

Proficiency Effects and Compensation in Advanced Second-Language Reading

Philip Shaw and Alan McMillion, Stockholm University

1. Introduction

Many, if not most, university courses within the Scandinavian countries use English language textbooks and articles to a greater or lesser extent, although lectures and seminar discussions may be in the L1 (Swedish, Danish, etc.). Scandinavian students generally use the same textbooks as their L1 English counterparts in Britain/US. Students in this position are especially interesting given that the number of L2 users of English in this position is steadily increasing around the world (Graddol 2006).

It is often taken for granted that because Scandinavian students' spoken production and listening comprehension is genuinely advanced, then their reception of educational texts is likely to be at a level at least as advanced. This assumption is poorly founded (Cummins 1984), and their receptive proficiency is worth investigating for this reason alone. But rather less research has been carried out on such advanced users compared to research on beginner and intermediate levels of L2 readers (the focus often being on L2 acquisition/learning rather than L2 use) and there are several reasons why this group of L2 users is particularly interesting. For example: it is unclear what processing differences (as opposed to merely slower processing) may exist between L1 and advanced L2 groups, and insights here might inform us about potentially useful pedagogical intervention; and it is unclear what allows some intermediate readers to become advanced and what allows some advanced readers to become native-like (in reception, production, or both), while others, regardless of the length of time or intensity of their L2 use, do not move beyond the advanced-but-not-native-like level.

This paper therefore focuses on the reading and comprehension skills of advanced L2 users, paying attention to both the quality of the comprehension 'product' (the mental model of text content constructed) and speed of processing. More specifically, it focuses on first-year biology students at Stockholm University and equivalent students at Reading and Edinburgh Universities in the UK. This is part of a larger

study on the “advanced L2 user”¹ Although it may not be immediately clear who or what this label refers to, for current purposes we assume that students of biology at Stockholm University, most of whom have had 9-10 years of English at primary, secondary and upper secondary school, are representative of the group. That is, advanced readers are L2 users of English who have reached a level of proficiency adequate for reading advanced textbooks in English.

For university students a large proportion of time is spent reading textbooks and related material, so it seems plausible to assume that an important factor in whether they succeed in their studies using L2 English is their level of reading proficiency. (It might seem to be the most important factor, but evidence suggests that Swedish students do not in fact read their textbooks all that much; see McMillion and Shaw forthcoming.) We do not necessarily know much about their reading proficiency. Although the Swedish National Tests for entering university studies include tests of reading comprehension of both Swedish and English texts, these texts are not always of the same level and type as those encountered in university programs. Moreover, it is not clear whether high scores on such tests correlate with learning and recall or whether they only provide an index to relatively rapid text modeling under pressure (the test situation). In other words, whatever it is that the Swedish National tests are measuring, it may not correlate strongly with recall measures. Reading for learning and long-term recall, which presumably is the kind of reading students engage in when studying textbooks, may be different from the kind of reading that is effective in answering comprehension questions immediately after reading an otherwise irrelevant text (see e.g. Carver 1990). Thus, any answer to the question of whether Swedish university students comprehend textbook English at a level comparable to their L1 counterparts will depend, at least in part, on how reading comprehension is operationalized.

In this context the relevant general questions concerning advanced L2 reading proficiency would include the following:

¹ This project is part of the ‘AAA’ (Advanced Second Language Users) programme at Stockholm University. We would like to thank the Bank of Sweden Tercentenary Fund for funding this project.

- How well does the overall reading comprehension of advanced L2 users compare with that of L1 readers that have comparable education and social backgrounds?
- What are the product and processing differences between advanced L2 readers and L1 readers? In particular, are there differences other than mere speed of processing?
- Is there a difference between comprehension as measured by standard comprehension tests (read a text and then answer questions about the text, under time pressure) and comprehension which implies more robust modeling of the text so that recall and learning are enhanced?

With these questions as a point of departure, this paper will discuss one of several studies carried out at Stockholm University. This study focuses on differences between advanced L2 English readers and L1 readers of roughly the same educational background, on both product and processing tasks.

2. Background

Success in reading in a foreign language has been ascribed to a combination of three groups of factors:

1. appropriate background knowledge relating to the text topic, genre, cultural assumptions (e.g. Steffensen, Joag-Dev, and Anderson 1979) etc.
2. effective reading skills, strategies, attitudes, and expectations (van den Broek et al. 1995) for reading in the given language.
3. proficiency in the syntax, lexis, etc. of the language being read (Alderson 1984, Bernhardt 2005, van Gelderen et al. 2004)

In the present study we compare a group of Swedish university students with an equivalent British group. Both have to read expository English-language material without strong cultural content which seems likely to be equally accessible in either culture. Since some of the texts presuppose some general knowledge, we expect that background knowledge will be a factor differentiating within the groups, but not necessarily between them. We have no reason to believe that either group is better-informed on general topics, that is, but we do suppose that individuals in both groups will vary considerably from one another on

this parameter. Nor have major differences in text structure been reported between Swedish and English expository text (cf Fløttum et al. 2006, who find Norwegian academic prose to be more like English than French).

