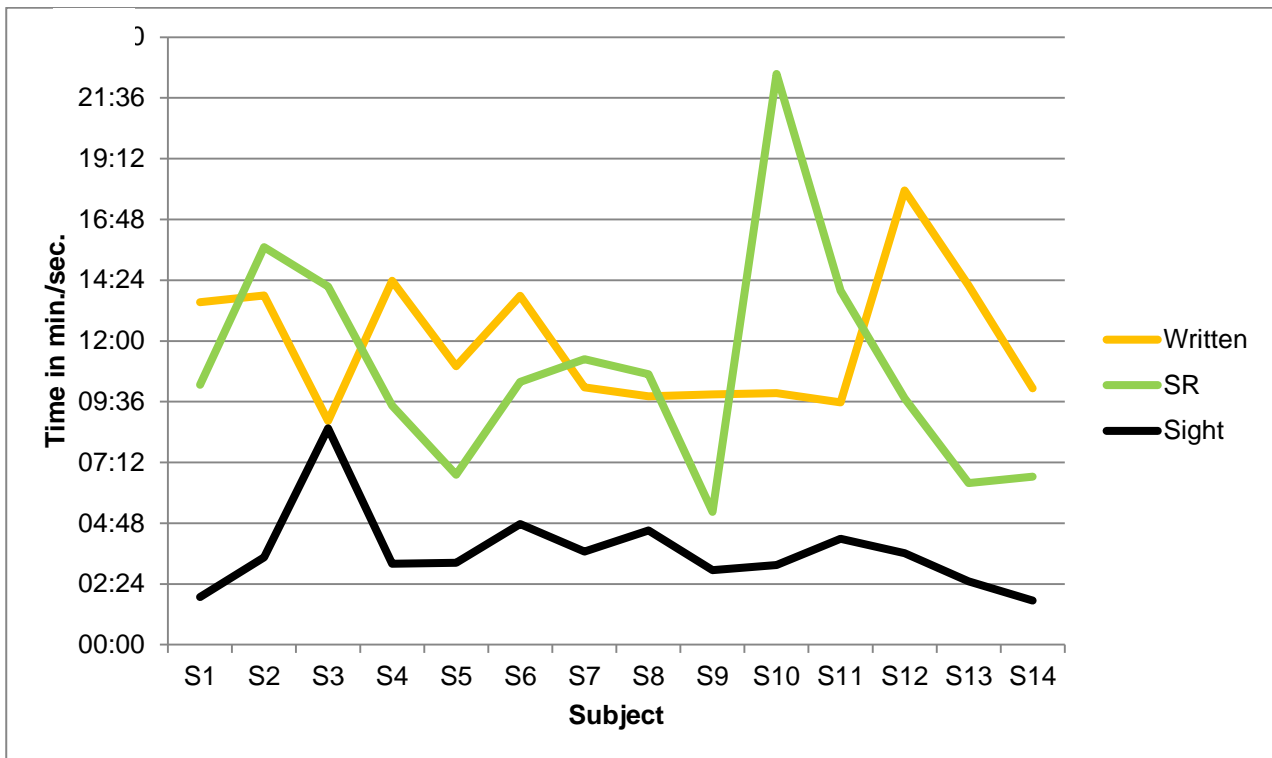


Figure 2: Individual task times in phase 2

Figures 3 and 4 show the ratings of the translation quality for the *individual* participants in the two phases. In phase 1, three out of 14 students (S1, S6, S7) produce better SR translations than written translations, while in the second round this holds true for six participants (S2, S4, S5, S7, S11, S12).

