Fluency in second language writing: the effects of enhanced speed of lexical retrieval
Snellings, P.J.F.

Citation for published version (APA):
CHAPTER 5

AN EXPERIMENTAL COMPUTER PROGRAM TO TRAIN LEXICAL RETRIEVAL

ABSTRACT
Word recognition and lexical retrieval are essential aspects of reading and writing, respectively. Improving the speed of these processes is important for both young children who are learning to read and older students who have to read and write in a second language (L2). We showed that a computer tool, the Nelson-Writing training program, was effective in enhancing both lexical retrieval and lexical access in a classroom context. In addition, the resulting increased fluency in lexical retrieval had positive effects on the use of lexical material productively, as well as on the expression of essential content elements in writing. The program provides feedback on the basis of measures of reaction time (RT). So far, this type of information has been predominantly used for testing objectives and has been restricted to laboratory conditions. Because this is the first training program that specifically focuses on speed of lexical retrieval, it may serve as a useful tool for future research into this area. In addition, as it has proven its potential in an educational setting, it can also be adapted for educational purposes. This paper gives a detailed description of the design principles underlying the Nelson-W program, the way it functioned in a classroom experimental setting and examples of exercises used in an experimental application.

INTRODUCTION
Efficient and fast lexical retrieval is important for fluent language production. Spending too much attention on lexical retrieval may have adverse effects on the writing and speaking process because working memory (WM) capacity is severely limited (for a discussion of attention limits see Kellogg, 1999; Lea & Levy, 1999). The effort involved in lexical retrieval, which is part of task difficulty, is an important factor in determining to what extent cognitive capabilities are taxed. If lexical retrieval proceeds with limited conscious effort, WM capacity will be less heavily taxed. In contrast, when a second language (L2) is involved, lexical retrieval becomes problematic. As a result, more attention for lexical retrieval results in less attention that can be devoted to planning, monitoring or other cognitive operations used in writing (Hayes, 1996; Kellogg, 1996). Already in 1980, Flower and Hayes (Flower & Hayes, 1980) suggested that by making sentence production processes more efficient (e.g. with sentence-combining exercises), the writer has time to concentrate on other important constraints of writing. In the case of comprehension, it has been shown that word recognition is an important aspect of reading proficiency (LaBerge & Samuels, 1974; Perfetti, 1985). In addition, it has been argued that poor readers have deficient phonological skills and as a result are slower in word recognition with negative effects on reading ability (e.g. Wagner & Torgesen, 1987; Yap & van der Leij, 1993). Previous research suggested that remedial activities that aim to improve speed are more effective than activities with an emphasis on accuracy (Van den Bosch et al., 1995). Because word recognition is slower in the L2, claims have been made as to the importance of speed of word recognition for reading proficiency in L2 as well (e.g. Segalowitz et al., 1991). In the current chapter, we will focus on fluent language production in the L2. To investigate whether an experimental speed up of lexical retrieval was possible we developed a computer tool with an emphasis on speed. This computer tool, called the Nelson-Writing (henceforth, Nelson-W) program, will be the topic

of this chapter. We used this Nelson-W program to experimentally enhance the development of lexical access and lexical retrieval skills.

In one study, we looked into the possibility to increase lexical retrieval speed and speed of lexical access with this newly developed computer tool (Snellings et al., 2002, Chapter 2). Students were trained in class for 4 weeks. As some components involved in lexical access and lexical retrieval are similar (Levelt, 1989), we had expected effects of our training in lexical retrieval to surface on a measure of lexical access (a lexical decision task) as well, although the effects were expected to be less pronounced than on a direct measure of lexical retrieval. Results showed that the experimental training resulted in better accuracy and in faster reaction times (RTs) on the lexical decision task. In the case of our measure of lexical retrieval, the Written Productive Translation Task (WPTT), we expected larger effects. As predicted, our results on the WPTT showed a significant difference; the trained students were both more correct and faster on words trained in their condition in comparison to students who were not trained on these words. In sum, we showed that a 4-week, highly focused training aimed at speed-up of responses, resulted in both higher hit scores and faster RTs on lexical retrieval. The study demonstrated that the training with the Nelson-W program effectively improved an essential sub-process involved in L2 written production.