The second dimension of difference – reading skills, strategies, attitudes, and expectations – is what Fukkink, Hulstijn, and Simis (2005) refer to as metacognitive knowledge. Metacognitive knowledge is regarded as representing language-independent, higher-order processing skills that are transferable (normally with positive results) from L1 to L2 reading. There is, however, evidence that these higher-order processing skills are not necessarily all language-independent (Koda 2004), with some being only shared across related languages, so that there will be some language-related differences at this level between L1 and L2 readers. However Fukkink et al. found a high degree of transfer of metacognitive knowledge from Dutch-language to English-language reading, and it is likely that there are similar transfers for the English-Swedish language pair.

This implies that in our case the relevant differences between the L1 and advanced L2 readers are primarily linguistic. There are several levels of language knowledge where differences between L1 and L2 users can be located. Clearly, L2 readers' knowledge of L2 lexis, syntax, characters, etc. may be incomplete and lead to miscomprehension. But Perfetti's (1999) model of reading comprehension reminds us that the other dimension of difference is processing. Even when L2 knowledge is fairly complete, processing may be slower or qualitatively different. In short, differences between L2 and L1 readers with similar cultural/textual/attitudinal backgrounds may be due to the influence of automatized procedures transferred from L1 (which may need to be inhibited), to limited linguistic knowledge, or to L2 processing being less automatized.

Bernhardt (2005) proposes a compensatory perspective on L2 proficiency, claiming that knowledge sources (particularly L1 literacy and L2 knowledge) are not strictly additive but interact synergistically: "the higher the L1 literacy level, the more it is available to buttress impoverished second language processes, ... the more word knowledge is developed, the more it frees up resources to operate on more complex syntactic patterns, and so forth" (2005:140). The construct L1 literacy seems to be similar to metacognitive knowledge – a set of habits,

behaviors and strategies for processing texts. In addition to generative processing behaviors, L1 literacy must include strategies for dealing with the various kinds of obstacles and problems that occur in text processing. In other words, L1 literacy includes behaviors and strategies for dealing with problems conditioned by inefficient processing and lack of textual coherence cues. What is transferred from L1 reading to L2 reading, then, is not only those processes that contribute to reading efficiency but also those behaviors and strategies used to handle processing difficulties.

Although Bernhardt (2005) discusses the transfer of literacy skills, such as character recognition, vocabulary recognition and access, text structure, and expectations concerning text structure, she does not analyze them into component processes (2005:140). However, it need not be a unitary L1 literacy that compensates for weak L2 knowledge, but a set of skills or processes, some highly automatized, some less so or even explicit, declarative, and conscious, that are intricately interwoven. It can be assumed that these occur at different levels of the reading process. For example, character recognition is certainly a low-level skill where automatized recognition processes are easily transferred from the reading of L1 to the reading of L2 texts. For cognate languages such as Swedish and English, it does not constitute an obstacle for the L2 reader (although Swedish does have three letters, å, ä and ö which English lacks). Hence, concerning initial learning of an L2 Bernhardt says “[...] there is already some literacy knowledge on the part of all readers especially from cognate languages” (2005:139). Vocabulary recognition and lexical semantic access are processes that could present L2 readers with obstacles to smooth reading, either because their L2 vocabularies do not include many of the words in a text or that their meanings are not well entrenched or too vague with insufficient supporting context. But this can also be a problem for weaker L1 readers.

Several researchers have found that L2 readers are measurably slower than their L1 counterparts. Segalowitz, Poulsen and Komoda (1991) found that advanced French-English bilinguals read considerably slower in L2 than in L1, and that slower overall reading in L2 seems to be a general phenomenon. Fraser (2007:387) points out that there is no necessary link between proficiency and processing speed. Processing speed seems in fact to lag behind proficiency. L2 users’ processing on various levels is usually slower than L1 users’ and this remains true even for quite advanced L2 users.

The slower whole-skill rates usually found for L2 readers are often regarded as implying that L1 users have more automatized decoding that requires less attention and thus leaves more capacity free for higher cognitive processing (Geva 1993). Slower whole-skill reading is certainly at least partly due to slower low-level processing. It is a very robust finding of word recognition studies that L2 users process individual uncontextualized words more slowly.