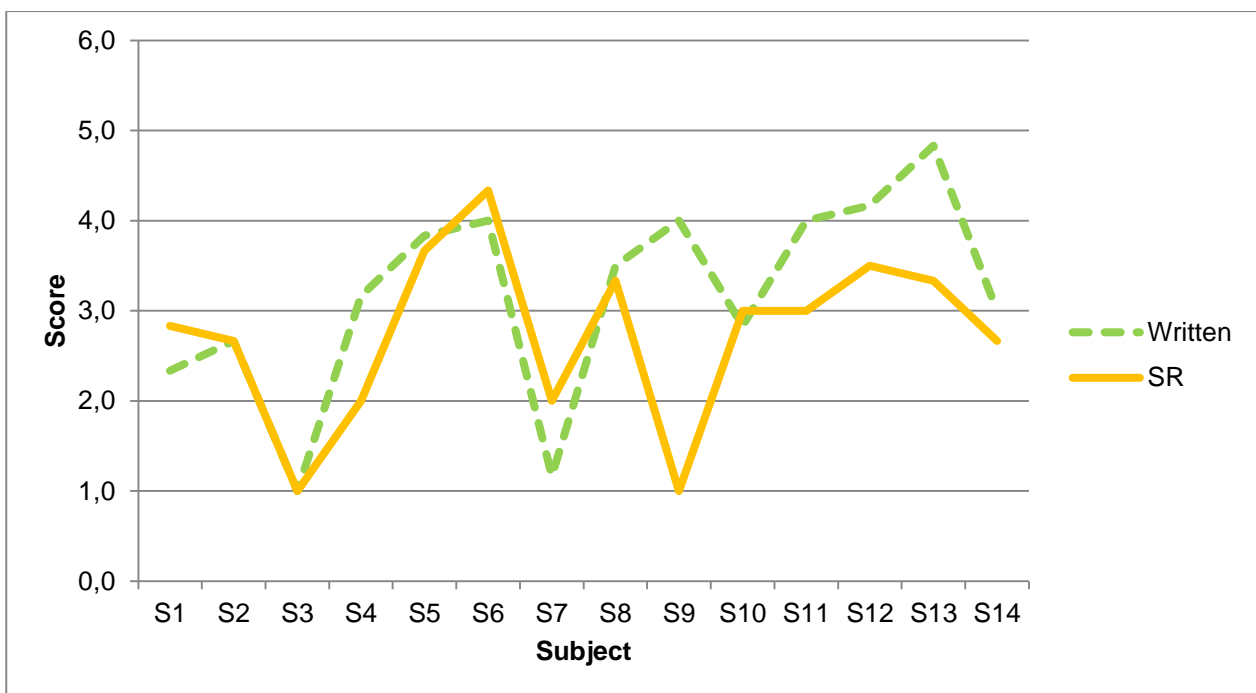


Figure 3: Individual quality ratings in phase 1 (written vs. SR)

