Differences between strong and weak readers on response times, accuracy, and complexity measures: An explorative study

Ingrid P. J. Voermans
Radboud University Nijmegen, Learning and Plasticity

Abstract
In this study participated 18 weak readers and 21 strong readers from the third grade of Dutch primary schools with respect to the acquirement of reading skills. A dynamic point of view was taken, and more specifically: the phonologic coherence model. The children started with a pretest, then followed by a training of ten sessions. Afterwards, a transfer test was taken. Each session, participants had to read 50 consonant-vowel-consonant words from a computer screen, both existing words and pseudowords. Differences between strong and weak readers were analyzed, as were differences between existing words and pseudowords, as well as progress over time. A more explorative question was whether or not differences would be present on measures of complexity. Results indicated that strong readers read faster than weak readers and they did so making fewer errors in reading both existing and pseudowords. Both strong readers and weak readers read existing words faster and more accurate than pseudowords. The children did not make any progress during and after the training. No differences on measures of complexity were found.
Keywords: phonologic coherence hypothesis, novice readers, reading problems, reading fluency, recurrence quantification analysis

In our modern society, we are surrounded by written language. From the menu in a restaurant to the instruction leaflet for medication: Lots of information is written. In education, reading therefore plays a major role (Wentink, Hoogenboom, & Cox, 2009). For being able to *learn*, children have to be able to *read*. For instance, school subjects like history or geography lessons are hard if one has trouble reading. To put it rather mildly, it is quite a struggle to live daily life, if one is unable to read well (Hellendoorn & Ruijssenaars, 2000).

In the current study, a dynamic approach was chosen to investigate the reading skills of novice readers. First, an example of an often cited (not dynamic) theory on reading development will be offered, namely the theory of Chall (1983 and 1996b, as cited in Kuhn & Stahl, 2003). Then, an overview will be given of the problems of theories that describe the process of learning to read in hierarchical stages, as the theory of Chall does. After that, a possible solution to these problems will be offered: The phonologic coherence hypothesis (Van Orden, Pennington, & Stone, 1990).

How do Children Learn to Read?

Different theories exist about how children acquire reading skills. A theory often cited is the theory of Chall (1996b, as cited in Kuhn & Stahl, 2003). According to Chall, the process of learning to read can be divided into six different stages, with the first stage already starting before children start school (Chall, 1983 and 1996b, as cited in both Blok, Oostdam, Otter, & Overmaat, 2002 and Kuhn & Stahl, 2003). In this first stage, called *Stage 0*, *pre-reading*, *early reading*, or *emergent literacy*, the child becomes familiar with oral and written language in his or her daily environment. For instance, the child learns that when you read, you should start at the top of the page and move from the left to the right. When the child gets formal reading instructions, a transition to the next stage, *Stage 1*, or the *initial stage of conventional literacy*, takes place, in which the child learns to decode words. The child learns what the letters look like and which graphemes and phonemes belong together. *Stage 2,* also called *confirmation and fluency* or *ungluing from print*, consists of the acquirement of fluency in decoding and therefore in reading. The child learns to automatically recognize words often used. According to Chall, this is an important step, because it requires less space in working memory. Free space can be used to focus on the content of a text. Thus, to be able to focus on the meaning of texts, fluency is a prerequisite. When spending a lot of energy on decoding, the child cannot focus on the content, and therefore it cannot learn from a text. When a child is not able to learn from a text, the child will have problems in both its school career and daily life. To acquire fluency, practice is required. With frequent reading of sentences and texts, recognition, fluency, and comprehension improve (Adams, 1990 as cited in Blok et al., 2002). Feedback on the words read might also stimulate the process of learning to read. Meta-analyses (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Schimmel, 1983 as cited in Bangert-Drowns et al., 1991) reveal that learning improves, when a child is told whether or not a word is read correctly and learning improves particularly when feedback is offered immediately after reading a word (Kulik & Kulik, 1988). The amount of information provided during feedback does not seem to make any difference, probably because children are most interested in the correct answer (Bangert-Drowns et al., 1991). As a consequence, they might not listen to any further explanation afterwards. The second stage of Chall’s theory appears to be the most important and will therefore be the focus of the current research. For completeness the remaining stages of Chall’s theory will be discussed briefly. In *Stage 3, reading for learning the new,* the child has to read many texts (Kuhn & Stahl, 2003) and a transition is made from reading for fun to reading to learn. In *Stage 4*, called *multiple viewpoints*, the child reads texts of different sources about one subject and learns to critically evaluate these texts. In the final stage, *Stage 5*, *construction and reconstruction*, the child learns to compare different sources in order to form an own opinion.

Predictors for Reading Fluency

The main predictors for reading fluency are *phonological awareness*, *letter knowledge,* and *speech abilities* (Wentink et al., 2009). *Phonological awareness* is the awareness of sound structures in spoken language. A child with a well-developed phonological awareness knows for instance that the words ‘hat’ and ‘her’ start with the same phoneme /h/. Also, the child is aware of the fact that ‘hat’ and ‘cat’ rhyme. Another aspect of phonologic awareness is that the child knows that, if the phoneme /p/ is removed from ‘pair’, the sound /air/ remains. Hence, the child is aware that language consists of phonemes and these phonemes are represented by graphemes (Liberman, 1971 as cited in Blachman, 2000). For the awareness of the relation between phonemes and graphemes, *letter knowledge* is necessary: A child has to know which phonemes and graphemes belong together. In other words: Which sound is represented by which letter and vice versa (Wentink et al., 2009)? When letter knowledge has developed, a child is able to gain speed in reading. Last, *speech abilities* are important, because of the relation between the amount of speech and the vocabulary of a child. The more language a child produces, the more extended its vocabulary becomes. In the first years of reading education, a child is taught to read high-frequency words (Van Weerdenburg, Verhoeven, Van Balkom, & Bosman, 2009). Lexical retrieval of words the child is familiar with, that is, recall words from the memory, is faster than decoding new words.

Differences between Strong and Weak Readers

In contrast to the English language, the Dutch language has a relatively transparent orthography (Verhoeven & Van Leeuwe, 2009). In (relatively) transparent orthographies, the differences in the level of reading skills of novice readers reveal themselves mostly in speed, rather than in accuracy (Davies, Rodríguez-Ferreiro, Suárez, & Cuetos, 2013; Suárez-Coalla & Cuetos, 2012; Wimmer, 1993). The finding that weak readers read slower than strong readers, but with the same amount of errors, holds for existing words, and for pseudowords alike (Wimmer, 1993). For pseudowords, the difference in speed between strong and weak readers is even larger.

A great deal of research has been done to unravel why some children struggle to acquire sufficient reading skills (for meta-analyses see, for example, Fletcher, 2009; Reid, Fawcett, Manis, & Siegel, 2008; Vellutino, Fletcher, Snowling, & Scanlon, 2004). In their meta-analysis, Vellutino et al. (2004) discuss several suggestions. Compared to strong readers, weak readers are supposed to have less developed phonological knowledge, less developed verbal memory, and more difficulties in the retrieval of knowledge from memory. Other suggestions are that weak readers have smaller temporoparietal brain areas or differences in the size of the corpus callosum. Despite the suggestions, no unequivocal conclusion has been drawn (Vellutino et al., 2004). This cannot be a surprise, because learning to read is a multi-sensory process (Blomert, 2011). Moreover, to pronounce a single utterance, around seventy muscles have to work together (Turvey, 2007). As a consequence, the problems in achieving reading skills might be better explained by the *interdependence* of the cognitive components involved in reading than by the components themselves (Van Orden & Holden, 2002; Rueckl, 2002; Van Orden & Kloos, 2003). Assuming interdependence, a more adequate approach to reading would be to take a dynamic stance, in which the relationships among components are studied, rather than the components themselves. After years of searching, yet not finding, *the* one factor that explains all reading problems, the time has come to take a new course: Explore the reading process from a dynamic point of view (Van Orden, Jansen op de Haar, Bosman, 2010).

Why a dynamic perspective?

Viewing reading from a dynamic perspective has several advantages. First, the perspective is focused on development (Bosman & Van Orden, 2003). As a consequence, it offers the possibility of change. A dynamic theory on reading might be able to explain both how children learn to read and why some children struggle. However, the struggling children are not doomed to have reading problems for the rest of their lives. Viewing reading from a dynamic perspective offers possibilities to acquire reading fluency, even after a slow start.

Second, most other perspectives, including the theory of Chall (1983 and 1996b, as cited in Kuhn & Stahl, 2003), describe learning processes in hierarchical stages, but it remains unexplained how to get from one stage to the next. In a dynamic perspective, there are no hierarchical stages, yet states are important. A state can be stable or instable (Smith & Thelen, 2003). When someone is *acquiring* a certain skill, for instance reading a word, the state the child is in, is instable, while it is changing, for example, from sounding out letter by letter towards a more integrated reading of the word. The more acquired the skill, the more stable the system becomes. The stability of a system’s state can be deduced from measures of complexity, for example *recurrence rate*, *determinism*, *entropy*, and *meanline* (Wijnants, Hasselman, Cox, Bosman, & Van Orden, 2012). These will be explained in detail in the Method section. In short, the process of getting from one stage to the next can be described when taking a dynamic perspective on development.