Readers obviously differ in working memory span. Individual working memory span is important for all reading (Haarmann, Davelaar and Usher 2003, Abu-Rabia and Siegel 2002), since it facilitates the cumulative development of the textbase and situation model. If L2 users have less automatized processing, demands on memory are higher, so a qualitative difference between L1 and L2 readers might be that L2 readers experience more demands on memory.

Walczyk's (2000) model of compensatory processing among weaker and stronger L1 readers is very revealing. Walczyk claims that the interplay between automatized and control processes (his term) in reading is intricate and variable, but given the range of individual variability in many sub-skill processes involved in reading, readers use a variety of strategies and behaviors to compensate for inefficient automatized processing. His model predicts that the range of reading product among L1 readers of similar educational background will be fairly narrow, but, because of less efficient processing, weak readers will need more time to reach this level of comprehension product. Similar ultimate levels of accuracy will be achieved at different speeds. Walczyk points out (2000:563) that in consequence the magnitude of correlations between processing efficiency and comprehension product will be greatest when reading occurs in situations where there are severe time constraints (e.g. test situations). That is, where time is limited so that only the fastest processors complete the task, the difference between weak and strong readers will show up as a difference of product, but where reading is self-timed the same difference will show up in speed.

Walczyk's model has been applied to L1 readers and it seems reasonable to investigate the hypothesis that many (though not necessarily all) aspects of advanced L2 reading will be similar or identical to corresponding L1 aspects. We may hypothesise that causes of lower levels of automaticity among the L2 readers will be similar to the causes of lower levels of automaticity among weaker L1 readers, e.g.

shallow vocabulary knowledge or small vocabulary size, non-automatized collocational processing, narrow working memory span, inefficient syntactic parsing, etc. However, unlike the weak L1 readers, many of the L2 readers are efficient L1 readers and have functional compensatory behaviors and strategies from their L1 processing that can be used for dealing with various types of reading difficulties in the L2. Consequently, we might hypothesise different patterns due to the availability of L1 metaknowledge/literacy/strategies.

3. Research questions

Broadly, our primary aim is to compare advanced L2 reading product and processing with that of L1 users of similar educational background and reading purposes. More specifically, in order to investigate whether some tasks call on strategies that are differentially available to L2 readers, we want to compare results on tasks that require long-term storage of text information with tasks that do not make such demands on the reader.

Our starting research questions are the following:

On four different types of test, are there differences between mean scores for reading accuracy between the L1 and advanced L2 groups of English readers?

On two of these types, are there differences between the mean processing speeds of the two groups?

How does the relation between mean speed and mean accuracy vary between L1 and L2 groups across different types of test?

4. Method

We selected first-year biology students at Stockholm University as our target population. Although the majority of these subjects had Swedish as their L1 and as their home language, a minority had other home languages and occasionally other L1 languages. To date we have administered paper tests to a total of 100 Swedish first year biology students. Of these we have further tested 15 using computer-based tests.

Our L1 sample consists of 49 British subjects, first-year students of biology at Reading and Edinburgh Universities. All had had their secondary education in Britain, and nearly all were native speakers of English. All were tested on the same paper tests under the same conditions as the Swedes, and 19 (from Edinburgh) also took computer-based tests.

Our battery of tests consists of a set of paper-based tests and a set of computer-based tests using the program Superlab.² The relevant paper-based tests are the following:

Comprehension test. An often-used comprehension test format consists of short texts followed by a set of multiple-choice questions, given with time limits. Our studies made use of one test with this format in order to have a task type that was comparable with other studies. The texts for this test were approximately 200 words long and take up various academic-type topics (economics, biology, etc.). The texts were intended for the general educated reader; some came from the *Economist* and the *Guardian*, others were of comparable difficulty. They therefore presupposed some background knowledge and possession of this knowledge might be supposed to help some L2 readers to compensate for any language deficiencies.

There were ten texts and for each text there was a set of four multiple-choice questions with five alternatives. Each question set included one of each of the following types: (a) targeting information retrievable from the text, (b) targeting information retrievable from the text but paraphrased, (c) targeting information requiring an elaborative inference, i.e. an inference which adds information not in the text, (d) targeting gist or summary information.

The informants were given 25 minutes to complete the test. The aspect of interest was accuracy, but under time pressure (not unlike the situation many students find themselves in on a regular basis, i.e. a fixed amount of text to read and understand within a limited amount of time).

² Superlab version 4.0.3c, April 23, 2007 release, Copyright Cedrus Corporation 1991-2007, San Pedro, California.

The results of pilot versions correlated well with other text comprehension tests, including the Swedish National Tests.

