

A comparative study of two methods of synthetic phonics instruction for learning how to read: Jolly Phonics and THRASS

Carol Callinan & Emile van der Zee

The National Strategy for Primary Schools in England (2006) advocates synthetic phonics as a means for teaching children to read. No studies exist to date comparing the effectiveness of different commercially available synthetic phonics methods. This case study compared two schools at which Jolly Phonics (JP) was taught with one school at which THRASS (Teaching Handwriting, Reading and Spelling Skills) was taught at Reception level (4 to 5 years) over a one-year period. Reading ability for words and non-words as well as short-term memory ability for words and phonemes improved in all schools. However, reading ability improved more in one JP school compared to the THRASS school, with no differences between the other JP school and the THRASS school. This paper considers how particular variables may mask instruction method effects, and advocates taking such factors into account for a more comprehensive future evaluation of synthetic phonics methods.

THE ABILITY TO READ is one of the most important skills that children attain during their early years, and is a necessary pre-requisite for children's continuing education (e.g. McGuinness, 1998, 2005, 2004). This paper considers the teaching of reading in English, and investigates how reading ability for words and non-words as well as short-term memory ability for words and phonemes improves in Reception year children in three different schools using two different methods of reading instruction.

Research investigating the structure of the English spelling system demonstrates that this system is difficult for children to learn (Goswami, 2005; Zeigler & Goswami, 2005). English contains many to many mappings for spelling (graphemes) to sound (phonemes); for example, the letter 'c' can be pronounced as /k/ ('cat'), /s/ ('city'), or /tʃ/ ('church'). The UK Government expects primary schools in England to use the synthetic phonics method (The Primary Framework for Literacy and Mathematics, 2006; Rose Review, 2006). And indeed, research has shown that synthetic phonics can be successful (Johnston & Watson, 2005, 1997; Watson & Johnston, 1998). However, empirical research has focused on

contrasting synthetic phonics teaching with analytic phonics, where words and word parts are taught before individual graphemes. In this paper we contrast two commercially available packages for teaching synthetic principles: Jolly Phonics or JP (Lloyd, 1992), and Teaching Handwriting, Reading and Spelling Skills or THRASS (Davies & Ritchie, 1998; Davies, 2006).

JP offers guidance for the first nine weeks of tuition. After this, teaching is determined by the teacher. The programme is delivered in 15-minute sessions by introducing at least one new phoneme per day. Children are tutored in 42 of the 44 phonemes in English (as defined by Flesch, 1955), and receive instructions in 46 of the most common graphemes. The presented phonemes maximise use for constructing new words (see Table 1). Phonemes are associated with kinaesthetic activities; for example, imitating a light switch being on or off for the phoneme /o/. The programme uses both synthetic and analytic principles.

THRASS offers guidance for the first three years of tuition. The programme uses pictures in relation to two, three and four letter graphemes. The programme contains

Table 1: The JP schedule of grapheme tuition, the grouping is intended to maximise opportunities for word blending.

| | | | | | | | | |
|-----------|-------------|-----------|-------------|--------------------|-------------------|----------------------|-------------|----------------------------------|
| Week 1 | Week 2 | Week 3 | Week 4 | Week 5 | Week 6 | Week 7 | Week 8 | Week 9 |
| s a t i p | n c/k e h r | m d g o u | l f b a i j | oa ie ee or z w | ng v y x oo OO | ch sh th th qu ou | oi ue er ar | Main alternative spellings |

10 stages (see Table 2) and is also delivered in 15-minute sessions each day. Children receive instructions for all 44 English phonemes and 120 of the most common graphemes. The programme is presented in a fixed order. During lessons teachers will utilise pictures to help children identify phonemes while showing different possible spellings. The programme initially uses synthetic principles and later analytic strategies (see Table 2).

In this study participants completed two reading measures: one for words (the Burt Reading Test Revised, 1974) and one for non-words (Miskin, 2006). While both measures test reading ability, the last measure, which was originally developed for use with second language learners, tests accurate decoding ability and reveals whether children are using their phonic skills to

decode words rather than a whole word decoding strategy. Although we expected an improvement on both measures (Johnston & Watson, 2005), due to the method of instruction we anticipated that THRASS taught participants would exhibit higher levels of non-word decoding ability.

Short-term memory tests are normally not used in combination with reading measures. However, memory span for words is associated with reading task performance and the ability to learn reading skills (Griffith & Snowling, 2002; Dufva, Niemi & Voeten, 2001). In addition, deficits in verbal short-term memory may reveal underlying problems with – among other things – phonological processing (Gathercole et al., 2004). We, therefore, used two newly-constructed verbal short-term memory tests: one test measuring word string recall, and

Table 2: The 10 stages of THRASS teaching, Stages 1 to 9 pertain to the THRASS picture chart.

