

Convergent or divergent development of word recognition
in beginning reading?⁴

⁴Snel, M.J., Aarnoutse, C.A.J., Terwel, J., & van der Veld, W.M. (submitted).
Convergent or divergent development of word recognition in beginning reading?

Abstract

In this study we analyzed the development of beginning reading. We have measured word recognition of 178 children 4 times during grade 1. Our research questions were: How does word recognition in general develop during grade 1? How does word recognition develop during grade 1, for pupils initially classified as poor and non-poor readers? And are pupils tied to one category, i.e. poor reader or non-poor reader, during grade 1?

The results of this study showed that the development of word recognition in grade 1 is roughly linear and that poor and non-poor readers show a parallel development of word recognition. In addition, pupils are not tied to the classification made at the beginning of grade 1. About 40% of the poor readers in November manage to become non-poor readers in May, and about 10% of the non-poor readers in November become poor readers in May. Therefore, it is very important that in grade 1 word recognition skills are measured frequently in order to adapt the amount and type of reading instruction to the pupils reading skills.

5.1 Introduction

Teaching children to read is a complex task. The purpose of beginning reading instruction is to help children master the many challenges of the written word, including knowledge of the alphabetic system, the ability to decode new words, a vocabulary that allows words to be read at sight, and the ability to construct, integrate and remember the meanings of words in text. The purpose of this study was to determine the development of first grade pupils with regard to an important aspect of beginning reading, namely word recognition. Word recognition is the process of converting a sequence of letters into sounds for the identification of a word. This decoding process can initially be quite slow. As words are encountered more and more frequently, word recognition becomes more automatic and thereby faster. Specific patterns of letters, morphemes, and words get directly linked to their representations in the mental lexicon (Coltheart, 1978). As a result, the initially arduous process of word recognition gives way to smooth reading and comprehension.

5.1.1 Beginning reading instruction

Beginning reading instruction in the Netherlands starts in grade 1, around the age of six years. In the first half of grade 1 all of the regular grapheme-phoneme correspondences are taught. The instructions are focused on the reading of monosyllabic words of the type Consonant Vowel (CV), Vowel Consonant (VC) and Consonant Vowel Consonant (CVC) (Verhoeven & Aarnoutse, 1999). The second half of grade 1 is focused on the automatization of word recognition skills. During this period, the pupils are taught via repeated practice to recognise words more directly in memory and as a result reading becomes faster and more

fluent (Gentry, 2006). At the end of grade 1, pupils can read phonologically regular words of the type CCVC, CVCC and CCVCC; they can read short words with different spelling patterns and also polysyllabic words; they can call upon a wide variety of word-identification techniques and they can automatically recognize many words (Verhoeven & Aarnoutse, 1999).

5.1.2 Beginning reading programme Learning to read safely

Primary school teachers can call upon different programmes to teach reading. The most frequently used programme for beginning reading instruction in the Netherlands is Learning to read Safely (Mommers, Verhoeven, Koekebacker, van der Linden, Stegeman, & Warnaar, 2003). This programme is largely based on phonics, but emphasizes the entire language experience. During the first half of grade 1, the focus of this programme is the three steps in 'the fundamental reading operation': 1) linking graphemes to phonemes in the direction of reading and thus from left to right, which requires the visual analysis of graphemes, linking them to phonemes, and remembering phonemes in sequence; 2) auditory synthesis or the merging of phonemes; and 3) assignment of meaning. After this, during the second half of grade 1, more complex and varied texts are offered and thereby increase the pupils reading fluently.

In the Learning to Read Safely programme a distinction is made between readers, average readers, and poor readers, based on the pupil's reading skills. Each group is given different activities adapted to their own reading level, next to the common activities for all groups. Pupils who are able to read a simple book at the beginning of the year are classified as readers. They are given extra difficult activities that stimulate reading comprehension. Pupils who are able to follow the instructions without serious problems are classified as average readers. They are only participating in the common reading activities. Pupils who have serious problems with reading are classified as poor readers. They are receiving additional guidance of the teacher. The aim of the differentiation is to adapt the instructions to the reading skills in order to keep all pupils motivated and to minimize the number of pupils who fall behind.