Recall task. This task was based on a text of about 800 words, taken directly from a biology textbook comparable to the one these subjects were using in their courses, and thus aimed more precisely at them as audience than the texts in the multiple-choice test. It therefore presupposed quite a lot of background knowledge, but it was knowledge that the subjects could be presumed to have more uniformly than that required in the multiple-choice test. The subjects were allowed 10 minutes to read the text and then, without having the text to refer to, were given 10 minutes to write a summary of it. The Swedish subjects were told that they could write their summary in either Swedish or English (see Brantmeier 2006 on answering test questions in one's L2). The task has content validity, in that it is like what students have to do, but it is not a pure reading task and the results will be affected by memory differences, and differences in various productive language skills.

The test was scored by developing a key listing all the propositions in the text (and intersubjectively validating this among three researchers). The subjects' handwritten summaries were then keyed into a file, and the propositional content of each was compared with the key separately by two evaluators. The number of correctly recalled propositions for each summary was tallied, scoring 1 for more central and 0.5 for less central points.

The underlying hypothesis behind this task is that to optimize recall of a rather complex text, in this case concerning the plant hormones gibberellins, a reader would need to create a mental model (or models) of the text. This model would naturally have stronger links for causally linked factors (e.g. that gibberellins caused dwarf plants to produce normal growth) and temporal links (e.g. when various discoveries related to gibberellins were made) and weaker links to information that was of a more incidental nature (that one scientist was Japanese, another American). Although the recall-protocols were time-consuming to evaluate, they provided valuable information concerning the range of abilities of the subjects in text modeling and recall.

In addition to these product tests a number of processing tests were administered using the Superlab program. Among the various tests, the following two are relevant to the current discussion.

Sentence coherence. One sentence at a time was shown on the screen. Each sentence consisted of two clauses that were explicitly related via a subordinating conjunction. The informants were to decide whether the clauses created a coherent meaningful statement. For example, the sentence ‘Fish is becoming expensive because the moon revolves around the earth’ would be incoherent while ‘I need to borrow some money because I left my wallet at home’ would be coherent. The aim was to test the speed of processing when relating two simple clauses.

Scenarios. This test consisted of two sentences being shown on the screen. When the mouse button was pressed these sentences disappeared and a third appeared. The task was to decide whether the third sentence was coherent with the preceding two sentences. Together the sentences build a scenario or simple situation. The aim was to test the model-building skill of the informants. Unlike the single sentence coherence test, this task requires the informant to recall the basic scenario from the two sentences which are no longer on the screen and must be encoded, presumably as some type of situation model (see Zwaan & Radvansky 1998, Van den Broek et al 1995, 2002). In addition to the basic language processing involved, the subject must keep in working memory for a few seconds the situation described in the first two sentences while decoding the third sentence.

In both tests all informants were told that they should answer as fast as possible compatible with accuracy. In fact the results show that subjects generally prioritized in the same way: average accuracy was about the same for both groups.

On the paper tests where subjects achieve varied accuracy in a fixed time: their achievement can be described by one number. On the computer-based ones, however, they work in their own time, and may still be more or less accurate: their achievement is described by a number for time and a number for accuracy.. In order to be able to compare these two types of test we calculate two figures based on the results – one number called ‘efficiency’ and another called ‘relative L2 efficiency ratio’. The first is a measure, specific to each test, of the subjects’ accuracy in relation to time. The second is a measure of how close L2

subjects were to L1 performance in terms of both speed and accuracy on the given test, and it should be comparable across tests.

Efficiency is calculated as accuracy divided by response time. For the reading comprehension the efficiency is the score out of the total possible, 40, since the time available for making a response was the same for all subjects. Similarly, for the recall task the efficiency is the score out of the maximum achieved by any subject since the time was again a fixed quantity. The computer-based tests give us the exact time taken for each question by each subject, so the efficiency here is the ratio of total number of correct choices to average response time in seconds. These figures provide four different and mutually incompatible measures, but which have in common that they measure accuracy per time. They are only meaningful as input to the next figure, the relative L2 efficiency ratio. This is the ratio of the L2 readers' efficiency score to that of the L1s, the product of dividing L2 efficiency by L1 efficiency. This figure can be used to compare the L1 advantage on one test with that on another.

5. Results

Comprehension test. (multiple-choice). In general, this test was experienced as rather difficult by all the subjects (on the basis of their recorded comments afterwards), with only a minority of the subjects finishing all ten texts within the allotted time. Results shown here are based on 100 Swedish subjects. The average overall score was 17.80 for the British subjects (SD: 4.41) and 13.24 (SD: 4.43) for the Swedish subjects, so that the L1 readers appeared to do much better.

The data indicate that different subjects applied different strategies to answering the questions. Many of the British subjects tried to get through all the questions in the allotted time and consequently had only chance level accuracy on many of the later texts. Many of the Swedish subjects and several of the British used the strategy of spending more time on each question, with the general result of high accuracy on the questions answered.