In a subsequent study, we looked at the effects this increased speed in lexical access and lexical retrieval had on writing quality (Snellings et al., submitted, Chapter 4). The results for two sets of trained words showed that in each case students in the trained words condition could use more words productively than students who were not trained. In the case of both sets of words the size of the effects was medium with 26% of the variance in scores due to the training in the case of the first set of words and 24% in the case of the other set. Regarding the expression of essential content elements we found that students in one trained condition used more content elements than the untrained students. A total of 6% of the scores could be accounted for by the training. Although the differences in the second trained condition were not statistically significant, the direction of the differences in expressed content elements was similar to that in the first trained condition. These results suggest that when students had more attention to spare because lexical retrieval of the trained words proceeds effortlessly, this was beneficial to their writing processes. Besides testing the predictions on used words and expressed content elements, we explored whether there were any effects of the training on the scores of Global text quality. These data showed that even though the results were in the expected direction in the case of one writing assignment, the means of students in the trained words conditions on both tasks were not significantly higher. It appeared that the effect of enhanced lexical retrieval was not large enough as to significantly influence scores on a global text level. Still, the Nelson-W program succeeded in improving the ability to use vocabulary productively in written text and in one instance also improved the detailed expression of content, both important aspects of writing.

Clearly, the program we developed has been successful in increasing both lexical access and lexical retrieval, skills that are essential for fluent reading and writing. So far, other experimental training programs with a focus on speed have been developed for remediation of poor reading skills (Frederiksen, Warren, & Rosebery, 1985; Van den Bosch et al., 1995). As a consequence, they have focused on lexical access or the access to sub-lexical units. In contrast, the current aim was to increase lexical retrieval for productive purposes. Although lexical access and lexical retrieval share certain processes (Levelt et al., 1999), language production requires additional skills. For this reason, the current program requires students not only to exhibit receptive knowledge but also to use their knowledge in written production. In addition, this program has different feedback possibilities than existing training software, including both feedback on correctness and on speed (reaction times) on the same items in completed exercises. So far, programs that do make use of reaction time have
generally been used individually and in laboratory conditions whereas the Nelson-W program can be used with a group in a classroom. Because this is the first program training specifically lexical retrieval and it includes new features that have not been available in previous programs, it may serve as a useful tool for future research into this area. Moreover, it has already proven its potential in an educational setting and can be adapted for use in educational practice. In this chapter we will provide a detailed description of the design principles underlying the program, describe an implementation in a classroom experimental setting and give examples of exercises used in this experimental application. As a point of reference, we will also refer to a successful computer program with a focus on speed of lexical access (Frederiksen et al., 1985). Finally, we will have a more detailed look at the logfiles derived from one productive exercise (the translation exercise). The translation exercise is particularly interesting because it aims directly at the target skill of lexical retrieval. Another reason to have a closer look at the translation exercise is that students did this exercise four times, which enables us to illustrate the individual student’s development in some detail.

THE NELSON-WRITING TRAINING PROGRAM

As a point of reference we will first give a short description of a program that is similar in its aims to the Nelson-W program. The program chosen is a remediation program that has proven to be successful, and is an example of many programs of this kind (e.g. Cohen, Torgesen, & Torgesen, 1988). Frederiksen developed the program for children with a reading disability (Das-Smaal, Klapwijk, & van der Leij, 1996; Frederiksen et al., 1985). It aimed to enhance the perception of multi-letter units in words, while stimulating speed of recognition and accuracy at the same time. As pointed out by Das-Smaal, Klapwijk and Van Der Leij (1996), this program may not only increase perceptual coding skill but the resulting enhanced efficiency may decrease demands on WM and hence circumvent memory limitations. As a result, capacity can be dedicated to other components of the reading process. Therefore, even though the focus is on word recognition instead of lexical retrieval, the theoretical rationale is in fact similar to that behind the Nelson program and it is interesting to look at its design and features.

The Frederiksen program called “SPEED” resembles a car race game. Students are required to determine whether or not a target multi-letter unit is present in words shown in rapid succession. The overall goal of “SPEED” is to speed up detection without sacrificing accuracy until performance on a unit reaches a criterion speed. Immediate feedback is given on both speed and accuracy. Before practicing a particular unit, a target speed is presented on the screen. Students have to increase their speed from the initial speed until the target speed is reached for that particular unit, after which a new unit is presented. Initial speed may be set at 60 words per minute (wpm), for example, and increases in the direction of the target speed set 50 wpm higher (110 wpm). The rate of increase depends on the student’s performance. Whenever a child responds correctly within the given time limit, speed of word presentation is increased and an error light is extinguished. Whenever a mistake is made or the display time has passed, an error light is shown and display time is increased, reducing the required speed. In addition, allowing no more than four subsequent errors stimulates accuracy. If the criterion speed is not reached in one run, students continue, with the initial speed in the new run 30 wpm less than the final performance level on the previous run and the target speed 20 wpm more.