Third, the dynamic approach has yielded promising results to explain the development of children on several domains, for instance language development (Bassano & Van Geert, 2007), socio-emotional development (Lewis, Zimmerman, Hollenstein, & Lamey, 2004), cognitive development, and motor development (Smith & Thelen, 2003). Perhaps a dynamic perspective might also make a valuable contribution to research on reading problems. The phonologic coherence hypothesis is such a dynamic model on reading (Bosman & Van Orden, 2003).

Phonologic Coherence Hypothesis

The phonologic coherence hypothesis views reading from a dynamic system approach (Van Orden et al., 1990). According to this hypothesis, reading can be modelled as a recurrent network (Bosman & Van Orden, 2003). The phonologic coherence hypothesis stems from neural network theory or connectionism. The metaphor used in network theory is the brain. The neurons are seen as nodes. All nodes are interconnected and are able to activate or inhibit each other.

In the phonologic coherence hypothesis it is assumed that three node families exist: Letter nodes, phoneme nodes, and semantic nodes (Bosman & Van Orden, 2003; Van Orden et al., 2010). This means that all letter nodes form one family, all phoneme nodes form one family, and all semantic nodes form one family. All nodes have a bidirectional connection with all nodes of their own family and with all nodes of the other families. The nodes of the same family inhibit each other. Outside their own respective families the nodes activate each other. For instance, when a letter node gets activated, it inhibits other letter nodes, but activates phoneme nodes and semantic nodes.

Within the phonologic coherence model, a child is seen as a system. When a written word is offered to the system (as is the case with reading), the letter nodes get activated (Bosman & Van Orden, 2003). These letter nodes activate the phoneme nodes and the semantic nodes. Next, the phoneme nodes activate the letter nodes and the semantic nodes again. The semantic nodes activate also both the letter nodes and the phoneme nodes. Subsequently the letter nodes activate the phoneme nodes and semantic nodes again, forming feedback loops (see Figure 1). These loops are temporary stable and result in a coherent dynamic whole, called *resonance* (Van Orden et al., 2010). Resonance helps to stabilize the state a system is in, while it strengthens the connections between the activated nodes. When resonance is obtained, the network has built a phonologic interpretation and/or a meaningful interpretation of the offered word. The order in which the nodes are activated, depends on the context. In reading, the letter nodes are activated first, whereas in spelling, the phoneme nodes are.

Differences exist between the strength of the connections between the three node families (Van Orden et al., 2010). In typically developing readers, the connections between letter nodes and phoneme nodes are the strongest, followed by the connections between the phoneme nodes and semantic nodes. The weakest connections exist between the letter nodes and semantic nodes. The strong connections between the letter nodes and phoneme nodes have evolved from the rather consistent relation between letters and phonemes in transparent orthographies, like Dutch. The letter L, for instance, is always pronounced as /l/ and the phoneme /l/ is always written with the letter L. However, a phoneme or a letter gives little information about the meaning of the word, resulting in a less strong connection with the semantic nodes.

Reading a word will activate both correct and incorrect letter nodes and phoneme nodes (Van Orden et al., 2010). Reading for instance the word HI, will activate the letter nodes H_1 and I_2 (the subscript refers to the position of the letter in the word), and subsequently the correct phoneme nodes /h_1/ and /$a^I{}_2$/, but the letter node I will also activate the incorrect phoneme node /i/ (as in HIT). The activated phoneme nodes will activate the correct letter nodes H and I, but also the incorrect Y, because the phoneme /$a^I{}_2$/ can also be written with the letter Y, as in MY. All incorrect nodes have to be inhibited, before the word HI can be read correctly. In spelling, the same process becomes activated, but instead of starting with activation of the letter nodes, the process starts with activation of the phoneme nodes.

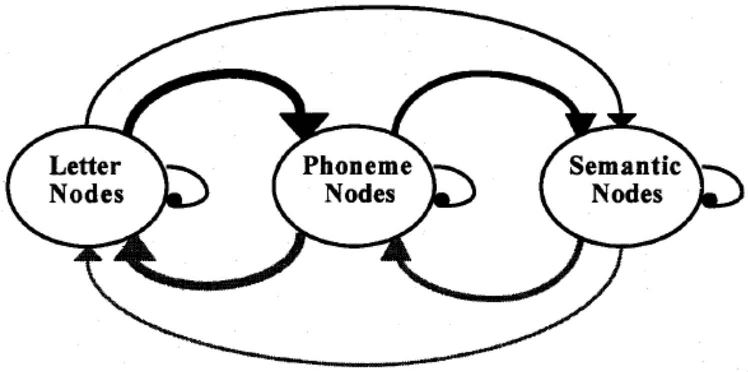

Figure 1. The feedback loops between the node families. The lines ending in an arrow indicate activation of other nodes, while the lines ending in a dot represent inhibition. The boldness of the lines indicates the strength of the connection: The bolder the line, the stronger the connection (Van Orden et al., 2010, pp. 141).

Now it is explained how the network works with reading, it is just a little step to understand how this network comes into being. As in the brain, a strong connection arises between cooperative neurons, the connection between letters and phonemes becomes stronger after frequently offering a letter with the corresponding sound (Bosman & Van Orden, 2003). The connection between the letter M and the phoneme /m/ strengthens, because the M is always pronounced as a /m/ and the /m/ is written with the M. At the same time, the connection between the /m/ and for instance the letter E becomes stronger than the connection between the /m/ and the letter Q, because it is more probable the /m/ is followed by an E than by a Q. When the network has been formed sufficiently (i.e., relations are formed between the phoneme nodes, the letter nodes, and the semantic nodes), the network becomes able to read new words and even pseudowords.

In contrast to other theories on the process of learning to read, for instance the theory of Chall (1983 and 1996b, as cited in Kuhn & Stahl, 2003), suggestions for reading instruction can be deduced from the phonologic coherence model (Bosman & Van Orden, 2003). This is a huge advantage, because the model has an immediate practical application. Obviously, phonology is important in instruction, because phonology is both a predictor for reading fluency and the most important factor in the process of learning to read. Another suggestion for reading instruction is placing the learning process in a meaningful context. After all, the model reveals the fundamental interconnectedness among phonemes, letters, and semantics. When offering a meaningful context, children have more information that they can use. Also, in reading instruction the whole word should be offered to the child, because relations are stronger between often cooperating nodes (see the example above: The /m/ is more likely to be followed by E than by Q). Next, words offered earlier to the system (the child in this case), influence the processing of newly offered words. As a consequence, it should be considered which words should be taught to the child first and which words later. Should reading start with only high-consistent words, or should certain letters be mastered first, before offering new letters? The phonologic coherence hypothesis does not give information regarding the most adequate order of presenting words, it only says that order matters. Empirical research is necessary to establish this order for each different language (Bosman & Van Orden, 2003). A final suggestion for reading instruction is to be aware of the manner in which the stimulus (the word) is offered to the system (the child). The way a stimulus is offered to the system, influences the process within the system and as a consequence the output (the reading of the word). More specifically, this means that it is important to adapt the reading instruction to the child and its environment. For instance, one needs to be aware of specific qualities of a child, language environment and social-economical environment. Furthermore, reading instruction should also take into account collaboration between different senses (for instance, *hearing* a phoneme and simultaneously *seeing* the corresponding grapheme), while learning to read is a multi-sensory process. If the instruction is not tailored to the child, the quality of the output (reading) will suffer. Research is necessary to decide how the stimuli should be offered (Bosman & Van Orden, 2003).

Several causes of reading problems can be deduced from the phonologic coherence hypothesis (Bosman & Van Orden, 2003). It might be that for weak readers the connections between some letter nodes and phoneme nodes are not bidirectional. As a consequence, no feedback loops are formed and without these feedback loops, the connection between the letter and the phoneme does not get strengthened. The state of the system remains instable. The assumption that some connections are not bidirectional might also explain why struggling readers have troubles reading pseudowords. For existing words, the connection with the semantic nodes might provide support in reading a word, but with non-existent words, this is not possible. Another possible explanation for the problems in reading skills of some children, is an instable connection between the nodes, so the child needs more practice to create stable connections. It is also possible that problems arise because the activation of the nodes expires too rapidly.

Apart from providing possible causes of reading problems, the phonologic coherence hypothesis also offers suggestions for interventions for struggling readers. In almost all interventions, clinicians attempt to strengthen the connection between letters and phonemes (Bosman & Van Orden, 2003). According to the phonologic coherence hypothesis this type of intervention can be supplemented by another type of intervention: In addition to focus on strengthening the weak relation between phonemes and letters, intervention could also rely on and strengthen the relation between phonology and semantics. This means that intervention should take place in a meaningful context.