Figure 1 gives the cumulative total scores for Swedish and British subjects across the test. Thus it shows that after one question the average score for both groups was around 2.1, after two questions the average

total score was around 4.3, etc. It can be seen that up to text 4 the two groups had almost identical average scores, but they start to diverge at text 5

251658240

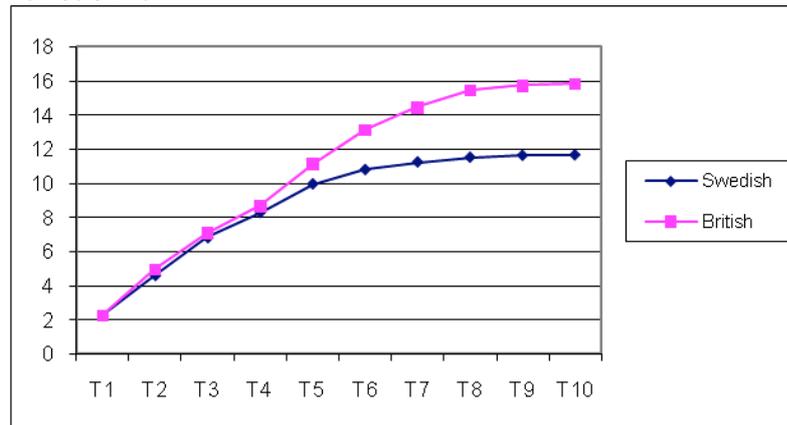


Figure 1. Cumulative total scores, text by text.

This suggests that the main difference between the groups is the time it takes them to produce a result, and not in reading skills as such. The British subjects in general read much faster, answering on average 34 questions, with two-fifths (19 out of 49) answering all, while Swedes on average answered 22.5, and only 2 out of 100 managed all the questions. Since the groups are comparable in terms of product until text 5, the figure suggests that reading under time pressure is the main factor differentiating the groups. Text 5 is in fact the first text which substantial numbers of subjects did not attempt.

It should be noted that different strategies can be adopted to deal with the difficulty of the task under time pressure. One can either do only the questions one has time for, or one can use the last few minutes to go through and answer every question at random, hoping to pick up extra points by chance. Several of the weaker British subjects adopted the second strategy, but apparently none of the Swedes

To assess whether the reading process was comparable between groups, the correlation was calculated between the average scores per question on the first four texts (which most attempted) for the two groups. The resulting coefficient was 0.89; that is, questions which were

relatively easy for one group were relatively easy for the other, and vice-versa. The pattern of correct answers to the questions was, it could be said, similar between groups.

Recall On average the British subjects recalled 13.2 propositions (SD 4.4) and the Swedes 8.5 (SD 2.8). The British subjects thus performed substantially better on this task.

Sentence processing On the previous two paper-based comprehension and recall tests a fixed time was associated with very different overall accuracy scores.³ On this test the position was reversed: the accuracy of the Swedish subjects was nearly as high as that of the British subjects, but their speed was much lower.

Table 1. Results of sentence processing test.

Variable measured	British	Swedish
Mean accuracy (correctly classified sentences/50)	46.13 SD: 2.3	45.27 SD: 3.4
Speed (seconds from presentation to decision)	3.64 SD: 1.0	4.62 SD: 1.7

Nineteen British subjects were compared with 15 Swedish (See Method above). Both groups achieved more or less complete accuracy, but the British subjects were more than a quarter faster than the Swedes.

Scenarios (Situation modeling and text modeling). On this test, too,, the accuracy of the Swedish subjects was nearly as high as that of the British subjects, but their speed was much lower.

³ By looking at the cumulative test scores, as in Figure 1, we can see that the overall score difference on the multiple-choice comprehension test is likely to be a product of the fixed time: in the first questions, where time pressure was least, accuracy scores were not very different between the groups.

Table 2. Results of scenarios test

Variable measured	British	Swedish
Mean accuracy (correctly classified scenarios/32)	30.37 SD: 1.6	29.60 SD: 2.3
Speed 1 (seconds from presentation to request for third sentence)	2.39 SD: 0.7	2.94 SD: 0.9
Speed 2 (seconds from presentation of third sentence to decision)	5.40 SD: 1.9	6.61 SD: 2.4

Again 19 British subjects were compared with 15 Swedish. Both groups achieved more or less complete accuracy but the British subjects were more than a quarter faster than the Swedes.

Ratios for all scores. Table 3 gives the efficiency (average score divided by average time taken) and relative L2 efficiency ratios for the four tests. Note that efficiency and raw score are the same for the paper tests because all subjects had the same time. The relative L2 efficiency ratios for the four different comprehension tasks are very similar. The recall protocol relative L2 efficiency score is, however somewhat lower than the others.