As pointed out before, although the Nelson-W program has the same focus on speed and accuracy as the Frederiksen program, it nevertheless has different aims as it concerns lexical retrieval rather than lexical access. Furthermore, the focus is on students with normal
language abilities learning the L2 as opposed to children reading in their mother tongue and having a reading disability. The Nelson-W software can be used outside the laboratory in a classroom setting, training a group of learners at the same time. For an impression of the layout of the program and screen captures we refer to the programmer’s website: http://members.ams.chello.nl/b.roovere/ (click the NELSON tab). In what follows we will describe the general design principles underlying the Nelson-W program and we will refer to the Frederiksen program when appropriate. Next, we will discuss experiences from an experimental intervention in an educational context and describe the specific exercises used. Finally, we will discuss the logfiles from the translation exercises and suggest some interesting issues concerning the learning process that they may provide insights into.

GENERAL DESIGN PRINCIPLES

The design principles outlined in this section were at the basis of the experimental intervention and the four different exercise types we used. In short, we developed a Block exercise with a focus on word order, a Correction exercise that involves correction of semantic anomalies and a Detection exercise that involves detection of semantic anomalies. Finally, we developed a translation exercise in which students have to translate Dutch words in English carrier sentences into English. See below for more detailed information on these exercise types.

1) A focus on speed
Because our aim is to stimulate fast lexical retrieval for production purposes, it is important to encourage students to respond as quickly as possible (Van den Bosch et al., 1995). Consequently, we explicitly encourage students to answer as fast as possible while keeping correctness in mind, a similar principle as in the Frederiksen program. To achieve this, we stress attention to speed in the on screen instructions at the beginning of every exercise. In addition, feedback displays the exact speed in seconds as well as speed of previous reaction on the same word as a baseline for reference to their progress. The procedure is as follows; a timer starts when the target stimulus has appeared and stops when the response key is pressed. Emphasis on speed is also obtained by complementing the numerical information with two different icons, the choice of icon depending on whether the response is faster or slower than the previous response on that same item. Whenever a student is faster we show a fighter jet plane, whereas we show a rowing boat for a slower response. In Figure 1 the screen captures show two different feedback screens for the translation exercise. A student has typed the English translation “while” of the Dutch equivalent “terwijl” in the white rectangle and has pressed the Return key. In addition, an equal sign behind the English carrier sentence precedes the correct answer that is given regardless of the correctness of the student’s answer. A tick is shown to indicate that the answer was correct, as well as current speed (“snelheid”) and previous speed (“vorige keer”). In the first screen capture a rowing boat appears because the response was slower than last time. In the second screen capture a fighter jet plane is presented because the response was faster than last time. In addition to the speed feedback, it is also possible to restrict presentation time, as a further incentive to answer as fast as possible. This “flashcard principle” entails that the item disappears from the screen after a specified time but students can still give their answer.
Figure 1. Screen captures of feedback in the Translation exercise. In the first screen capture the response is *slower* than the previous time and *a rowing boat* is shown. In the second screen capture the response is *faster* than the previous time and *a fighter jet plane* is shown. Both screen captures show the English translation the student has typed in the white rectangle: "while" for the Dutch equivalent "terwijl". An equal sign behind the English carrier sentence precedes the correct answer that is given regardless of the correctness of the student’s answer. A tick indicates that the answer was correct, and at the bottom right-hand corner current speed ("snelheid") and previous speed ("vorige keer") are shown. Underneath the rowing boat and fighter jet plane icons instruction is given to press the spacebar to continue ("Druk op spatie om verder te gaan").
terwijl I was having lunch the doorbell rang = while

1.8 sec snelheid
2.5 sec vorige keer

Druk op spatie om verder te gaan 21-2-00 16:38:
In the Frederiksen program only immediate feedback was provided on current accuracy and speed, and as long as the target speed was not reached, students could improve their speed. In the Nelson-W program, the first time a student does an exercise, feedback displays individual average reaction time until that moment as a preliminary baseline. This serves as an incentive for speed but whether the response is faster or slower than the average reaction time is rather dependent on the difficulty of a particular item. Easy or short items are bound to be faster than the average response whereas the opposite holds for longer or more difficult items. Yet, once a student has answered an item correctly, that time is taken as the target time that has to be improved. Every time that same item is practiced the baseline is the individual speed on the last correct response on that same item. As a consequence, both slow and faster students are constantly encouraged to improve their speed. If the answer is incorrect, no speed feedback is provided and a single red cross appears. At the end of each exercise we show students a results screen in which both current results (Percentage correct, Average reaction time) and results from the previous sit of the same exercise (also Percentage correct and Average reaction time) are presented.