In short, the phonologic coherence hypothesis is a *comprehensive* model, with a focus on development, that is able to explain the development of reading skills, can describe how to get from one stage to the next, gives suggestions for reading instruction, can explain why some children struggle in achieving reading skills, and offers suggestions for intervention in struggling readers. Because all these elements can be deduced from this single model, it is a good candidate to replace the existing theories, that usually only explain elements of reading processes and reading behaviour.

The current study

As said, in our literate society, being able to read fluently is important for being able to function well. Because of this importance, research is necessary to find out how children learn to read and why for some children it is no problem at all to acquire reading skills, while others struggle so hard. What are the differences between these children? Knowing about the underlying processes, it may become clearer why and when the process stagnates, and it might become clear what the aim of an intervention should be. Intervening as soon as possible will prevent children from getting behind in learning, getting frustrated and losing their motivation for reading (Wentink et al., 2009).

In the current study, it was hypothesized, based on phonologic coherence model, that all children, both strong and weak readers, would gain speed in reading and make fewer errors, after frequent repetition, because repetition strengthens the connections between letter nodes and phoneme nodes. It was questioned whether or not differences in the development of speed and accuracy existed between strong and weak readers and between reading existing words versus pseudowords. Furthermore, it was questioned whether or not transfer would occur, from reading well-known words during training, to other, non-trained, words. The expectations were that weakly reading children would make the same number of errors as strongly reading children, yet that weak readers would be slower than strong readers. In relatively transparent orthographies, which Dutch is, differences in the level of readings skills of novice readers reveal themselves mostly in speed. For pseudowords, the difference in speed between strong and weak readers was expected to be even larger than for existing words, because the phonologic coherence model states that, while reading pseudowords, no connection with semantics could be used. Another, more explorative question, is whether or not differences would exist between strong and weak readers and between reading words and pseudowords with regard to measures of complexity.

Method

Participants

The current study is part of a larger study regarding the process of learning to read. In the Dutch school system, children are at the age of six when entering first grade, where they learn to read. In the overall study, several conditions were used. For the current study, the data of one particular condition was used. This condition consisted of 39 third-grade children of ten different village schools. All children attended regular education. The children were selected, based on their scores on a test for word decoding skills, called the *Drie-Minuten-Toets* (*DMT* [Three-Minutes-Test]; Jongen & Krom, 2009). Half of the selected children were strong readers, that is, children with the highest scores on the DMT. The other half of the children were weak readers, that is, children with the lowest scores on the DMT.

The group of weak readers consisted of eighteen children: Thirteen boys and five girls. The mean age in this group was 83 months (*SD* = 4.26). The group of strong readers consisted of 21 children, of which fourteen were boys and seven were girls. The mean age was 82.29 months (*SD* = 3.99). All children had Dutch as their first and only language.

Materials

Selection.

For the selection of children the scores on the DMT (Jongen & Krom, 2009) were used. By using the DMT, the level of word decoding skills can be determined. The DMT consists of three cards, of which the first two have to be administered halfway first grade. Both lists consist of five rows, each containing thirty words. All words on the first two cards have only one syllable. On the first card vowel-consonant (VC) words (as, oom [ashes, uncle]) are listed, CV words (fee, bui [fairy, shower]) and CVC words (vak, wit [job, white]). The second card consists of CCV words (plooi, vlaai [fold, flan]), CCVC words (smid, bloem [smith, flower]) and CVCC words (tand, jurk [tooth, dress]). Also CCVCC words (grond, vlecht [ground, braid], CCCVC words (strik, spreeuw [bow, starling]) and CVCCC words (barst, dienst [burst, service]) are presented on the second card.

Training materials.

During the training, two lists of words (List I and List II) were used. Both lists consisted of 25 existing words and 25 pseudowords. All words were CVC words and had a clear letter-sound correspondence (i.e., the pronunciation of the graphemes within the word is the same as when the graphemes are pronounced separately). The graphemes used, were: a, e, i, o, u, g, k, l, m, n, p, r, s and t. The frequency of graphemes used on List I was approximately the same as on List II. The pseudowords were deduced from the existing words, in a way that for both word types all graphemes were used approximately the same number of times. The words had to be read from a computer screen and the child wore a headset consisting of a headphone with attached microphone.

Procedure

Selection.

The DMT was administered in January or February by the participating schools. The scores were used for the selection of the children for the current study. In the DMT the child was asked to read aloud the words on the first card as accurately and as fast as possible. The child was given one minute to read the words. After this minute the test was stopped by the test leader. The same procedure was followed for the second card. The rough score was determined by subtracting the words wrongly read from the total number of words read. Using a table, the rough scores were converted to standard scores. The 25% best readers received an A-score, the next 25% get a B-score, and the following 25% had a C-score. The next 15% received a D-score and the 10% weakest readers had an E-score. The children with an A-score and the children with a D- or an E-score were selected for the current study.

Training procedure.

The training took place in March and April. It started with a pretest. During the pretest, the words of List II were used. After the pretest, ten training sessions took place, with the words of List I. Thereafter, a transfer session took place. At the transfer session, the words of List II were used again. Every session lasted around ten to fifteen minutes. The pretest, the training sessions and the transfer session took place within three weeks. One to four sessions per child per day were held, depending on the time available on a day.

Each session took place in a quiet room. The child sat in front of the computer screen, the test leader sat behind the child. During the training, the child met Tom, a puppet on the computer. Tom asked the child to read aloud the word that appeared on the screen. After reading the word, the child received feedback from Tom. He told the child whether the word was read correctly or wrongly and then the word was repeated (for instance: ‘Well done, it’s cat’). Whether a word was read correctly or wrongly was decided by the test leader. She pressed a key on the keyboard, after which the feedback followed in the headphone. Apart from judgment on the correctness of the word, a key press also registered the response time, that is, the time the child needed to read aloud the word from the moment the word had appeared on the screen.

The block of existing words and the block of pseudowords were presented each session. During one session the block of existing words was presented first followed by the block of pseudowords, during the next session, this order was reversed. No breaks took place between the two blocks of words. Also, the words within each block were offered consecutively, without any breaks in between. Within each block of words (existing or pseudo) the words were offered in a random order and also between participants the order of presentation was random. During the first session half of the participants started with reading existing words, the other half started with reading pseudowords.

Data-analysis

All analyses were executed with analyses of variance with repeated measures. The data were analyzed in three different ways: per *block of words*, per *session*, and *overall*. For the analyses per block of words (all existing words within one session versus all pseudowords within one session), differences in measures of central tendency were analyzed. The analyses were conducted for *response times* (i.e., the time a child needed to read a word) and for the *percentage correct* (i.e., the percentage words read correctly within one block of words). In the Results section, details of the analyses will be presented. In the analyses per session (all words within one session, i.e., both existing and pseudowords) and the overall-analyses (all existing words in the pretest, all training sessions, and transfer session versus all pseudowords in the pretest, all training sessions, and transfer session), measures of complexity were analyzed. To do so, Recurrence Quantification Analysis (RQA) was used.

RQA is a nonlinear technique attempting to model the repeating behavior of a system (Wijnants et al., 2012). To do this, a phase space (see Figure 2 for an example), is reconstructed, based on one single time series. The fact that this reconstruction can be based on a single time series follows from Takens' theorem (Takens, 1981). In the current study, the time series consisted of the response times a child needed to read the words. Looking at the amount of recurring patterns, it can be deduced whether the system is stable or changing (Wijnants et al., 2012). A stable system returns to one or more recurrent points in the phase space.

RQA provided information about the recurring patterns in the system, using *recurrence rate, determinism, entropy* and *meanline* (Wijnants et al., 2012). Based on the phase space, a graphical representation is made, called the *recurrence plot* (see Figure 3). This recurrence plot is used to calculate the above-mentioned RQA measures. The amount of repeating behavior in the system can be deduced from the recurrence rate. The recurrence rate is computed by dividing the number of points recurring to a recurrent point by the total number of possible recurring points. Determinism gives information about the amount of recurrent points being part of a recurring pattern. In this context, a pattern is a certain behavior that is recurring in the same order as before in the time series. In the recurrence plot this is visible as a diagonal line. To compute determinism, the number of recurrent points that