5.1.3 Differences in reading development

The performance of poor versus non-poor readers has been studied for decades (e.g. Aarnoutse, Mommers, Smits, & Van Leeuwe, 1986; Jacobson, 1999; Leppänen, Niemi, Aunola, & Nurmi, 2004). The development of poor versus non-poor readers can theoretically be divergent, convergent, or parallel. These developmental trajectories should be interpreted under the condition that poor readers start at a lower level than non-poor readers. When non-poor readers improve their reading skills at faster pace than poor readers do, the development is divergent. This type of development process has been called the Matthew effect (Merton, 1968), also known as the rich getting richer and the poor get-

ting poorer. Bast and Reitsma (1998) found evidence for a Matthew effect. Morgan, Farkas and Hibel (2008) found evidence for a one-sided Matthew effect where the poor became poorer. The opposite is when poor readers improve their reading skills at a faster pace than non-poor readers do; in that case the development is convergent. Parrila, Aunola, Leskinen, Nurmi and Kirby (2005), and Aarnoutse, van Leeuwe, Voeten and Oud (2001) found support for a convergent development. Finally, when poor and non-poor readers develop at the same pace, then the development is parallel. Verhoeven and van Leeuwe (2003) and Aarnoutse and van Leeuwe (2000) found support for a parallel development. The common theme in all these studies is that they deal with the development across several grades. In contrast, we will study the development of poor and non-poor readers in grade 1.

The idea of a common instruction program together with an instruction program adapted for readers, average readers, and poor readers, is to minimize the number of pupils who fall behind and to keep all pupils motivated. The assignment of pupils to a specific category is very important for the individual pupil, because it will determine the amount and type of reading instruction. In the Learning to Read Safely programme the teacher decides to which category a pupil is assigned. At the beginning of the year, this decision is largely based on the performance on kindergarten tests and the readings skills. Juel (1988) and Smith (1997) found support for the stability of early classification. This finding paints a rather grim outlook for poor readers, as they have little chance on becoming a non-poor reader. However, not all scholars share this perspective. In a study covering 6 grades of primary school, Phillips, Norris, Osmond and Maynard, (2002) found that reading categories are more porous than might be concluded from other evidence (Juel, 1988; Smith, 1997). Lohman and Korb (2006) discuss several reasons why pupils are likely to be incorrectly classified, e.g. due to measurement errors, differential growth rates, and changes in the norming population.

5.1.4 Research questions

Reading development is commonly studied across several grades. We are, however, interested in the development of early reading, i.e. reading development in grade 1. During the first half of grade 1 reading instruction focuses on learning to decode words and to read short sentences, while in the second half of grade 1 the focus is on learning to read more complex words and on automatization of reading. Our first research question therefore is: How does word recognition develop during grade 1? Our second research question is: How does word recognition develop during grade 1, for pupils initially classified as poor and non-poor readers? There exists a controversy about the stability of the classification. Some scholars believe that pupils who are classified as poor readers always remain poor readers, or at least always belong to the lower percentiles of reading skills. Other scholars contradict this idea. Therefore, our third research question is: Are pupils tied to one category, i.e. poor readers or non-poor readers, during grade 1?

5.2 Methods

5.2.1 Participants

Our participants were 178 pupils from 8 Catholic schools located near the city of Utrecht in the Netherlands. The average age of the pupils at the start of grade 1 (August) was 6 years and 6 months (SD = 5.1 months). The sample consisted of 51% boys and 39% pupils from minority groups, mostly Turkish or Moroccan. The schools in our sample were asked to participate by college pupils who were doing their internships at those schools. All schools adopted the Learning to read safely programme (Mommers et al., 2003).