As the Frederiksen program was intended for children with reading problems, a criterion speed was specified that presumably was high enough to fluently read sentences. In contrast to the Frederiksen program, our aim is to train both poor writers and those writers who are already more proficient. As a consequence, we cannot specify one criterion speed for all students. In addition, evidence is lacking for the exact level of speed required for different types of students.

2) The importance of correctness
In the on screen instructions at the beginning of every exercise we explicitly encourage students to keep correctness in mind. In addition, the feedback uses two different icons to indicate whether the response is correct (a green tick) or incorrect (a red cross). Whenever an answer is incorrect, information on speed is not shown on the screen, only the red cross appears to stress that correctness has to be taken into account. It is very important that whenever students make an error they carefully examine the correct answer, especially in the case of the translation exercise where spelling has to be correct as well. Therefore, to further ensure that students focus on correctness, the program incorporates an additional feature. Before using the program, it is possible to specify the minimal number of times a student must try to give the correct answer. For example, if this number is set at two, the student cannot move to the next item before having tried two times. In the case of a translation exercise they should not just register the correct answer but also pay attention to its correct spelling. In particular, we would like to avoid that students only glance at the correct answer in the feedback and quickly move on to the next item. Therefore, for the translation exercise we recommend a setting of three. Whenever students make a mistake and keep on making it, it forces them to look carefully at the feedback for at least three times and to check whether their response is not only the target word but is also correctly spelled. Finally, it has to be noted that giving correct feedback is difficult in the case of the translation exercise, as several alternatives are possible. In our discussion of the translation exercise we will go into this in more detail.

3) Flexibility of stimuli and instructions
We programmed the Nelson-W program in such a way as to provide optimal flexibility for each individual user. It is possible to use written stimuli, as well as pictures and sounds. The written stimuli can be entered in simple text files (.txt extension), the pictures are in bitmap (.bmp) format and the sounds in wave sound (.wav). In addition, a combination of written text, a picture and a sound can be used simultaneously in one stimulus.
The program further allows for easy construction and adaptation of stimuli. In the case of the translation exercise, the input is not restricted to single words and if the aim is to train collocations, combinations of words can also be entered. Because the only requirement is that stimuli are entered in a text file, most languages are possible, as long as the character set is installed on the computer used. Before using the program, the user can personally specify the desired settings for each exercise by ticking the appropriate settings in a menu or by typing in the required settings. In this way, the user can personalize the settings for each of the exercises, and accordingly adapt the instructions, effectively creating different variants. If an exercise has been done before, example items can be skipped the next time the exercise is done, preventing redundancies. If a sole focus on correctness is desirable, reaction times can be omitted. In contrast, if the focus is only on speed, error feedback in both item feedback and the result screen at the end of the exercise can be omitted. Finally, if considered necessary for speed up or as an additional incentive to speed up, presentation time can be specified. After specifying all the settings at the user’s computer, a copy routine in the menu can be used to put the program on installation disks.

4) The avoidance of sequential learning
To make sure students learn the items individually and are able to use them on their own, the order in which items are trained in all exercises is randomized each time before the students commence an exercise. In the block exercise the order of the two constituents is also randomized. In this way students do not learn sequences of items, and previous items in a list do not trigger items that are to follow. This is important because the discussion in Anderson (2000) makes clear that transfer may be problematic if expertise has to be applied in different contexts. In this particular situation, students have to use isolated items outside the list in writing. By randomizing the order of items students are presumably better able to retrieve them when they are needed in writing.

5) Progression from receptive to productive learning
Under the assumption that productive knowledge is conditional upon receptive knowledge, the Nelson-W program is designed to incorporate exercises aimed at both kinds of knowledge. In our experimental application, each specific word can be trained with four different types of exercise, moving from receptive skills to the target production skills. It is also possible to consistently train the same meaning of the words used, so that mappings of stimuli to responses are consistent rather than varied (Schneider & Shiffrin, 1977; Segalowitz & Gatbonton, 1995). The aim of the present program is to increase lexical retrieval speed. In our approach we therefore started with a receptive exercise focusing on the correct position of the target words in the sentence (Block exercise). The focus then changed to the meaning of the items but once more using a receptive format. Students have to decide whether the target item makes sense in the context of a sentence and are either required to choose an alternative if it does not (Correction exercise), or to indicate which of two words in a sentence makes the sentence odd (Detection exercise). Finally, students practice the specific skills necessary for productive language skills. In the translation exercise students have to translate Dutch words in an English carrier sentence as fast as possible in suitable English equivalents.