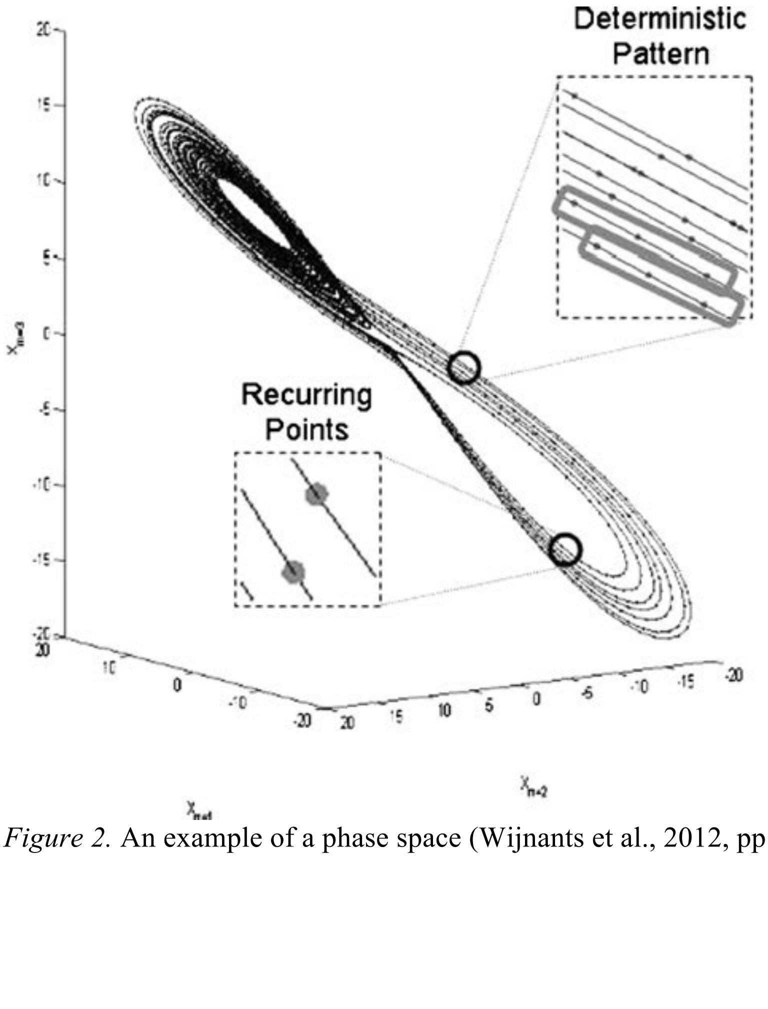

Figure 2. An example of a phase space (Wijnants et al., 2012, pp. 105).

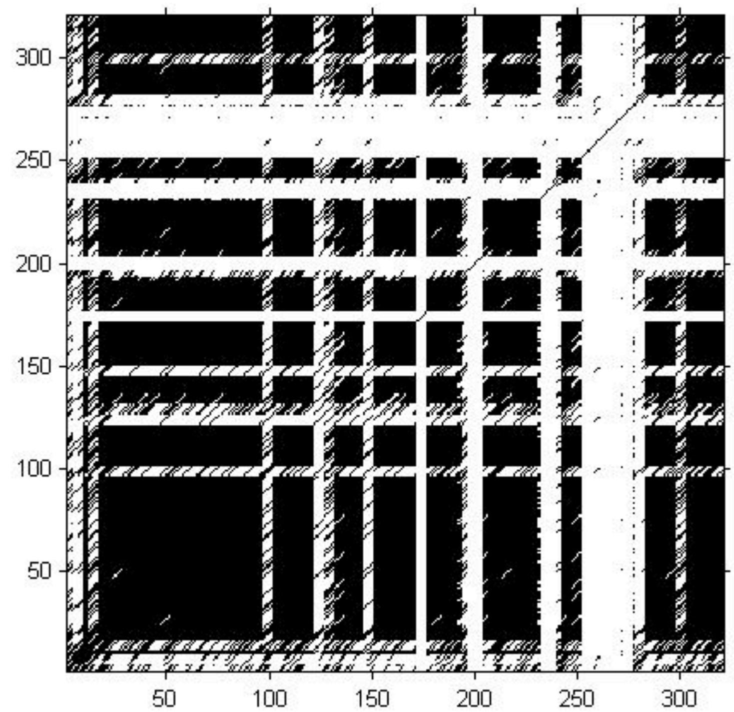

Figure 3. The recurrence plot of participant number 101 (weak reader) while reading 300 existing words.

are part of a pattern is divided by the total number of recurrent points. Entropy is a measure to determine how often certain patterns recur (Wijnants, Cox, Hasselman, Bosman, & Van Orden, 2012) and the meanline is a measure for the mean length of these recurring patterns (the mean length of the diagonal lines in the recurrence plot; Wijnants et al., 2012). High figures of recurrence rate, determinism, and meanline are indicators of a stable system, whereas a high figure for entropy indicates more chaos in the system (Lichtwarck-Aschoff, Hasselman, Cox, Pepler, & Granic, 2012).

Before starting the analyses, the *lag (t), embedding dimensions (m)* and *radius (e)* had to be defined (Wijnants et al., 2012). The *lag* was defined based on the *Average Mutual Information function (AMI)*. This function tries to predict a next data point, based on the current data point (Fraser & Swinney, 1986). In the current study a lag of 1 turned out to fit well.

After choosing the lag, the *embedding dimension* was determined, based on the *False Nearest Neighbor analysis* (Kennel, Brown, & Abarbanel, 1992). Using too few dimensions, data points seem to be close to each other in a low-dimensional phase space, although they are not in a phase space with more dimensions. As a consequence the number of recurrent points becomes too high. Incorporating too many dimensions does not bring new information into the system. In the current study an embedding dimension of 3 was chosen for the analyses per session. An embedding dimension of 5 was chosen for the overall analyses.

Finally the *radius* was chosen. The radius decides whether or not a point is called recurrent. With a radius too high all points become recurrent, with a radius too small only exact points are recognized as recurrent (Wijnants et al., 2012). In the current study the radius was set on 20% of the maximum phase space distance (Riley, Balasubramaniam, & Turvey, 1999).

Results

Before starting the analyses, all outliers (3 *SD* from the Mean) were removed, and the time series for the measures of complexity were normalized. This normalization was done by subtracting the mean and then dividing by the *SD*. Also, the assumptions for analysis of variance with repeated measures were checked. In some analyses involving the variable time, the assumption of sphericity was violated. According to Girden (1992), epsilon had to be examined in these cases. Because all epsilons appeared to be smaller than .75, the Greenhouse-Geisser correction was used. Where the assumption of sphericity was violated, it is mentioned in the description of the analysis concerned. As explained before, de measures of central tendency were analyzed per *block of words*, the measures of complexity were analyzed per *session* and *overall*. Alpha was set on .05.

Analyses per block of words

Response times.

The performance of each child was reflected in 24 time series, each consisting of 25 reaction times. Half of the time series contained the reaction times for reading existing words, the other half of the time series consisted of the reaction times for reading pseudowords. The first time series and the last time series contained the response times of the pretest and the transfer session. The intermediate ten time series were of the training sessions. Each of the time series was averaged per child. The mean reaction times for reading the words were input for a repeated measures ANOVA with reading type (strong vs. weak) as a between-subject factor, and word type (existing vs. pseudo) and time (session number) as within-subject factors. The training sessions were analyzed separately from the pretest and transfer session.

Training sessions.

In the analyses of response times in the training sessions, the assumption of sphericity was violated. Therefore, the Greenhouse-Geisser correction was used.

For the response times of the children per block of words during training, a main effect of word type was found ($F(1, 37) = 39.50$, $p < .001$; see Appendix A for the means of the analyses per block of words). Regardless of reading type, children read the pseudowords slower than the existing words. Also, a main effect of reading type appeared to be present ($F(1, 37) = 83.97$, $p <.001$). Compared to strong readers, weak readers read slower on average, regardless of word type. The interaction effect between word type and reading type was significant too ($F(1, 37) = 13.03$, $p < .01$; see Figure 4 for the interaction). The difference in response times was higher for the pseudowords (strong readers: $M = 2016$ ($SD = 202$); weak readers: $M = 4771$ ($SD = 218$)) than for the existing words (strong readers: $M = 1909$ ($SD = 188$); weak readers: $M = 4375$ ($SD = 204$)). No within-subject effect of time was present ($F < 1$), indicating that the children did not gain speed during the training in reading the words, regardless of word type. None of the tested interactions between word type and time ($F(9, 333) = 1.46$, $p = .21$), reading type and time ($F(9,333) = 1.05$, $p = .38$), and word type, reading type, and time ($F(9, 333) = 1.19$, $p = .32$) were significant.

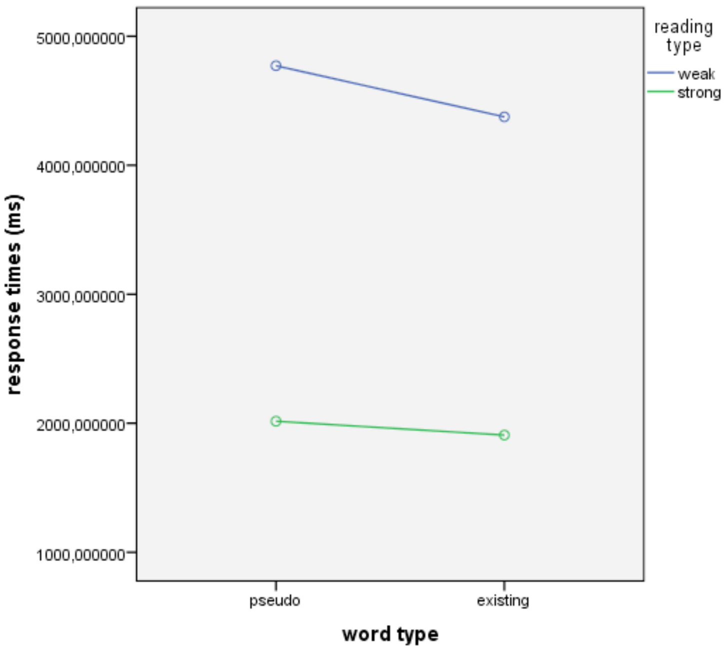

Figure 4. Interaction effect of the response times in the analyses of the training per block of words.