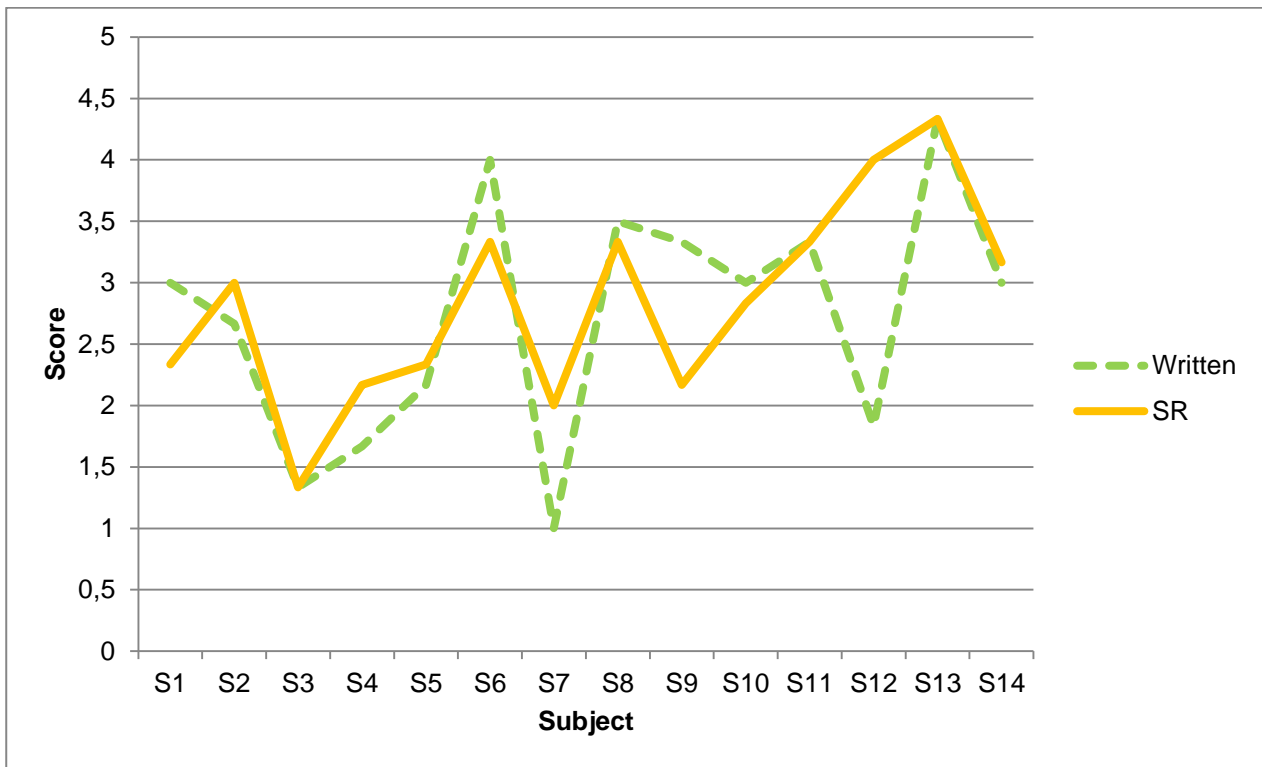


Figure 4: Individual quality ratings in phase 2 (written vs. SR)

3.2.2 *Number and types of misrecognitions (Research question 3)*. The number of misrecognitions in both the experimental and control groups are represented in Table 3. The first phase exhibited 173 misrecognitions. The number of errors for the individual participants ranged from 4 to 32. If one looks at the experimental and control groups, it will be noted that the number of misrecognitions is higher in the control group. If one then considers phase 2, it can be seen that the total number of misrecognitions has increased, from 173 to 248. The rise in errors is notable in the experimental group while the figures for the two phases remain stable in the control group. This comes as a surprise as we thought experience with the program would result in a decrease in the number of errors. For a possible explanation, we refer back to the longer task times found particularly for those students who had taken the program home and learnt to use the oral commands, albeit ineffectively, and therefore spent more time attempting to correct errors than did the control group. Presumably this reflects a stage in the students' development, and we predict that once they have become more familiar with the software, or the software has adapted to their way of speaking, the number of errors will be substantially reduced.

Table 3: Number of misrecognitions in the experimental and control groups. Phases 1 and 2 compared.

Phase 1	Misrecognitions	Range
Experimental (N=7)	75	4-18
Control (N=7)	98	4-32
All students (N=14)	173	4-32

Phase 2	Misrecognitions	Range
Experimental (N=7)	152	10-39
Control (N=7)	96	2-30



All students (N=14)	248	2-39
---------------------	-----	------

Table 4: Types of error (means of 14 students). Phases 1 and 2 compared.

	Homophones	Hesitations	Mispron.	Dragon	Total
Phase 1	8%	24%	54%	14%	100%
Phase 2	6%	14%	38%	42%	100%

Looking at the distribution of the recognition problems (Table 4), it will be seen that the percentage of homophones remains more or less the same across time, the percentage of hesitations decreases, and so does the percentage of mispronunciations. Therefore, relatively speaking, there is an increase in the number of errors which can be attributed to the SR system in the second phase. The lower percentage of hesitations and word boundary problems could be caused by two factors. Firstly, that the students have been exposed to far more interpreting training, which has enhanced their oral competence and maybe also reduced the number of hesitations. A second possible reason is that even after very limited training with the program, they may simply have become better at dictating to the SR program and now realise how to avoid misrecognitions resulting from hesitations. The same may be true with regard to the reduction in the number of mispronunciations. Our retrospective interviews (section 3.4) reveal that a number of students feel that they have become more aware of their pronunciation problems in the course of training the SR program. It is difficult to argue with a machine, and perhaps it is therefore easier to accept that one's pronunciation is inadequate in certain respects when the immediate consequence of an error is flashed back on a screen rather than when it is pointed out by a teacher.