Table 3: Overall results of tasks

Task	Score British	Score Swedish	Efficiency British	Efficiency Swedish	Relative L2 efficiency
Reading comprehension	17.80	13.24			0.74
Recall protocol units	13.24	8.52			0.64
Sentence accuracy	46.13	45.27	12.68	9.80	0.77
Speed	3.64	4.62			
Scenarios accuracy	30.46	29.60			
Speed 1	2.39	2.94	12.74	10.06	0.79
Speed 2	5.40	6.61	5.64	4.48	0.79

It will be noted that the reading comprehension test score used here is the overall one, strongly influenced by the Swedes not having finished the test. This score is used because it represents achievement in a known time; any other measure, such as average number right of those attempted, represents achievement in an unknown time per question,

generally longer for the Swedes than the British subjects, and cannot give us an efficiency measure.

6. Discussion

Although these tests were carried out on convenience samples of volunteers, there is no reason to believe that there are significant differences between the two groups other than their first language and educational culture. Possible influences from educational culture, which merely add to the uncertainty of the results, include test-wiseness and summarizing skill. The British students were probably more used to doing multiple-choice tests and the weaker subjects were more inclined than the Swedes to complete the whole test at random; the average number of questions attempted by British students with a given overall score was higher than that of Swedes with that score, and they often had very few correct answers on the last questions attempted. As for summarizing, it is possible that the Swedish education system puts (even) less stress on this than the British.

The paper tests were carried out on relatively large samples and the computer tests on moderate-sized subsets of those samples. In both cases the members of the computer-test subset were likely to be rather more proficient readers of English than the average of the whole group, because it was the more proficient subjects who tended to volunteer for further tests.

In every test there were Swedish subjects who scored well above the average score of the British subjects, and British subjects who scored well below the average of the Swedish subjects, so that one could well argue that background knowledge and L1 literacy/reading skills/metacognitive knowledge were more important variables than L1/L2 status. Nevertheless, the results show consistent differences between L2 and L1 readers.

To sum up the results, our first and second research questions asked whether there would be differences between mean scores for reading accuracy on our various tests between the L1 and advanced L2 groups and whether there would be speed differences between the groups. We found differences in accuracy when the tests had to be performed in a limited time, but not in self-timed tasks. The results of the multiple-

choice reading comprehension test suggest that Swedish subjects achieve less reading product in a given time than British ones. The results of the parts of the multiple-choice done by all subjects (the first four texts) and the sentence and scenario processing suggest that Swedish subjects achieve the same accuracy of product as the British ones if (1) they are provided with additional time for these tasks, and (2) the texts are relatively short (about 200 words or less). When reading for learning of longer texts, the Swedish subjects did not, for whatever reason, reach the same level of success as the British subjects.

The third question asked how the relation between mean time taken and mean accuracy varies between L1 and L2 groups across different types of test. Here the somewhat surprising answer was that this relation was not greatly affected by test type. On fixed-time tasks the L2 readers seemed to achieve about three quarters as much as the L1 readers in a given time, and on self-timed tasks the L1 readers seemed to take about three-quarters as long to achieve a given product.

At least for the various types of text examined here – from 1 to 800 words -- the Swedish subjects read as well (or as badly) as the British ones, but required more time. This is in conformity with Walczyk's hypothesis (2000) concerning faster and slower first-language readers, viz. that slower readers were slower because they made use of more controlled processes to reach the same result as faster ones. Bernhardt's proficiency threshold (2005), above which L1 literacy compensates for lack of L2 proficiency, is in agreement with Walczyk's hypothesis in that for both L1 and L2 readers more controlled processes are used when automatized processing is lacking. That is, above a certain threshold one can indeed achieve native-like accuracy, but not at native-like speed. The 'compensation' used by advanced L2 readers is similar to that of slower L1 readers. Both groups rely on controlled, conscious processing to pick up the slack from automatized processing. Slower (educated) L1 readers can be supposed to lack automatization at several levels of processing and more conscious, controlled processing is used, slowing down the overall comprehension process. This would also be true of advanced L2 readers, who have weaker automatization of word recognition, parsing, pronoun recovery, etc. Again, more controlled processing is used to 'compensate' for lack of automatized processing, but this requires more time.

It can be assumed that the advanced L2 readers have higher-level coherence strategies transferred from their L1, i.e. transfer of L1 literacy (Berhardt 2005). On the assumption that all levels of processing of the advanced L2 readers will be less automatized than comparable processing among the L1 readers, it seems reasonable to assume that advanced L2 readers would make considerable use of their higher-level coherence processes (i.e. their literacy skills). An indication that advanced L2 readers may especially rely on transferred literacy skills is the result that the advanced L2 readers attained L1 level accuracy scores on the sentence and scenario tasks (see Table 3 above). The processing difficulty of these tasks primarily concerns inter-clausal relationships, i.e. literacy and coherence.