Pretest and transfer.

Just as in the training sessions, main effects of both word type and reading type were present in the data of the pretest and transfer session. Regardless of their reading skills, children read the existing words on average faster than the pseudowords ($F(1, 37) = 10.32$, $p < .01$). Also, strong readers were on average faster than weak readers ($F(1, 37) = 90.12$, $p < .001$), regardless of which words (existing or pseudo) they read. However, the interaction between word type and reading type was no longer significant ($F(1, 37) = 1.75$, $p = .19$), meaning that the difference in response times between strong and weak readers, did not vary across word types. A within-subject effect of time was not found either ($F(1, 37) = 2.05$, $p = .16$). In the transfer session the words were not read faster than in the pretest. No interaction effects were found between word type and time ($F(1, 37) = 2.11$, $p = .16$), reading type and time ($F(1, 37) = 2.29$, $p = .14$), and word type, reading type, and time ($F < 1$).

Percentage correct.

The percentages correct were computed for each session, but separately for existing words and pseudowords. As a consequence, the percentages correct were computed per 25 words. The percentages correct were input for a repeated measures ANOVA with reading type (strong vs. weak) as a between-subject factor, and word type (existing vs. pseudo) and time (session number) as within-subject factors. The training sessions were analyzed separately from the pretest and transfer session.

Training sessions.

In the analyses of the percentages correct, the assumption of sphericity was violated. Therefore, the Greenhouse-Geisser correction was used.

For the percentage correct read words a main effect of word type was found ($F(1, 37) = 31.28$, $p < .001$). The existing words were read with significantly fewer errors than the pseudowords, regardless of reading type. The main effect of reading type appeared to be significant too ($F(1, 37) = 23.61$, $p < .001$). Strong readers made fewer errors than weak readers, regardless of word type. An interaction effect was also found ($F(1, 37) = 17.44$, $p < .001$; see Figure 5 for this interaction effect). The difference in percentage correct was higher for pseudowords (strong readers: *M* = 98.95 (*SD* = 1.23); weak readers: *M* = 89.89 (*SD* = 1.33)) than for existing words (strong readers: *M* = 99.37 (*SD* = .98); weak readers: *M* = 92.78 (*SD* = 1.05)). Yet, no within-subject effect of time was found ($F(9, 333) = 1.18$, $p = .32$), indicating that the children did not make less errors during the training. None of the tested interactions between word type and time ($F(9, 333) = 1.17$, $p = .33$), reading type and time ($F(9,333) = 1.23$, $p = .30$), and word type, reading type, and time ($F < 1$) were found to be significant.

Pretest and transfer.

Like in the training, a main effect of word type was present in the pretest and transfer session ($F(1, 37) = 32.94$, $p < .001$). Again, while reading existing words, the children made fewer errors than while reading pseudowords, regardless of the reading skills of the child. Also, a main effect of reading type was found ($F(1, 37) = 21.97$, $p < .001$). Strong readers made fewer errors compared to weak readers, regardless of word type. An interaction effect was found too ($F(1, 37) = 16.86$, $p < .001$; see Figure 6 for this interaction effect). For the pseudowords (strong readers: *M* = 98.48 (*SD* = 1.24); weak readers: *M* = 89.11 (*SD* = 1.34)) the difference in percentage correct was larger than for the existing words (strong readers: *M* = 99.05 (*SD* = 1.10); weak readers: *M* = 92.56 (*SD* = 1.19)). There was no within-subject effect of time
($F < 1$), that is: regarding the percentage of words read correctly, no significant differences between the pretest and the transfer session were present. No interaction effects were present between word type and time ($F(1, 37) = 2.19$, $p = .15$), reading type and time ($F < 1$), and word type, reading type, and time ($F(1, 37) = 3.15$, $p = .08$).

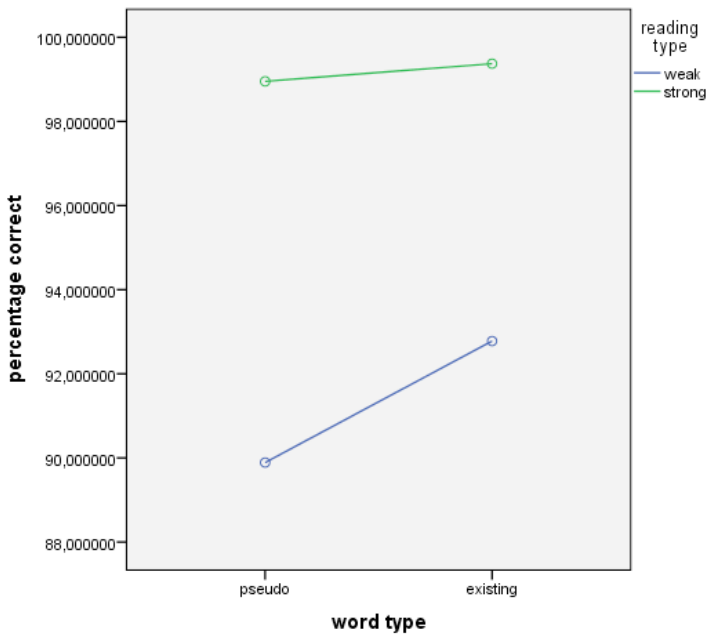

Figure 5. Interaction effect of the percentage correct in the analyses of the training per block of words.

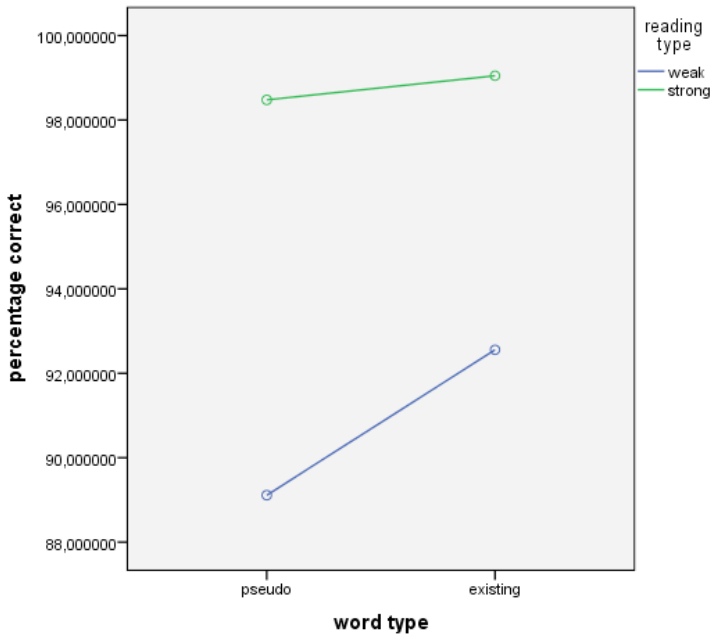

Figure 6. Interaction effect of the percentage correct in the analyses of the pretest and transfer per block of words.

Analyses per session

In the analyses per session, measures of complexity were analyzed. The performance of each child was reflected in twelve time series, each consisting of 50 reaction times. The first time series and the last time series contained the response times in the pretest and the transfer session. The intermediate ten time series were of the training sessions. On each time series measures of complexity were computed and averaged. The mean recurrence rate, determinism, entropy, and meanline were input for a repeated measures ANOVA with reading type (strong vs. weak) as a between-subject factor and time (session number) as a within-subject factor. The training sessions were analyzed separately from the pretest and transfer session.

Training sessions.

In the analyses of the measures of complexity in the training, the assumption of sphericity was violated. As a consequence, the Greenhouse-Geisser correction was used.

Looking at the recurrence rate in the training sessions, neither the main effect of reading type ($F(1, 37) = 1.75$, $p = .19$; see Appendix B for the means of the analyses per session), nor the within-subject effect of time ($F(9, 333) = 1.01$, $p = .42$), or the interaction effect between the two factors ($F(9, 333) = 1.30$, $p = .26$) reached a significant level. The same was true for determinism, entropy, and meanline (reading type, time, and their interaction, all p's $> .05$). [1]

Pretest and transfer.

Between the pretest and the transfer test, no differences were found in recurrence rate, determinism, entropy, and meanline (reading type, time and their interaction, all p's $> .05$).