5.2.2 Measures

To study the development of the pupil's word recognition skills, we measured the speed of word recognition (WR). This is a measure of the pupil's ability to decode printed words (Aarnoutse & Kapinga, 2007). The pupil is presented a card with a list of 100 words of increasing difficulty. The unrelated words range from single syllable words to multi-syllabic words. The child is asked to read the words aloud as quickly and accurately as he or she can but without pressure. The test score is the number of words read correctly in 90 seconds. The test-retest correlation is .86 (Aarnoutse & Kapinga, 2007).

5.2.3 Procedure

We administered word recognition (WR) tests 4 times (November, January, March and May) in grade 1. The college pupils administered the WR-tests as part of their internship. They received special training sessions on how to administer these WR-tests. All WR-tests were administered with individual pupils in a separate and quiet room. All tests were done simultaneously (in the same week) on all 8 schools.

5.2.4 Data analyses

The data were analyzed in IBM statistics version 20. We used Repeated Measures ANOVA to analyse the development of word recognition of all pupils (N=178), as well as the development of poor readers versus non-poor readers. Pupils were classified as poor or non-poor readers using the WR-test of November. For the classification we used the percentile distribution in the manual (Aarnoutse & Kapinga, 2007). We choose the 20th percentile to create the groups of poor and non-poor readers. According to the manual the cut-off score for the November test is 15 words. Applying this cut-off score to our sample resulted in 18.5% poor readers (n=33) and 81.5% non-poor readers (n=145).

In order to analyse whether pupils are tied to one category during grade 1, we compared the classification in November with the classifications in January, March and May. Any disagreement in the consecutive classifications indicates that pupils are not tied to their classification, and thus indicates a classification error at January, March, or May compared to November. The classifications of January, March, or May are also based on a cut-off at the 20th percentile according to the manual; however, because pupils improve over time the actual cut-off scores will be higher. According to the manual, the cut-off score in January is 23 words, in March it is 35 words, and in May it is 44 words. If we apply these cut-off scores to our sample it results in 24.7% poor readers ($n=44$) in January, in 15.2% poor readers ($n=27$) in March, and in 19.1% poor readers ($n=34$) in May.

5.3 Results

Table 1 shows the univariate statistics of the WR tests for all pupils, as well as for the poor and non-poor readers. The average scores of the WR tests increase over time, for all pupils as well as for poor and non-poor readers. The biggest improvement occurs between January and March.

Table 1: Univariate statistics (mean (SD)) of the WR tests.

	All pupils ($n=178$)		Poor readers ($n=33$)		Non-poor readers ($n=145$)	
WR1 (November)	28.03	(16.89)	11.70	(2.35)	31.75	(16.57)
WR2 (January)	39.52	(19.59)	19.15	(6.27)	44.15	(18.60)
WR3 (March)	54.51	(19.75)	31.30	(10.29)	59.79	(17.45)
WR4 (May)	63.65	(20.70)	40.70	(16.32)	68.87	(17.86)

Figure 1 shows that the development of Word Recognition for all pupils ($n=178$) is approximately linear during grade 1. After each WR test (approximately 8 weeks in between), the scores increase roughly 10 words. A repeated measures ANOVA was performed to test whether there is a significant improvement over time. Mauchly's test indicated that the assumption of sphericity had been violated, $\text{CHI}2(5)=197.02$, $p = .00$, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity (epsilon = .58). The results show that time has a significant effect on the WR test scores, $F(1.74, 308.33) = 733.16$, $p = .00$. We tested the significance of the improvement between successive measures using the contrast: repeated. The difference between November and January was significant ($F(1, 177) = 308.73$, $p = .00$), the difference between January and March was

significant ($F(1, 177) = 587.89, p = .00$), and the difference between March and May was also significant ($F(1, 177) = 249.37, p = .00$).