We can go further and consider the types of skill that the two groups seem to manifest. Similar general results seem to have arisen from four rather different tasks, and it is useful to see what these involve.

Reading a short text and answering a set of questions about it requires the construction of a text model. However, such text models can be fairly loose, i.e. it is not necessary to store details since the reader can always go back and check the text. For questions concerning specific content, if the text model is weak or has gaps, looking back through the text, i.e. skimming for specific information, will be a viable approach to finding the appropriate information. One strategy for this kind of task is first to read for general orientation (laying foundations, e.g. Gernsbacher 1990), then to read the questions to see what information is crucial, and then to re-read the text with these questions in mind. Interviews with some of our L1 and L2 subjects confirm that this was what they did. So the task may require what Carver (1990) calls “skimming” reading followed by “scanning”. It seems plausible that reading for this type of comprehension test is similar to problem solving, and hence not the same kind of reading which is done for studying or learning.

By and large, the L2 group was as accurate as the L1 group on this task. Even for questions requiring inferencing or summarizing the L2 group performed comparatively well (though neither group did as well as on these questions as on simpler ones). Thus lack of language knowledge was not causing the L2 readers to answer some questions incorrectly that L1 readers generally answered correctly. On the other hand, it was making them slower at skimming and scanning.

In contrast, the recall task relies on the construction of a model that can be stored and reproduced (if any substantial number of propositions is to be recalled). The productive element is a confounding factor, in that one can imagine subjects varying in ability to verbalise the model stored or experience of so doing. We believe this to be a fairly small factor, because our measure of protocol quality did not demand coherence, organization, or even consistency of language. Background knowledge has to be used to judge what is likely to be significant and to augment explicit text cues concerning how propositions are related to one another. The propositions in the text have to be remembered either as isolated details or as parts of a coherent model (and as noted above some propositions lend themselves to one and some to the other). This is a task that relies less on question-answering strategies and more on study-reading strategies. Although it superficially relies on memory to a considerable extent, what becomes retrievable from memory will depend on the kind of model that the reader constructs. In addition to a coherent overall model, it also involves the kind of reading that Carver (1990) calls “a learning process” and “a memorizing process” Carver claims that for “learning” type reading readers must rehearse facts, regress and repeat in order to better anchor important aspects of the model and thereby to succeed in remembering the ideas for a recall protocol.

Researchers who lay stress on the situated nature of literacy and the importance of background knowledge (Lea and Street 1998) might expect different results from the recall task than from the multiple-choice one, with perhaps less impact from L1/L2 status. In fact, however, more uniform background knowledge did not produce more similar results across the L1-L2 divide or was outweighed by the extra memory burden of the recall task. While the nature of the current recall test makes it impossible to confirm or deny this, it was our impression that subjects were under less time pressure in this task than in the multiple-choice one. The slightly lower scores for the L2 group on this task may be due more to either the recall or the study-reading element than the time constraint. But this is a question for future research.

The timed computer-based tasks hardly allow for different reading strategies and only pose modest memory loads. The scenarios task is similar to the recall task just discussed in that both require storing information for later use (a kind of reading for learning/memorizing). They are otherwise quite different in that the former only requires

remembering a very short scenario (two sentences) and the latter requires modeling an 800-word text. In any case, both require the construction of a mental model of the text since it must be recalled later (i.e. without having access to the text for reviewing or rereading). In contrast, the sentence coherence task only requires reference of this mental model to everyday real-world knowledge of typical relations among phenomena and does not make any recall demands on the reader. Nevertheless the difference in reading efficiency seems to be of the same order in these two tasks as in the reading comprehension (where we know from coparison of scores on the first four questions with overall scores that speed is the main factor in the difference) and the recall task (where we suspect that time is not the main factor).

Nevertheless, the broader implication is that the common view that advanced L2 readers can achieve nearly the same comprehension product as their L1 counterparts seems only to apply when L2 readers actually take the additional time required to “read for learning” that is necessary to achieve a text model comparable to that of L1 students. If, however, university students in Scandinavia, as advanced L2 readers of English language textbooks, are less than 75% as efficient as the L1 readers of the same textbooks, but cannot put in the necessary extra time then clearly they are at a disadvantage when competing with their L1 counterparts.

References

- Abu-Rabia, Salim, and Siegel, Linda. 2002. “Reading, Syntactic, Orthographic, and Working Memory Skills of Bilingual Arabic-English Speaking Canadian Children.” *Journal of Psycholinguistic Research*. 31 (6), 661-678.
- Bernhardt, Elizabeth. 2005. “Progress and Procrastination in Second Language Reading.” *Annual Review of Applied Linguistics*. 25, 133-150.
- Brantmeier, Cindy. 2006. “The Effects of Language of Assessment and L2 Reading Performance on Advanced Readers’ Recall.” *The Reading Matrix*. 6 (1), 1-17.
- Carver, Ronald P. 1990. *Reading Rate: A Review of Research and Theory*. New York: Academic Press.