[1] Different values for the parameters were tested, but this did not make any differences for the results.

Overall-analyses

In the overall-analyses, measures of complexity were analyzed. The performance of each child was reflected in two time series, both consisting of 300 reaction times. One time series contained the reaction times for reading existing words in the pretest, all training sessions, and the transfer session. The other time series consisted of the reaction times for reading pseudowords in the pretest, all training sessions, and the transfer session. Measures of complexity of each time series were computed and averaged. The mean recurrence rate, determinism, entropy, and meanline were input for a repeated measures ANOVA with reading type (strong vs. weak) as a between-subject factor and word type (existing vs. pseudo) as a within-subject factor.

For the recurrence rate, no main effect was present for either word type (existing versus pseudowords, $F < 1$; see Appendix C for the means of the overall-analyses) or for the reading type ($F(1, 37) = 1.73$, $p = .20$). However, the interaction between word type and reading type appeared to be significant ($F(1, 37) = 4.15$, $p < .05$; see Figure 7 for this interaction effect). The influence of reading type on recurrence rate is smaller for existing words (strong readers: $M = .17$ ($SD = .02$); weak readers: $M = .15$ ($SD = .02$)), than for pseudowords (strong readers: $M = .20$ ($SD = .03$); weak readers: $M = .13$ ($SD = .03$)). For determinism, entropy, and meanline, none of the results were significant (word type, reading type, and their interaction, all p's > .05).

Discussion

In the introduction, several hypotheses were formulated. The children in the study had to read consonant-vowel-consonant words. First, it was expected that the time the children would need to read these words would decrease after frequent repetition and that the children would make fewer errors. This hypothesis was not confirmed. Both strong and weak readers did not gain speed and did not make fewer errors while reading the same words during ten training sessions. Moreover, no differences existed between the pretest and the transfer session. This means that, words presented for the first time during the pretest were not read faster or with fewer errors in comparison to the second time they were presented, ten sessions later. The second expectation was that strong readers would read faster than weak readers, and that this difference would be even larger for pseudowords than for existing words. This hypothesis was confirmed by the data. However, the expectation that weak readers would make the same number of errors as strong readers has to be rejected. The weak readers made more errors while reading existing words than strong readers did. For the pseudowords, this difference between strong and weak readers was even larger.

According to the theory of Chall (1983 and 1996b, as cited in both Blok et al., 2002 and Kuhn & Stahl, 2003), children learn, with practice, to automatically recognize often-used words. As it is faster to automatically recognize words than having to decode every word, practice leads to faster reading. According to the phonologic coherence hypothesis, children gain speed in reading, because with repetition, the connection between the letter nodes and the phoneme nodes strengthens (Bosman & Van Orden, 2003). However, the children in the current study did not read faster over time. A possible explanation is that the children already knew the words before the pretest. This pretest was held in March. While most schools used

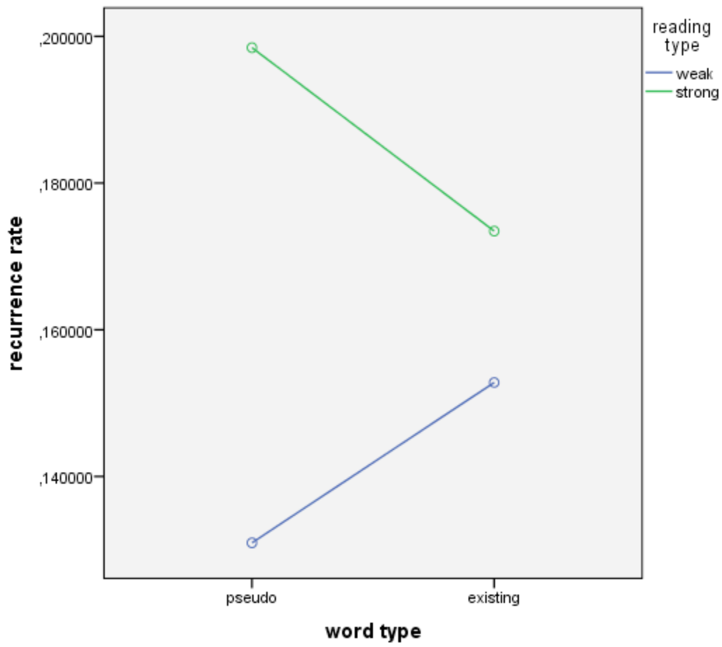

Figure 7. Interaction effect of the recurrence rate in the overall-analyses.

the reading program ‘Veilig Leren Lezen (‘Learning to Read Safely’; Verhoeven & Van Leeuwe, 2009), the children were already taught more difficult words, like CVCC words, and CCVCC words. According to the reading program, all CVC words should already be mastered. When a child already recognizes a word immediately, it is harder to gain speed, than while learning a new word (Kuhn & Stahl, 2003). So, it is possible that a ceiling effect was present for reaction times. According to the meta-analysis of Kuhn and Stahl (2003), greater gains in achievement were found, when more difficult words were used. Moreover, studies with easier materials, did not find any significant results of repetition. Whether or not a ceiling effect caused the finding in the current study that speed in reading did not increase, is hard to say, while the test leader had to press a key, before the time was registered. As a consequence, the reaction times of the test leader were included in the response times of the participant. The test leader tried to react as fast as possible, so for the *relative* response times it probably did not matter, while the reaction times of the test leader will be approximately the same for all words and all children. However, the *absolute* reaction times are higher, because the test leader needed time to press the key. Therefore, the exact time a child needed to read a word, remains unclear. In future research, it might be recommendable to use an amplification voice key. In contrast to other types of voice keys, the amplification voice key can measure the response times of the child with millisecond precision (Duyck et al., 2008). Previous research has shown that this procedure can be used reliably with children as young as 6 years old (Bosman & De Groot, 1991, 1995).

Our finding that the children did not read more accurate after training was surprising, because in other studies, both repetition (Stoddard, Valcante, Sindelar, O’Shea, & Algozzine, 1993) and feedback (Bangert-Drowns et al., 1991) improved performance. The finding of the current study might also be the result of a ceiling effect. In the pretest, the children already made only a few errors. Strong readers already read 99% of the words correctly, for weak readers this was 91%. As a consequence, little improvement on accuracy was possible. To avoid a ceiling effect, the level of the task should have been adapted to the reading skills of the children (Korat, 2010). In the current study, this would mean that more difficult words should have been used. Only CVC words (like ‘kat’ [cat]) were used, while in the selection of the children for the study by means of the DMT, more difficult one-syllable words were used, for instance CCVCC words (like 'grond' [ground]) and CVCCC words (like 'barst' [burst]). The lack of progress in both speed and accuracy is also reflected in the lack of significant differences in the complexity measures. Measures of complexity say something about the stability of a system (Smith & Thelen, 2003). A system that is changing is less stable than a system that already has mastered a certain skill. No differences were found between strong readers and weak readers on the measures of complexity. This implies that there might be no differences between the two groups on the stability of the process. For weak readers, the process of reading words is not more or less stable than for strong readers.

Also surprising was the finding that the weak readers made more errors in reading words than strong readers did. This is unexpected, because in languages with clear letter-sound correspondences, as in Dutch, differences in the level of reading skills of novice readers reveal themselves mostly in speed (Wimmer, 1993; Everatt & Elbeheri, 2008). In previous research, the number of errors appeared to be the same for both strong and weak readers. However, two types of dyslexia are suggested in the literature (Van der Schoot et al., 2003; see also Hendriks & Kolk, 1997): *Spellers* and *guessers*. Spellers read slowly, but accurately, guessers read fast, yet with many errors. A possible explanation for the finding that weak readers made more errors while reading words than strong readers is that the group of participants in the current study included many guessers. Having guessing children in the group of participants, might lead to a difference between strong and weak children in the amount of errors made in reading the words. However, this explanation cannot explain why response times of the weak readers did differ from the response times of the strong readers. In the current study, differences between strong and weak readers were found for both response times and accuracy. Another explanation stems from the phonologic coherence hypothesis. According to this model, it is possible that for weak readers, the connection between letter nodes and phoneme nodes is not bidirectional. As a consequence, no feedback loops can be formed, causing problems in reading the word correctly. The finding that the difference in accuracy between strong and weak readers is even larger for pseudowords than for existing words, is a confirmation of the phonologic coherence hypothesis. That is, according to this model, children might use the semantics of a word, while reading (Bosman & Van Orden, 2003). While reading pseudowords, semantics cannot be used, resulting in problems with reading these words, at children who have poorly developed connections between letter nodes and phoneme nodes.

Another, more explorative question, was whether or not a difference would exist between strong and weak readers and between words and pseudowords on the measures of complexity. The only difference found with regard to the complexity measures was in the recurrence rate in the over-all analyses (one time series of all existing words and one time series of all pseudowords read during the pretest, complete training, and transfer session). The effect of reading type (strong readers versus weak readers) on recurrence rate differed for word type. That is, while reading existing words, the difference between the two reading types in the average recurrence rate is smaller than while reading pseudowords. Recurrence rate is a measure that provides information on the stability of a state of the system. The higher the recurrence rate, the more stable the state of the system is. For strong readers, the difference between reading pseudowords and reading existing words on the stability of the state of the system is smaller than for weak readers.