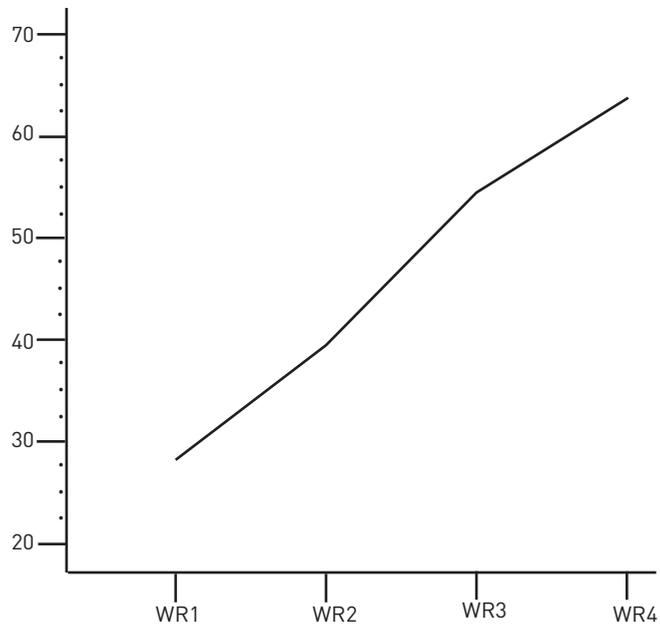


Figure 1: Development of word recognition for all pupils

Figure 2 shows that the development of Word Recognition for poor ($n=33$) and non-poor readers ($n=145$) is also approximately linear. After each WR test the number of words read in 90 seconds increases about 10 words. We performed a repeated measures ANOVA with poor/non-poor as a between groups factor, to test whether the development of poor and non-poor readers is convergent, divergent, or parallel. If the effect of time is moderated by poor/non-poor, then the development is either convergent or divergent. The factor poor/non-poor was created from the score on WR1, therefore the variable WR1 is not included in the analysis. Mauchly's test indicated that the assumption of sphericity had been violated, $\text{CHI}^2(5)=88.73, p = .00$, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\epsilon = .72$). The results show that time has a significant effect on the WR test scores, $F(1.43, 251.84) = 315.88, p = .00$. Furthermore, the effect of time is not moderated by poor/non-poor, $F(1.43, 251.84) = 2.16, p = .13$. This implies that the development is the same for poor and non-poor readers.

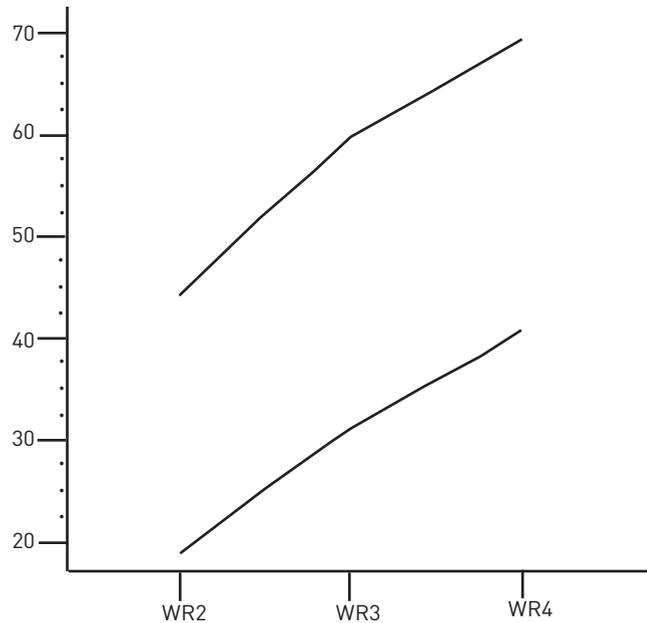


Figure 2: Development of word recognition for poor (bottom line) and non-poor readers (top line)