6) Administrative flexibility
If an experiment involves different schools, it is possible to choose the specific school the program is going to be used at for that particular session. The program incorporates a database so that students only have to type in their identification number. Once they have typed their number their name automatically appears on the screen. This prevents them from making typing errors (or using capital letters alternately) and ensures students’ feedback is based on
their own past performance (correctness and reaction time). If students from different classes work on the same computer, specifying the class beforehand by the user, makes sure that the name of the student who worked on this machine the previous time will appear on the screen automatically. The program is designed in such a way that it can also function in a network environment. Using the program in this environment has important advantages. First, data are less likely to be lost because they are not just saved on one pc. Second, students do not have to work on the same computer each session; entering the student’s personal code automatically retrieves the appropriate information on previous reaction times to be used in individualized feedback.

7) The ability to monitor the individual’s training processes
To be able to keep track of each student’s personal progress the Nelson-W program saves a personal logfile for each student (see below). A record is kept of the computer that the student works on, name of the exercise, date and time. The student’s personal identification number, name, school and class are also saved. In the logfiles, a P for practice item or an N for a target training-item precedes every item. The time to complete the item follows, then a letter combination indicating correct (“gd” for the Dutch “goed”) or incorrect (“ft” for the Dutch “fout”). Next, the typed response is displayed (e.g. “missing”), followed by the context sentence including the target item in Dutch and finally, the intended correct response (e.g. “missed”).

```
Desktop-id : Laptop 15
Toetsnaam (“name of the exercise”): translateta
Datum (“date”): 29-11-99
Tijd (“time”): 14:31:36

Leerlingnr (“student identification nr.”): 88888
Naam (“student’s name”): Lex
School (“name of the school”): De Domtoren
Klas (“class”): 3

P 4476 gd first [eerst] he ate a sandwich, then he left = first
N 5016 gd while [terwijl] I was having lunch the doorbell rang = while
N 7260 ft missing she fired the gun but [miste] = missed
```

As we will show in more detail later, these data provide information on the way a student reacts at each moment in time; from the first time he/she responds to an item to the last time the same item is practiced.

CONTENT SPECIFIC ASPECTS

We will now describe the types of exercises used in our experimental training (Snellings et al., 2002, Chapter 2).

**Block exercise**
This exercise can be used equally well in a training that focuses on enhancing lexical access or in training with a focus on enhancing lexical retrieval skills. The aim of this exercise is to both strengthen the receptive knowledge of words and to train syntagmatic knowledge.
(collocational restrictions). Students have to complete a short sentence correctly as fast as possible by choosing the next to follow from two constituents, one of which results in an incorrect word order, as in the following example:

After some time
woke up she
Z M

Students have to press either Z or M for the constituent they think will be the first one following After some time.

Correction exercise
To strengthen the connections between the concepts and the lexical forms, this exercise has a strong focus on the meaning of the words involved. This is relevant both for lexical access and lexical retrieval skills. In the correction exercise, students are asked to judge words for appropriateness of use and if the context requires it, to choose a better substitute from two alternatives. This exercise aims at enhancing the connections between the meanings and the lexical forms. Students have to focus on the meaning of the words in a short sentence. They have to detect semantic anomalies as fast as possible and can indicate which of two alternatives will make sense (or accept the present alternative marked “X”), as in the following example:

Sue felt very ill. After some ages (X) she felt better though.

time centuries
Z M

Students have to indicate whether ages (X) is acceptable, or whether replacing it with time (Z) or centuries (M) will yield a more semantically acceptable sentence. Here, time (Z) would be the correct answer.

Detection exercise
Similar to the correction exercise, the detection exercise is suitable for enhancing both lexical access and lexical retrieval skills. Students again have to decide whether the meaning of the words and expressions in the specific context presented is appropriate. They have to focus on the meaning of the words in a short sentence and detect semantic anomalies as fast as possible, indicating which word causes the anomaly, without being able to choose from alternatives, as in the following example:

We went to the pool while(Z) we arrived(M) home.

Students have to indicate whether while (Z) or arrived (M) makes the sentence odd. Here, the word while makes the sentence odd as a word like after is more likely.