### *3.3 Quantitative conclusions*

We can draw the following quantitative conclusions. In the first phase, there was a small time saving in translation with SR as compared with written translation. This does not increase after the training period. On the contrary: our prediction of a reduction of task times in this modality was not confirmed. First, this is because the students who were given the program used it only very sparingly. In hindsight, we may have underestimated the follow-up and control procedures necessary to ensure that the students actually used the program to the extent we had prescribed. Most likely, more time, effort and guidance is required before improvements kick in and differences between the two groups can be registered. We feel convinced that productivity will improve with experience (Dragsted, Hansen and Sørensen 2009 : 304), and that our measurements were collected at too early a stage in their development, at a point when their limited interaction with the program appears to slow them down rather than speed them up. Secondly, the output quality of the SR translations was more similar to that of written translation; in fact, in certain cases, it was even better. The third conclusion was that the same four types of misrecognitions recurred in the second round of experiments but that the distribution was somewhat different, with relatively fewer hesitations and mispronunciations in the second phase. We could observe a higher percentage of misrecognitions in the experimental group as compared with the control group. As we have seen earlier, this is probably owing to the students in the experimental group now practising the use of oral commands, and therefore struggling to correct errors in the second phase. Note, however, that there was great inter-individual variation in both groups.

### *3.4 Retrospective interviews*

As stated above, we also conducted retrospective interviews. The students were asked what they thought of translating with SR versus the written modality. Some of their comments are quoted below. Recurring statements indicate that they consider that: one obtains a better overall picture; one has to plan ahead what to say and this results in better quality; one doesn't waste time searching for the perfect solution; one doesn't translate word for word; one doesn't look up and check everything. In other words, students appear to feel that a good strategy is to translate the message first and then work on honing the text later. When asked about the advantages, they said: it saved time; the process was quicker; it removed the “stop go” effect; it was useful to see the text in print immediately; and that the first impulse was registered (cf. Chafe and Danielewicz 1987: 88). They also felt that both their spontaneous speech skills and their pronunciation had improved. When asked about the difficulties, the students stated that function words and numbers caused problems and also that the process was tiring. In addition, they thought certain texts to be more appropriate for the use of speech recognition than others.

In due course the students will become more aware of how to avoid misrecognitions and, with training, the system will adapt to the users' pronunciation and idiosyncracies. Though not unimportant, remedying these formal errors is much less of an issue than students' misrepresentations of the content. It would therefore be fruitful to undertake qualitative analyses of both the process and the product using the principles established by Göpferich (2010). By examining the quality of the solutions and the path to the solutions (i.e. the translation strategies adopted) in the different modalities, we will be able to see if SR translations do indeed result in students translating larger meaningful units rather than individual words.

### *Advantages of using SR*

- ❑ “Saves time, the process is quicker and different, often better than written (which is characterised by “stop go” + looking up too many words – often totally unnecessary).” (S9)
- ❑ “Quickly converted to print, less editing, think and plan more.” (S6)
- ❑ “Think more about contexts, more flow in speech, less formal = good for some text types.” (S5)
- ❑ “First impulse registered immediately, can instantly hear if this is adequate or not; useful to hear own pronunciation, one’s spontaneous speech improves (exploit first impulse), have corrected my pronunciation of the words SR program doesn’t understand” (S6)

### *Difficulties using SR*

- ❑ “Difficulties with function words, but it helps to take longer units in one go.” (S12)
- ❑ “Difficulties with numbers, but more problems in first experiment than second.” (S2)
- ❑ “I grew tired, requires high concentration to translate orally. Not enough confidence (yet), not good for legal texts, fine for other text types (e.g. mails)”. (S11)

## **4. Discussion and conclusion**

We believe that our study has shown that SR translation is a possible alternative or supplement to other translation modalities in translator training. It would appear to be a powerful pedagogical tool to help translation students think in larger chunks and take a more panoramic view. In our opinion, this is the reason – together with the additional translation and interpreting training they had received in the intervening months – for the convergence of the quality ratings of written and SR translation in the second round. A prominent feature in interpreting and other oral communication courses is to teach students how to plan their output and “think before they speak”. The benefits of using SR could therefore also rub off on written translation if it teaches the students to adopt the

motto "Think before you write". It would appear that it helps them trust their first intuition rather than pore over isolated words and spend ages trying to solve a problem.