In the previous analyses pupils were tied to their classification during grade 1. In order to test whether they are tied, we have analyzed the transitions (over time) from poor to non-poor and the other way around. We created additional classifications for January, March, and May that indicate whether a pupil is a poor or a non-poor reader at that moment. The cut-off scores mentioned in the 'data analysis' section were used to determine who's a poor reader and who's not. Table 2 shows the transitions of poor to non-poor and of non-poor to poor between November and January, between November and March, and between November and May. If pupils would be tied to their classification, all off-diagonal cells would be empty. However, these off-diagonal cells are not empty indicating that pupils are not tied to their classification during the year. That is classification errors are made when the November classification is used in January, March, and May. It is possible to compute the percentage of classification errors by dividing the number of 'incorrectly' classified pupils by the total number of pupils. This results in 15.2% classification errors between WR1 and WR2, 11.2% between WR1 and WR3, and 16.3% between WR1 and WR4. We used the CHI2 test to test whether the observed frequencies deviate from the expected frequencies under the condition that no classification errors are made. However, because that would result in zero frequencies for the off-diagonal cells – resulting in a division by zero - we used an expected frequency of 1 for those cells. It follows that the number of classification errors from November to January is significant ($\text{CHI}2(1) = 377.43, p < .00$), from November to March

is also significant ($\text{CHI}2(1) = 185.46, p < .00$), and from November to May is also significant ($\text{CHI}2(1) = 372.49, p < .00$).

Table 2: Transition of poor and non-poor readers from November to May.

	January		March		May	
	Poor	Non-poor	Poor	Non-poor	Poor	Non-poor
November						
Poor (n=33)	75.8%	24.2%	60.6%	39.4%	57.6%	42.4%
Non-poor (n=145)	13.1%	86.9%	4.8%	95.2%	10.3%	89.7%

5.4 Conclusions and Discussion

5.4.1 Conclusions

The development of word recognition in grade 1 is roughly linear. At average, pupils almost read twice as fast in May compared to November. This is quite a remarkable improvement.

We studied the development of word recognition for poor and non-poor readers. Pupils were classified as a poor reader when their score on word recognition was 15 or lower in November; other pupils were classified as non-poor readers. This cut-off value was taken from the manual (Aarnoutse & Kapinga, 2007) of the Word Recognition test and represented the 20th percentile. The results indicated that poor and non-poor readers develop at the same pace from January to May. We, therefore, conclude that poor and non-poor readers show a parallel development of word recognition. By definition, poor readers start at a lower level than non-poor readers. This paints a rather grim outlook for poor readers; they appear to be tied.

We studied whether pupils are indeed tied to their classification by inspecting the transitions between the November classification and successive classifications in January, March, and May. Based on the November classification 15.2% of the pupils are classified incorrectly in January, 11.2% are classified incorrectly in March, and 16.3% are classified incorrectly in May. This leads us to the conclusion that pupils are not tied to the classification made at the beginning of grade 1, i.e. after 10, weeks. About 40% of the poor readers in November manage to become non-poor readers in May, and about 10% of the non-poor readers in November become poor readers in May. The number of poor readers that become non-poor readers and the number of non-poor readers that become poor readers actually cancel each other out.

5.4.2 Discussion

There are many studies on the development of word recognition during primary school (e.g. Aarnoutse, Mommers, Smits, & van Leeuwe, 1986; Morgan, Farkas, & Hibel, 2008). To our knowledge our research is unique in that we study the development of reading in grade 1. At the beginning of grade 1, most pupils start with knowledge of the alphabet, but are not able to read. Our results show that after about 10 weeks, the average pupil reads almost 30 words in 90 seconds and 24 weeks later they read at average twice as fast. This is a remarkable achievement by both pupils and teachers. However, we all know that the average pupil does not exist and that pupils develop differently.