Translation exercise
To train lexical retrieval we needed a task that relied on productive processes. Picture naming tasks can be assumed to tap lexical access. In addition, the process tapped proceeds from meaning to word form just as in language production in writing. Because our focus was not restricted to single nouns and verbs, picture naming tasks were not considered to be appropriate. Rather, we needed a method that could incorporate nouns, pronouns, articles,
verbs, adverbs, adjectives and combinations of these categories. These words are difficult to depict unambiguously with pictures, and therefore we considered a word translation task. The Translation task taps similar processes as in picture naming and makes a focus on more words at the same time (in contrast to picture tasks) possible. The word translation task has learners fill in the correct translation within a meaningful sentence, a situation that closely matches lexical retrieval in written production. Students have to translate Dutch words in simple English carrier sentences. In our experiments, students started with exposure durations of 30 seconds. As an additional incentive to speed up, the third and fourth time an exercise was done, limited exposure durations of 15 seconds were imposed as recommended by Van den Bosch, van Bon and Schreuder (1990). In this so-called “flashcard version” (limited presentation time) the Dutch words and the English sentence disappeared from the screen after 15 seconds, stimulating the students to give faster answers. The following example is illustrative of the task:

voordat the bridge collapsed he had reached the other side

Translation: before

An additional option in the translation exercise is to specify a list of several alternative correct answers. Therefore, different correct answers are acceptable (e.g. “large” and “big”). In the correct feedback the typed answer is repeated if it is in the correct answers list (which includes words used by students in a pilot test). Still, in the translation exercise, giving appropriate feedback is complicated because errors can be diverse. They could be spelling errors, grammatical errors or words that are completely inappropriate. The current approach makes sure that whenever a mistake is made, all students receive the same feedback; the first correct answer in the correct answers list. As a result, in the case of errors all students are focused on the same lexical items and this is an advantage later on in testing because test results on the translation test are based on the same items.

Our experiences with the implementation of the program in a pilot study and in an experiment involving more students made clear that some aspects of the program need improving. In the next section we will discuss these experiences because we believe they are relevant to users implementing this or a similar program to enhance language production.

FUNCTIONING OF THE PROGRAM IN AN EXPERIMENTAL APPLICATION

In a pilot study involving 21 students (grade 8, age 13-14, pre-university education) the block and translation exercises were implemented in order to establish not only the proper functioning of both software and hardware but also the adequacy of the instructions and the exercises for our aims. On the basis of our experiences we clarified the instructions. In particular, students did not read the feedback because they assumed it reflected negatively on their reaction time. Therefore, the instructions had to make clear that reaction time was already measured the moment they pressed “Z” or “M” to indicate their choice of alternative (or the Return key in the case of translating), so that they could carefully read the feedback before pressing the spacebar for the next item. Similarly, in the case of the “flashcard version” we made clear to them that they could still answer after the sentence had disappeared from the screen. In the Block exercise, we also put the letters on the screen (“Z” and “M”), which were used to indicate the two alternatives, somewhat closer together to make both of them more salient. In the case of the Block exercise, questions asked immediately after the students had
finished the exercises showed that they focused both on correctness and speed, although they
paid more attention to item speed than to average speed. In the case of the Translation
exercise, due to its difficulty, most students focused on correctness instead of speed. Many
errors on the Translation task involved spelling errors. As the direction in average speed
(faster vs. slower) can also be easily deduced from the icons (fighter jet plane vs. rowing
boat), we emphasized the importance of the icons in the instruction.

In the experiment (Snellings et al., 2002, Chapter 2) we incorporated the findings from
the pilot study into the Nelson-W program and we included students from two different levels
in Dutch education. A total of 103 Dutch students participated, all in the beginning of their
third year of secondary education (grade 9, age 14-15). Students were sampled from four
different classes, two classes from lower general secondary education (n=47) and two classes
from pre-university education (n=56). Students were trained in a school setting in scheduled
English classes. Half of them worked on laptops, the other half on network machines.
Standard duration of each training was 50 minutes (except for one training of 45 minutes) and
there was a one week interval between each training session.

Students seemed to enjoy working with the program, presumably because of the game­
like structure and the fact that they could record personal improvements in speed and
accuracy. Because we had tested two exercises (Block and Translation) in the pilot study and
the underlying framework (time and correctness feedback) of all exercises was similar, there
were few problems in the final version. As expected, students from the lower educational
levels were more error prone, especially in the first session. Nevertheless, because of the
highly structured program and the clear corrective feedback this presented no problems. The
Detection exercise however, an exercise type that was not tested in the pilot study, appeared
to be particularly difficult for many students. In constructing the items we made sure that only
one of two highlighted words in the sentence made the sentence semantically odd (for
example: If you like heat you can go to the North-Pole). Unfortunately, it is often possible to
make a sentence semantically acceptable by replacing other words than the two highlighted
(in the example, replacing “heat” by “cold” would also make the sentence semantically
acceptable). To train a specific set of items, we wanted students to focus on one item only. To
achieve this, we developed an alternative format after we had conducted our experiment. In
this format only one item is highlighted, and students have to decide whether it is
semantically appropriate or not. The program randomly inserts either a word that is acceptable
or a word that makes the sentence semantically odd. In both cases students get feedback
showing the correct word in the sentence. Students either see: “If you like heat you go to the
North-Pole” or “If you like cold you go to the North-Pole”. In both cases, besides from
getting feedback on their answer (correct, incorrect and reaction time) students are presented
with the second sentence as feedback because only in this sentence the highlighted word
(“cold”) is semantically acceptable.