- Cummins, Jim. 1984. "Wanted: A Theoretical Framework for Relating Language Proficiency to Academic Achievement among Bilingual Students." In C. Rivera (ed.), *Language Proficiency and Academic Achievement*. Clevedon: Multilingual Matters. 20-27.
- Kjersti Fløttum, Trine Dahl, and Torodd Kinn. 2006 *Academic Voices: Across Languages and Disciplines*. Amsterdam: John Benjamins.
- Fraser, Carol A. 2007. "Reading Rate in L1 Mandarin Chinese and L2 English across Five Reading Tasks." *Modern Language Journal*, 91(3), 372-394.
- Fukkink Ruben G., Hulstijn, Jan., and Simis, Annegien. 2005. "Does Training in Second-Language Word Recognition Skills Affect Reading Comprehension? An Experimental Study." *The Modern Language Journal*, 89 (1), 54-75.
- Gernsbacher, Morton Ann. 1990. *Language Comprehension as Structure Building*. Hove and London: Lawrence Erlbaum Associates.
- Geva, Esther; Ryan, Ellen B. 1993. "Linguistic and Cognitive Correlates of Academic Skills in First and Second Languages." *Language Learning*, 43(1), 42.
- Graddol, David. 2006. *English Next*. British Council.
- Haarmann, Henk J, Davelaar, Eddy J and Usher, Marius, 2003. "Individual Differences in Semantic Short Term Memory Capacity and Reading Comprehension." *Journal of Memory and Language*, 48, 320-345.
- Koda, Keiko. 2004. *Insights into Second Language Reading: A Cross-Linguistic Approach*. Cambridge: Cambridge University Press.
- Lea, Mary, and Street, Brian. 1998. "Student Writing in Higher Education, an Academic Literacies Approach." *Studies in Higher Education*, 23 (2), 157-172.
- McMillion, Alan, and Shaw, Philip. forthcoming. "British and Swedish Students Reading English." In *Proceedings of the BALEAP Conference, Durham, April 2007*.
- Perfetti, Charles A. 1999. "Comprehending Written Language: A Blueprint of the Reader." In C. M. Brown & P. Hagoort (eds), *The neurocognition of language*. Oxford: Oxford University Press. 167-208.
- Steffensen, M.S., Joag-Dev, C. and Anderson, R.C. 1979. "A Cross-Cultural Perspective on Reading Comprehension." *Reading Research Quarterly*. 15 (1), 10-29.

- Segalowitz, Norman. 1986. "Skilled Reading in the Second Language." In Vaid, J. (ed), *Language Processing in Bilinguals: Psycholinguistic and Neuropsychological Perspectives*. London: Lawrence Erlbaum Associates. 3-19.
- Segalowitz, Norman, Poulsen, Catherine, & Komoda, Melvin. (1991) "Lower Level Components of Reading Skill in Higher Level Bilinguals: Implications for Reading Instruction." *AILA (Association Internationale de Linguistique Appliquée) Review*, 8, 15-30.
- van den Broek, Paul, Ridsen, Kirsten, and Husebye, Elizabeth. 1995. "The Role of Readers' Standards of Coherence in the Generation of Inferences during Reading." In Lorch, R. F. and O'Brien, E. J. (eds.) *Sources of Coherence in Reading*. Lawrence Erlbaum: Hillsdale, NJ, 353-373.
- van den Broek, Paul, Virtue, Sandra, Everson, Michelle Tzeng, Yuhtsuen and Sung, Yung-chi. 2002. "Comprehension and Memory of Science Texts: Inferential Processes and the construction of a Mental Representation." In Otero, Jose (ed) *Psychology of Science Text Comprehension*. Mahwah, NJ: Lawrence Erlbaum Associates, Incorporated. 131-154.
- van Gelderen, Amos, Schoonen, Rob, de Glopper, Kees, Hulstijn, Jan, Simil, Annegien, Snellings, Patrick, and Stevenson, Marie. 2004. "Linguistic Knowledge, Processing Speed, and Metacognitive Knowledge in First- and Second-Language Reading Comprehension: A Componential Analysis." *Journal of Educational Psychology*. 96 (1). 19-30.
- Walczyk, Jeffrey J. 2000. "The Interplay between Automatic and Control Processes in Reading." *Reading Research Quarterly*. 35 (4), 554-566.
- Zwaan, Rolf A., and Radvansky, Gabriel A 1998. "Situation models in language comprehension and memory." *Psychological Bulletin*, 123, 162-185.