Students say that using the software raises their awareness of potential pronunciation errors and that speaking their translation encourages them to deal with larger units, and thus translate overall meaning instead of individual words. If universities ensure that subject areas such as oral proficiency, phonetics, translation and interpreting are integrated into courses more meaningfully, students will see to a larger degree the relevance of the individual subjects, and discover that, to coin a phrase: "the whole is greater than the sum of the parts". An interdisciplinary approach will help Emma – our student who couldn't see the relevance of all the courses in her degree programme – appreciate the different components of the curriculum. So, rather than offering phonetics as an isolated subject, it would be preferable to link it to T & I training, since the reasons for needing a proper pronunciation become blatantly obvious once students find that interaction with the software only works if the users' pronunciation reaches a certain level.

Our findings indicate that, with training, SR translation can probably be made faster and can attain a standard which is as high as that of written translation. This means that Peter, our translator with the broken arm, and other translators who are troubled by physical impairments, may have a viable alternative to typing. But any translator may enjoy the mental and ergonomic variation of supplementing traditional keyboard activity with using SR.

Not only translators, but anybody who produces text in a foreign language, may benefit from SR technology. For those who, like David, are dyslexic it is an obvious alternative, but more generally, anybody who finds it challenging to deal with *written* language, notably in a foreign language, may

find it an advantage to make a quick first draft with an SR system (Leijten 2010: 965). Having problems with writing or spelling does not imply that one necessarily has problems with speaking and pronouncing.

We have seen that some people seem to be better at using SR software, and translating orally, than others, and, as can be seen from Figures 5 and 6, there would appear to be a relationship between the quality scores in sight and SR translation. This has to do with a variety of factors – for instance their pronunciation, their ability to dictate larger chunks and avoid hesitations, and their willingness to give up the working routines they have become accustomed to in written translation.



Figure 5: Individual quality ratings in phase 1 (sight vs. SR)



Figure 6: Individual quality ratings in phase 2 (sight vs. SR)

The readiness to embrace new technologies is another important prerequisite for succeeding with the program. Translators who have been practising for many years will have to invest considerable time and energy in order to change fundamentally their working habits and take on new tools. Universities have an important role to play in introducing SR in addition to other technologies as part of the translator’s toolbox. For too long universities like our own have engaged in traditional approaches to translator training, and have not addressed the changing preferences of students when obtaining and communicating information. Universities have tended to ignore the fact that young people read less than former generations, and that they are simply no longer attracted by black print on white paper. They are multimedial and expect classes in translation and other subjects to be more interactive and to incorporate media dealing with text, audio, animation, video (Hansen and Shlesinger 2007: 111). With the recent development of voice activation in smart phones and computers there is a golden opportunity to capture students’ interest. As speech becomes a more reliable and popular input method, using SR may grow into a more natural working mode also in translation. In sum, curricula have to be made more interdisciplinary to prepare students for life as

translators, where they have to be able to use a wide array of technological solutions that may enhance their performance.

Altogether using SR software can be seen both as a tool for translation research, teaching and practice. We hope to have shown the sound effects of using it in translation.