LOGFILE DATA FROM TRANSLATION EXERCISE

In this section we will illustrate the kind of information the logfiles recorded by the Nelson-W
program can provide us. They can give us a better insight into the development of lexical
retrieval. These data can also tell us something about the stability of the skill involved as the
computerized format enables us to look at the development in reaction times. Is there a stable
pattern to be found and if this is the case, does it depend on the type of words involved for a
particular student? In such analyses however, the relationship between reaction times and
correctness has to be taken into account. In Table 1 a negative score indicates an increase in
speed between the first time the translate exercise was completed and the fourth time.
Table 1. Increase in speed on the translation task: difference in reaction time (RT in ms) between time 4 and time 1 for 4 items.

<table>
<thead>
<tr>
<th>Item</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>first (4-1)</td>
<td>3373</td>
<td>-377</td>
<td>33</td>
<td>-1507</td>
</tr>
<tr>
<td>but (4-1)</td>
<td>-2132</td>
<td>-2829</td>
<td>-26</td>
<td>-1536</td>
</tr>
<tr>
<td>a rope (4-1)</td>
<td>-585</td>
<td>invalid*</td>
<td>-2723</td>
<td>-28</td>
</tr>
<tr>
<td>big fish (4-1)</td>
<td>-4081</td>
<td>-3795</td>
<td>-947</td>
<td>-3225</td>
</tr>
</tbody>
</table>

* No reaction time is reported because one of the two responses was incorrect.

Although the general trend seems to be an increase in speed from time 1 to time 4, it is evident from Table 1 that there are exceptions (student 1 and 3 on the item “first”). Clearly, it is important to have a closer look at the items and students involved and it has to be stressed that more data from one particular student have to be examined before drawing firm conclusions about general patterns. At the moment this is rather laborious because the information in the logfiles is in random order for each individual student and this order is as a consequence different each time the exercise is trained.

A more elaborate approach using the logfile data is to look at the development between two moments in time. We will now show the changes in speed for the item “a rope” from time 1 to time 2, from time 2 to time 3, and from time 3 to time 4 (between time 2 and time 3 there was a 1-week interval).

Table 2. Increase in speed on the translation task: difference in reaction time (RT in ms) from time to time for the item “a rope”.

<table>
<thead>
<tr>
<th>Time</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>1796</td>
<td>invalid*</td>
<td>-2788</td>
<td>367</td>
</tr>
<tr>
<td>3-2</td>
<td>-3223</td>
<td>invalid*</td>
<td>819</td>
<td>-437</td>
</tr>
<tr>
<td>4-3</td>
<td></td>
<td></td>
<td>-754</td>
<td>42</td>
</tr>
</tbody>
</table>

* No reaction time is reported because one of the two responses was incorrect.

In Table 2 a negative score indicates an increase in speed. In the case of “a rope” we see that student 1 improves the most from time 2 to 3. Student 3 improves most from time 1 to 2, and student 4 from 2 to 3. For drawing firm conclusions, more analysis is necessary but the point we would like to make here is that the Nelson-W program provides data on each individual moment in time and therefore makes such explorations and analyses possible.

DISCUSSION

Lexical access and lexical retrieval are important aspects of fluency in language skills. The Nelson-Writing program we reported on in this paper has proven effective in enhancing lexical retrieval skills that are important in writing (Snellings et al., 2002; Snellings et al., submitted, Chapter 2 and 4). The program is easy to use and can be adapted by individual users to meet individual preferences and different goals. In addition, the content is not restricted to one particular language. So, for both research and educational purposes, a whole range of languages can be considered. Although we used the program in an experimental context, the program has proven its potential in a classroom setting. We used the program in
class with a large group of students at the same time. It proved to be useful in training students of different levels of English proficiency. It also appeared to motivate both proficient and less proficient students, presumably because of the game-like structure and the individual feedback on both speed and correctness. For research purposes we trained the same set of stimuli in each group of students. Clearly, when using this program for educational purposes the choice of stimuli is all-important. Ideally, item difficulty should be adapted to the abilities of individual students. Although the program can also be used to learn unfamiliar words, its value lies in its focus on speed. The program provides a useful tool for teachers who want to train students to use words productively in writing. In contrast to the approach often taken in language classes, teachers do not have to stop when students know receptively what a word means, or when students are able to use the words productively only at the expense of considerable effort. That this is a worthwhile approach was shown by our experiment, in which trained students improved their ability to use a set of words productively and students in one trained condition were also better able to express the content of stories. Training students in words that they are already familiar with is a laborious task that all too easily results in uninspiring drills. In contrast, the Nelson-Writing program appears to be a motivating approach in focusing students on both speed and correctness and in this way enhancing their productive language skills.