In our study we made a distinction between poor readers and non-poor readers. Poor readers can read at average almost 12 words in November (after about 10 weeks). Non-poor readers already read approximately 3 times as fast in the same period, almost 32 words. Despite this big difference, poor and non-poor readers develop at the same pace between January and May. The causes for this parallel development are, however, unclear. Given the large difference in reading skills in November it is odd that the difference doesn't increase more after November. One of the causes might be the instruction programme Learning to Read Safely that is used in the schools in our study. Teachers are trained to adapt their instruction to individual differences, with a special attention to poor readers. This extra attention might prevent the poor readers from further falling behind, although it could also result in poor readers to catch up with the non-poor readers. At the same time, the readers in the non-poor readers category are given extra difficult activities that should stimulate reading comprehension. This could potentially speed up reading development for the (non-poor) readers. This extra practice might result in accelerated development of word recognition of the (non-poor) readers. It seems that the sum of these processes results in a parallel development of word recognition between January and May for poor and non-poor readers.

In a study covering 6 grades of primary school, Phillips et al. (2002) found that reading categories are more porous than might be concluded from other evidence (Juel, 1988; Smith, 1997). We have found support for this finding, because we observed significant classification errors, i.e. pupils classified as poor readers in November have a large chance of becoming non-poor readers during the remainder of grade 1, while non-poor readers have a small chance of becoming poor readers in the same period. Lohman and Korb (2006) discuss several reasons for these classification errors. First there is the possibility that pupils develop their reading skills at a different pace, which among others could be attributed to the specific instructions that are part of the Learning to Read Safely programme, without ignoring the innate differences that exist between pupils. Another explanation that is often mentioned in the context of classification is regression towards the mean (Campbell & Stanley, 1963; Furby, 1973). This process refers to the effect of measurement error, or unreliability, in the word recognition scores. As a result one can expect that some pupils are incorrectly classified as poor readers, while they should have been classified as non-poor readers, and vice versa.

5.4.3 Limitations

The results of this study are that poor and non-poor readers in grade 1 develop at the same pace between January and May. This result is in agreement with the findings of Verhoeven and van Leeuwe (2003) and Aarnoutse and van Leeuwe (2000). Nevertheless, this result does not exclude the possibility that during the remainder of primary school there is a convergent development (Parrila, et al., 2005; Aarnoutse, et al. 2001), or a divergent development (Bast & Reitsma, 1998; Morgan, et al., 2008).

This study would have benefited from an early measure of word recognition, e.g. in the first week of grade 1. We could have used that early measure to classify pupils as poor or non-poor readers. This would enable us to study the development of poor and non-poor readers in grade 1 using 4 measures instead of 3. With the current measures we can only cover the development of poor and non-poor readers in the second half of grade 1. There is, however, a practical difficulty in using a WR-test at the very beginning of grade one, only very little pupils are able to read a few words at the beginning of grade 1. As a result, we would end up with a large group of pupils who are poor readers. Unfortunately, we cannot tell exactly what the results of a very early measure of word recognition would be, because we don't have one. We suggest including such a measure in a future study.

5.4.4 Practical implications

Our results indicated that pupils are not tied to the classification (poor/non-poor) made after about 10 weeks (November) in grade 1. We showed that at the end of grade 1, more than 40% of the poor readers have become non-poor readers, and more than 10% of the non-poor readers have become poor readers. A fair amount of pupils are crossing the boundaries from poor to non-poor readers and vice versa. When the teacher sticks to the initial classification during grade 1 this system doesn't work, because some pupils are not receiving the attention they deserve. In order to overcome this issue, it is important to measure word recognition skills frequently and to take appropriate action on the outcome. In the Netherlands, many schools are already adopting this idea. These schools write an action plan for a class every three months. This action plan is among other things based on observations in the classroom and tests that measure readings skills. Based on these observations and reading skill tests each pupil is classified as either an: excellent reader, average reader, or a poor reader. In the action plan goals are formulated for each group and also how to achieve those goals. At the end of every three-month period the achievements are evaluated and pupils are classified again as an excellent reader, average reader, or a poor reader. In addition, new goals are set. This way of repeatedly evaluating achievements and repeated adaptation of action plans prevents classifications errors, and thus to prevent that pupils receive instructions that are not adapted to individual differences.