In a recent article of Wijnants et al. (2012) clear differences in complexity measures emerged between dyslexic readers (in the age of 7 to 8) and strong readers (in the age of 6 to 7, matched to the dyslexic readers on reading-age). Using these measures, it was even possible to differentiate between the severity of the reading impairment. Although the study of Wijnants et al. found clear differences on measures of complexity, it is unclear whether or not these differences also exist for younger children, who are learning to read. The weak readers in the current study have less reading experience than the children already diagnosed with dyslexia. Not all weak readers will eventually develop dyslexia, so the inclusion criteria for the participants in the current study were less restrictive. Another difference between the study of Wijnants et al. and the current study was the use of pseudowords in the current study. The analyses of the pseudowords implicated the use of semantics while reading words, as was stated by the phonologic coherence hypothesis. However, the major difference between the two studies is the feedback that was offered in the current study. According to many articles (for instance Bangert-Drowns et al., 1991; Schimmel, 1983 as cited in Bangert-Drowns et al., 1991; Kulik & Kulik, 1988), an advantage of giving feedback is that it improves the learning process. However, in complexity measures, it is desirable to capture a continuous process (Thelen, Schöner, Scheier, & Smith, 2001). In the current study, the feedback interrupted the process, and thus the time series. It might be that, because of the interruption of the process by feedback, no differences were found on the complexity measures between the readers (strong vs. weak), between words (existing vs. pseudo), or over time.

Another interruption of the process was caused by the fact that sessions are interrupted over time. The time series used in the overall analyses were constructed artificially by combining the time series of separate sessions into one. Hence, large variability in the sampling rate was present within these time series. The time between two consecutive reaction times could vary from only a few seconds, to as much as 4 days (from the last word of one session, to the first word of a next session). In contrast, for the separate time series per session, this variability was not present, yet these time series were only 50 reaction times long. A time series of 50 reaction times is usually too short for the purpose of studying the underlying dynamics. It is possible, that as a result of the interruption of the process, no differences were found in the measures of complexity, either between groups or over time. Whether or not any differences between strong and weak readers in measures of complexity exist at all, remains a question for future research. The same goes for the possible difference in complexity measures between the processes of reading existing words versus pseudowords.

In future research it is recommended to take into account the assumptions of the analyses used, when setting up an experiment. When using complexity measures, a time series should be captured continuously and with a constant sampling rate. Although feedback leads to learning, it might not be the right way to give feedback after each word read, when the dynamics of a system (the reading child) are explored.

To sum up, the children did not gain speed in reading the words and their accuracy did not improve. Weak readers did read slower than strong readers, but made more errors. The differences between strong and weak readers were larger for the pseudowords than for existing words. No differences in the measures of complexity existed between reading type or word type, probably due to choices made in the design of the experiment. In future research, the level of difficulty of the words should be better adapted to the reading skills of the children, in order to prevent ceiling effects. Also more attention could be paid to the design of the experiment, in order to catch the process that is going on.

References

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213-238. doi: 10.3102/00346543061002213

Bassano, D. & Van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: a quantitative approach to changes, transitions and intra-individual variability in early grammatical development. *Developmental Science, 10,* 588-612. doi:10.1111/j.1467-7687.2006.00629.x

Blachman, A. B. (2000). Phonological awareness. In P.D. Pearson, M.L. Kamil, & R. Barr (Eds.), *Handbook of Reading Research: Volume 3* (pp. 483-502). Mahwah, NJ: Erlbaum.

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72,* 101-130. doi:10.3102/00346543072001101

Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *Neuroimage, 57,* 695-703. doi:10.1016/j.neuroimage.2010.11.003

Bosman, A. M. T., & De Groot, A. M. B. (1991). De ontwikkeling van woordbeelden bij beginnende lezers en spellers. *Pedagogische Studiën, 68,* 199-215.

Bosman, A. M. T., & De Groot, A. M. B. (1995). Evidence for assembled phonology in beginning and fluent readers as assessed with the first-letter-naming task. *Journal of Experimental Child Psychology, 59,* 234-259. doi:10.1006/jecp.1995.1011

Bosman, A. M. T., & Van Orden, G. C. (2003). Het fonologisch coherentiemodel voor lezen en spellen. *Pedagogische Studiën, 80,* 391-406.

Davies, R., Rodríguez-Ferreiro, J., Suárez, P., & Cuetos, F. (2013). Lexical and sub-lexical effects on accuracy, reaction time and response duration: impaired and typical word and pseudoword reading in a transparent orthography. *Reading and Writing, 26,* 721-738. doi: 10.1007/s11145-012-9388-1

Duyck, W., Anseel, F., Szmalec, A., Mestdagh, P., Tavernier, A., & Hartsuiker, R. J. (2008). Improving accuracy in detecting acoustic onsets. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1317-1326. doi:10.1037/0096-1523.34.5.1317

Everatt, J., & Elbeheri, G. (2008). Chapter 21: dyslexia in different orthographies: variability in transparency. In: G. Reid, A. Fawcett, F. Manis & L. Siegel (Eds.), *The SAGE handbook of dyslexia* (pp. 427-439). Los Angeles, CA: SAGE.

Fletcher, J. M. (2009). Dyslexia: The evolution of a scientific concept. *Journal of the International Neuropsychological Society, 15,* 501-508. doi:10.1017/S1355617709090900

Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A, 33,* 1134–1140. doi:10.1103/PhysRevA.33.1134

Girden, E. (1992). *ANOVA: Repeated measures*. Newbury Park, CA: Sage.

Hellendoorn, J. & Ruijssenaars, W. (2000). Personal experiences and adjustment of Dutch adults with dyslexia. *Remedial and Special Education, 21,* 227-239. doi:10.1177/074193250002100405

Hendriks, A. W., & Kolk, H. H. J. (1997). Strategic control in developmental dyslexia. *Cognitive Neuropsychology, 14,* 321-366. doi:10.1080/026432997381510

Jongen, I., & Krom, R. (2009). *DMT en AVI.* Arnhem, the Netherlands: CITO.

Kennel, M. B., Brown, R., & Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A, 45,* 3403–3411. doi:10.1103/PhysRevA.45.3403

Korat, O. (2010). Reading electronic books as a support for vocabulary, story comprehension and word reading in kindergarten and first grade. *Computers & Education, 55,* 24-31. doi:10.1016/j.compedu.2009.11.014

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95,* 3-21. doi:10.1037/0022-0663.95.1.3

Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58,* 79-97. doi:10.3102/00346543058001079

Lewis, M., Zimmerman, S., Hollenstein, T., & Lamey, A. (2004). Reorganization in coping behavior at 1½ years: Dynamic systems and normative change. *Developmental Science, 7,* 56-73. doi:10.1111/j.1467-7687.2004.00323.x

Lichtwarck-Aschoff, A., Hasselman, F., Cox, R., Pepler, D., & Granic, I. (2012). A characteristic destabilization profile in parent-child interactions associated with treatment efficacy for aggressive children. *Nonlinear Dynamics, Psychology, and Life Sciences, 16,* 353-379.

Reid, G., Fawcett, A., Manis, F., & Siegel, L. (2008). *The SAGE handbook of dyslexia.* Los Angeles, CA: SAGE.

Riley, M. A., Balasubramaniam, R., & Turvey, M. T. (1999). Recurrence quantification analysis of postural fluctuations. *Gait & Posture, 11,* 12–24. doi:10.1016/S0966-6362(98)00044-7

Rueckl, J. G. (2002). The dynamics of visual word recognition. *Ecological Psychology, 14,* 5-19. doi:10.1207/S15326969ECO1401&2double_2

Smith, L. B. & Thelen, E. (2003). Development as a dynamic system. *TRENDS in Cognitive Sciences, 7,* 343-348. doi:10.1016/S1364-6613(03)00156-6

Stoddard, K., Valcante, G., Sindelar, P., O'Shea, L., & Algozzine, P. (1993). Increasing reading rate and comprehension: The effects of repeated readings, sentence segmentation, and intonation training. *Reading Research and Instruction, 32,* 53–65.

Suárez-Coalla, P., & Cuetos, F. (2012). Reading strategies in Spanish developmental dyslexics. *Annals of Dyslexia, 62,* 71-81. doi: 10.1007/s11881-011-0064-y

Takens, F. (1981). Detecting strange attractors in fluid turbulence. In D. A. Rand & L. S. Young (Eds.), *Dynamic Systems and Turbulence* (pp. 366–381). New York, NY: Springer.

Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences, 24,* 1-86. doi:10.1017/S0140525X01003910

Turvey, M. T. (2007). Action and perception at the level of synergies. *Human Movement Science, 26,* 657-697. doi:10.1016/j.humov.2007.04.002

Van der Schoot, M., Licht, R., Horsley, T. M., Aarts, L. T., Van Koert, B., & Sergeant, J. A. (2003). Inhibitory control during sentence reading in dyslexic children. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence, 10,* 173-188. doi:10.1080/09297040490911032

Van Orden, G., & Holden, J. G. (2002). Intentional contents and self control. *Ecological Psychology, 14,* 87-109. doi:10.1207/S15326969ECO1401&2double_5

Van Orden, G. C., Jansen op de Haar, M. A., & Bosman, A. M. T. (2010). Complex dynamic systems also predict dissociations, but they do not reduce to autonomous components. *Cognitive Neuropsychology, 14,* 131-165. doi:10.1080/026432997381646

Van Orden, G., & Kloos, H. (2003). The module mistake. *Cortex, 39,* 164-166. doi:10.1016/S0010-9452(08)70092-3

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522. doi:10.1037/0033-295X.97.4.488

Van Weerdenburg, M., Verhoeven L., Van Balkom, H., & Bosman, A. (2009). Cognitive and linguistic precursors to early literacy achievement in children with specific language impairment. *Scientific Studies of Reading, 13,* 484-507. doi:10.1080/10888430903162936

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2-40. doi:10.1046/j.0021-9630.2003.00305.x

Verhoeven, L., & Van Leeuwe, J. (2009). Modeling the growth of word-decoding skills: Evidence from Dutch. *Scientific Studies of Reading, 13,* 205-223. doi:10.1080/10888430902851356

Wentink, H., Hoogenboom, S., & Cox, A. (2009). *Leesonderwijs en leesbegeleiding voor leerlingen met ernstige spraak- en/of taalmoeilijkheden (ESM) – een katern bij de protocollen leesproblemen en dyslexie.* ’s-Hertogenbosch, the Netherlands: Masterplan Dyslexie.

Wijnants, M. L., Cox, R. F. A., Hasselman, F., Bosman, A. M. T., & Van Orden, G. (2012). A trade-off study revealing nested timescales of constraint. *Frontiers in Physiology, 3,* 1-15. doi:10.3389/fphys.2012.00116

Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., & Van Orden, G. (2012). An interaction-dominant perspective on reading fluency and dyslexia. *Annals of Dyslexia, 62,* 100-119. doi:10.1007/s11881-012-0067-3

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics, 14,* 1-33. doi:10.1017/S0142716400010122

Appendix A. The Mean Scores of the Analyses per Block of Words

Session number	Reading type	Word type	Response times in ms	% Correct
			M (SD)	M (SD)
1	Strong	Existing	1843 (292)	98.86 (2.24)
		Pseudo	2055 (369)	99.05 (1.75)
	Weak	Existing	4372 (1041)	93.11 (8.32)
		Pseudo	4787 (1150)	91.78 (6.50)
2	Strong	Existing	1978 (538)	99.43 (1.43)
		Pseudo	2168 (607)	98.48 (2.36)
	Weak	Existing	4204 (1016)	96.44 (5.80)
		Pseudo	4662 (1133)	90.00 (6.89)
3	Strong	Existing	1972 (508)	99.62 (1.20)
		Pseudo	2037 (482)	98.29 (2.39)
	Weak	Existing	4486 (1538)	90.67 (11.15)
		Pseudo	4753 (1307)	89.33 (8.89)
4	Strong	Existing	2072 (764)	99.24 (2.72)
		Pseudo	2026 (407)	99.05 (1.75)
	Weak	Existing	4135 (1119)	92.67 (8.15)
		Pseudo	4484 (1316)	88.67 (12.37)
5	Strong	Existing	1842 (345)	99. 62 (1.20)
		Pseudo	2085 (707)	98.67 (2.31)
	Weak	Existing	4573 (1424)	92.67 (8.92)
		Pseudo	4879 (1492)	89.33 (12.72)
6	Strong	Existing	1877 (577)	99.24 (1.61)
		Pseudo	1997 (494)	98.86 (2.24)
	Weak	Existing	4326 (1382)	91.56 (9.09)
		Pseudo	5028 (1881)	88.67 (12.96)
7	Strong	Existing	1899 (388)	99.62 (1.20)
		Pseudo	1976 (423)	99.05 (1.75)
	Weak	Existing	4392 (1347)	93.56 (6.74)
		Pseudo	4965 (1937)	91.33 (10.74)
8	Strong	Existing	1915 (524)	99.24 (1.61)
		Pseudo	1998 (744)	99.24 (2.05)
	Weak	Existing	4331 (1413)	91.11 (12.31)
		Pseudo	5004 (1549)	91.56 (8.22)
9	Strong	Existing	1811 (436)	99.43 (1.43)
		Pseudo	1905 (444)	99.43 (1.43)
	Weak	Existing	4528 (1972)	91.33 (11.08)
		Pseudo	4617 (1711)	88.00 (11.06)

10	Strong	Existing	1881 (421)	99.43 (1.43)
		Pseudo	1913 (469)	99.43 (1.43)
	Weak	Existing	4406 (1752)	94.67 (6.44)
		Pseudo	4533 (1334)	90.22 (11.50)
Pretest	Strong	Existing	1886 (297)	99.05 (1.75)
		Pseudo	2254 (483)	98.67 (2.63)
	Weak	Existing	4356 (891)	94.00 (6.47)
		Pseudo	4924 (636)	88.44 (8.22)
Transfer session	Strong	Existing	2041 (596)	99.05 (1.75)
		Pseudo	2058 (518)	98.29 (2.99)
	Weak	Existing	5178 (2324)	91.11 (9.95)
		Pseudo	5534 (2651)	89.78 (10.47)

Appendix B. The Mean Scores on the Complexity Measures of the Analyses per Session

Session number	Reading type	Recurrence rate	Determinism	Entropy	Meanline
		M(SD)	M(SD)	M(SD)	M(SD)
1	Strong	.17 (.08)	.74 (.09)	1.30 (.28)	3.08 (.52)
	Weak	.15 (.07)	.71 (.13)	1.22 (.27)	2.96 (.37)
2	Strong	.21 (.13)	.75 (.14)	1.43 (.45)	3.52 (1.80)
	Weak	.16 (.11)	.73 (.10)	1.25 (.33)	3.06 (.65)
3	Strong	.14 (.08)	.70 (.13)	1.16 (.31)	2.91 (.47)
	Weak	.14 (.05)	.71 (.09)	1.19 (.23)	2.90 (.37)
4	Strong	.17 (.07)	.73 (.11)	1.33 (.28)	3.19 (.62)
	Weak	.17 (.09)	.76 (.07)	1.30 (.23)	3.06 (.36)
5	Strong	.19 (.11)	.74 (.11)	1.39 (.37)	3.37 (1.06)
	Weak	.13 (.05)	.70 (.08)	1.14 (.17)	2.80 (.23)
6	Strong	.20 (.14)	.75 (.13)	1.40 (.44)	3.41 (1.00)
	Weak	.16 (.10)	.73 (.14)	1.34 (.38)	3.23 (.83)
7	Strong	.17 (.07)	.75 (.11)	1.34 (.31)	3.13 (.52)
	Weak	.17 (.09)	.74 (.09)	1.30 (.33)	3.19 (.80)
8	Strong	.16 (.06)	.75 (.07)	1.28 (.32)	3.08 (.44)
	Weak	.18 (.08)	.74 (.11)	1.33 (.32)	3.12 (.53)
9	Strong	.21 (.10)	.79 (.09)	1.44 (.34)	3.37 (.67)
	Weak	.16 (.07)	.72 (.11)	1.19 (.27)	2.91 (.39)
10	Strong	.17 (.06)	.74 (.08)	1.35 (.20)	3.12 (.34)
	Weak	.16 (.10)	.72 (.13)	1.25 (.41)	3.17 (.92)
Pretest	Strong	.17 (.11)	.75 (.12)	1.37 (.49)	3.47 (1.46)
	Weak	.16 (.06)	.73 (.07)	1.26 (.29)	3.06 (.48)
Transfer session	Strong	.20 (.14)	.75 (.11)	1.32 (.43)	3.23 (1.03)
	Weak	.19 (.09)	.76 (.11)	1.34 (.33)	3.16 (.63)

Appendix C. The Mean Scores on the Complexity Measures of the Overall-Analyses

Reading type	Word type	Recurrence rate	Determinism	Entropy	Meanline
		M(SD)	M(SD)	M(SD)	M(SD)
Strong	Existing	.17 (.11)	.88 (.05)	2.01 (.39)	4.63 (1.68)
	Pseudo	.20 (.14)	.88 (.05)	2.08 (.43)	4.92 (1.93)
Weak	Existing	.15 (.10)	.86 (.05)	1.93 (.35)	4.29 (1.31)
	Pseudo	.13 (.09)	.84 (.07)	1.84 (.39)	4.13 (1.29)