Apart from being a useful training instrument, The Nelson-W program provides important information about the learning process. The logfiles keep record of the responses given, whether they are correct, and how long it takes students to give their answers (reaction times). In the training, the order of the items is randomized and the logfiles capture the actual order in which the items were presented to each individual student. To make analysis easier, we aim at a future version of the program, in which logfiles include a numbered index of the items, so that the randomized order can easily be changed in a standard order. In a future version, the logfile data should also be complemented by a graphic display of the reaction times. This would enable both teachers and students to obtain a quick overview of the pattern of development in the logfiles.

The logfiles provide the tools to investigate claims about the individual development in reaction times for different types of learners, as these data show the progress from each moment in time to the next. These data also enable us to look into the effects of knowledge of specific words on development of accuracy and speed of retrieval. Do students show consistent behavior once they have learned the right responses or do they keep on making mistakes? If the latter is the case, it is informative to look into the kind of mistakes made. In addition, we can see for which words this is the case; which words are consistently mastered and which ones are not. Does speed increase only after a particular item has been answered consistently correct or do responses increase in speed before this point? One instance of the latter could be the occurrence of spelling errors that are simply the result of carelessness. In contrast, when students are still struggling to find an adequate answer, speed does not seem the main focus for these students, an assumption that was confirmed in the pilot study. Another interesting possibility is to determine whether students who start slower can benefit more from the training.

We have discussed the Frederiksen approach, in which a criterion speed was set for all students. As we do not aim at poor writers only but also at more proficient writers, one criterion speed for all students seems counterproductive. Moreover, evidence is lacking for the exact level of speed required for different types of students. For the training to have beneficial effects for more advanced students who are already fast at the outset of the training as well, specifying criterion speed beforehand could be justified only if this speed were above these advanced students' starting levels. Yet, such an approach would force slower students to aim for unattainable levels or spending too much time on single items. In sum, setting levels
beforehand may restrict fast students or overburden slow students. Nevertheless, we do believe that the criterion speed in the Frederiksen program could be used as an additional design principle if it were used in a slightly different way. Instead of setting a high criterion speed that is high enough to challenge proficient students, a minimal criterion speed could be set that ensures that low proficient writers are encouraged to reach this minimal level. If an appropriate level can be specified, such an approach would stimulate all students to become as fast as possible, while ensuring that less able students attain at least speeds above the specified minimum level.

Finally, we will suggest some possible exercise types that can be developed on the basis of the Nelson-W program and that can be incorporated in future versions to be used both for research and educational purposes. To practice grammar, the format of the Detection exercise can be used, asking students to indicate whether the used tense and verb forms are correct in that context. In addition to exercises focusing solely on receptive skills, the underlying framework could easily be used to focus on the productive use of grammar. For example, students could be asked to write words in the correct tense required by a specific context. Spelling problems could be practiced along the same lines as the grammar exercises. In this type of exercise, students would have to decide whether the correct spelling has been used. Alternatively, students would have to produce words with problematic spellings correctly in a specified context. In addition to the Block exercise, which focused on receptive knowledge of word order, another option would be to use the format of the translation exercise to train productive use of word order. Students would be presented with a scrambled sentence that they have to change into a possible English sentence. Alternatively, the beginning of a sentence could be displayed and the remaining part of the sentence is scrambled for students to place in the correct order.

The possibility to use pictures and sounds simultaneously to the written texts may strengthen the learning of target items in their written form because this cross-modal elaboration is expected to have positive effects on retrieval from memory (Paivio, 1986). Instead of using a combination of different modes it is also possible to restrict the stimuli to the visual or oral mode if either the written native language has to be avoided or speed of listening is the focus. In the visual mode for example, students could be presented with pictures and they have to write down the word that goes with it in its correct spelling. Alternatively, they would hear a certain word and have to use this oral input to type the written form down.

In this paper we have described the Nelson-W program for training speed of lexical retrieval in writing. We have made some suggestions for useful additions and optional new exercises that we believe are worthwhile to consider and that can readily be implemented in the Nelson-W framework. With these extensions to the existing program, the Nelson-W program could be a valuable asset to both language education and research into language production.