## References

- Agrifoglio, M. 2004. "Sight translation and interpreting: A comparative analysis of constraints and failures". *Interpreting* 6:1. 43–67.
- Chafe, W. and J. Danielewicz. 1987. "Properties of spoken and written language". R. Horowitz and S.J. Samuels, eds. *Comprehending oral and written language*. San Diego: Academic Press. 83–113.
- Collins, B. and I.M. Mees. 2008. *Practical phonetics and phonology*. 2nd edn. Abingdon: Routledge.
- Cruttenden, A. 1994. *Gimson's pronunciation of English*. Fifth edn. London: Arnold.
- Derwing, T.M., Munro, M., and Carbonaro, M. 2000. "Does popular speech recognition software work with ESL speech?" *TESOL Quarterly* 34:3. 592–603.
- Dragsted, B., I.M. Mees and I.G. Hansen. 2011. "Speaking your translation: students' first encounter with speech recognition technology". *Translation & Interpreting* 3:1. 10–43.
- Dragsted, B., I.G. Hansen and H.S. Sørensen. 2009. "Experts exposed". I.M. Mees, F. Alves and S. Göpferich, eds. *Methodology, technology and innovation in translation process research*. (Copenhagen Studies in Language 38) Copenhagen: Samfundslitteratur. 293–317.
- Gile, D. 1995. *Basic concepts and models for interpreter and translator training*. Amsterdam/Philadelphia: John Benjamins.
- Gile, D. 2004. "Translation research versus interpreting research: kinship, differences and prospects for partnership. C. Schäffner, ed. *Translation research and interpreting research: Traditions, gaps and synergies*. Clevedon: Multilingual Matters. 10–34.
- Göpferich, S. 2010. "The translation of instructive texts from a cognitive perspective: Novices and professionals compared". Göpferich, S., F. Alves and I.M. Mees, eds. *New approaches in*

- translation process research*. (Copenhagen Studies in Language 39) Copenhagen: Samfundslitteratur. 5-55.
- Hansen, I. G. and M. Shlesinger. 2007. “The silver lining: technology and self-study in the interpreting classroom”. *Interpreting* 9:1. 95–118.
- Jakobsen, A. L., and L. Schou. 1999. “Translog documentation”. G. Hansen, ed. *Probing the process in translation: Methods and Results*. (Copenhagen Studies in Language 24) Copenhagen: Samfundslitteratur. 151–186.
- Jurafsky, D. and J.H. Martin. 2000. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. New Jersey: Prentice Hall.
- Lambert, S. 2004. “Shared attention during sight translation, sight interpretation and simultaneous interpretation”. *Meta* 49:2. 294–306.
- Leijten, M., Janssen, D., & Van Waes, L. (2010) “Error correction strategies of professional speech recognition users: Three profiles. *Computers in Human Behavior* 26. 964–975.
- Leijten, M. & Van Waes, L. (2005) “Writing with speech recognition: The adaptation process of professional writers with and without dictating experience”. *Interacting with Computers* 17. 736–772.
- Luyckx, B., Delbeke, T., Van Waes, L., Leijten, M., & Remael, A. 2010. “Live subtitling with speech recognition: Causes and consequences of text reduction. Antwerp: *Artesis Working Papers in Translation Studies* 2010-1.
- Phillipson, R. 2003. *English-only Europe: Challenging language policy?* London: Routledge.
- Wells, J.C. 2008. *Longman dictionary of pronunciation*. 3rd edn. Harlow: Pearson Education.

Post print of *Target* 25:1 2013. pp. 140–154, DOI [10.1075/target.25.1.11mee](https://doi.org/10.1075/target.25.1.11mee)  
This article is under copyright and the publisher should be contacted for permission to re-use or reprint the material in any form.

Zong, C., and M. Seligman, M. 2005. “Toward practical spoken language translation”. *Machine Translation* 19:2. 113–137. Electronic Edition. [http://nlpr-web.ia.ac.cn/cip/ZongPublications/2006.06.23%20Machine%20Translation%20Journal\\_OnlinePDF.pdf](http://nlpr-web.ia.ac.cn/cip/ZongPublications/2006.06.23%20Machine%20Translation%20Journal_OnlinePDF.pdf) , 20 January 2012

Post print of *Target* 25:1 2013. pp. 140–154, DOI [10.1075/target.25.1.11mee](https://doi.org/10.1075/target.25.1.11mee)  
This article is under copyright and the publisher should be contacted for permission to re-use or reprint the material in any form.

## *Notes*

---

<sup>1</sup> Cf. Leijten et al. (2010: 969)

<sup>2</sup> *Dragon Naturally Speaking Preferred 10* (Nuance Communications, Inc.).

<sup>3</sup> <http://www.translog.dk>

<sup>4</sup> When creating a profile in the SR program, the user has the option to choose between different varieties of English.