| Stage Number | Age (Years) | Stage Title and Learning Aims |
|--------------|-------------|---|
| 1 | 3, 4, 5 | Picture location – locate pictures on the chart |
| 2 | 3, 4, 5 | Letter location – locate and name letters |
| 3 | 3, 4, 5 | Letter formation – name and form letters |
| 4 | 3, 4, 5 | Grapheme location – locate and name graphemes |
| 5 | 3, 4, 5 | Keyword location – locate and name 120 keywords |
| 6 | 5 | Phoneme location – locate and articulate 44 phonemes |
| 7 | 5 | Keyword synthesis – blend, read and spell keywords |
| 8 | 6 | Keygrapheme recall – visualise and spell graphemes |
| 9 | 6 | Keyword analysis – read, spell and analyse 120 keywords |
| 10 | 7 | THRASS 500 Tests – Read and spell THRASS 500 words |

another test measuring phoneme string recall. When children learn to read they need to accurately recall and synthesise words. We thus expected that short-term memory capacity for words and phonemes improves for all children during their first year of reading instruction. However, it is anticipated that – again due to instruction – THRASS participants would demonstrate better short-term memory recall for words and phonemes compared to JP instructed participants.

In general, girls perform better in reading tasks than boys (Brooks, Pugh & Schagen, 1996). In line with this, we expected that females score better than males on both reading and short-term memory measures.

In addition, we investigated whether having older siblings positively influenced participants' reading development (Rogoff, 1998; Tulviste, 1991). Older siblings may scaffold learning in younger siblings (Wood, 1989), or parents may have less time available supporting learning to read when older siblings are present.

Method

Participants

Twenty-four female and 30 male reception class children (mean age 4 years, 5 months); 16 from JP school 1, 18 from a THRASS school, and 20 from JP school 2.

Materials

Words and non-words of the Burt Reading Test Revised (1974) and the Miskin Non-word Test (2006) were printed onto laminated flash cards.

A short-term memory test for words was developed using words from Stuart et al.'s vocabulary (2003) (see Appendix 1). The test comprised nine blocks of word strings, each block increasing from one to nine words. Each block contained three strings of words.

A short-term memory test for phonemes was developed using 39 of the 44 English phonemes. The phoneme strings did not

create existing words (Appendix 2). Five diphthongs (e (air), I (ear), I (oy), u (oor) and au (ow)) were not included to ensure that participants only heard one sound at a time. The test comprised nine blocks of phoneme strings, each block increasing from one phoneme to nine. Each block contained three strings of an equal number of phonemes.

Participants' responses were recorded on a cassette or a voice recorder. Response sheets were generated for three participants who chose not to have their voices recorded.

A demographic questionnaire asked about a participant's first language, presence of older siblings, average time spent reading, date of birth and gender. Information about social-economic status could not be obtained.

A synthetic phonics programme checklist was used (Hepplewhite, 2005) to establish whether the phonics lessons indeed used the features associated with the phonics programmes.

Procedure

All sessions lasted 15 minutes or less. In September 2006 all tests were administered in one block. In November 2006 and June 2007 children were tested in two sessions: session one incorporated the Burt Reading Test Revised and the short-term memory test for word strings, and session two – on the same day – the short-term memory tests for phoneme strings and the Miskin non-word test.

Flashcards for the reading tests were presented in a fixed order. Spontaneous corrections were accepted. Responses to the Burt Reading Test Revised were correct if conforming to word structure; for example, 'ice' for 'is'. Responses for the Miskin non-word test were correct if conforming to phonemic word structure; for example, 'v ' or 'v k' in response to 'vok'. Word presentation continued until participants failed to name three consecutive words. The total number of items read constituted a participant's score.

The short-term memory tests were administered verbally. Participants were asked to repeat the experimenter. An initial rehearsal block of three strings, two items in length, was used to ensure that all participants understood the task. The tests were terminated if participants made three consecutive errors. The number of items in the last correct repetition constituted a participant's score.

One phonics lesson was observed and taped per class. These recordings confirmed that the schools were teaching the JP and THRASS methods as prescribed. Information regarding class size, time of tuition and presence or absence of teaching assistants was also collected.

All procedures followed the Society's guidelines for ethical approval. The testing was presented as reading and repetition games. Participants were tested individually in a quiet area.

Results

Nineteen children were removed from the analyses (three moved school, three had a different first language, four withdrew, seven were absent during one session, and two requested a break; eight were from JP school 1, six from the THRASS school, and five from JP school 2). Children were omitted from the final analysis in order to ensure matched sampling in relation to the variables that we investigated (for example, matched language ability skills). Seventeen out of the 35 remaining participants were female and 18 male. Only significant results are presented.

The analyses first focus on September, then on November and June.

An ANOVA on the Burt word reading test scores showed that participants were fully matched on their reading abilities in September. None of the children were able to do the Miskin non-word reading test. Another ANOVA revealed that participants were also fully matched on their short-term memory abilities for words. Finally, an ANOVA on short-term phoneme memory

scores showed that the THRASS school scored lower ($M=2.08$; $SE=0.67$) than JP school 1 ($M=3.38$; $SE=0.71$) and JP school 2 ($M=3.47$; $SE=0.42$).

A repeated measures ANOVA and post-hoc paired samples *t*-tests on the Burt word reading scores showed that reading ability scores significantly increased from September ($M=0.37$; $SE=0.24$), to November ($M=2.34$; $SE=0.57$), and again to June ($M=15.37$; $SE=1.94$) (see Figure 1), and were higher for participants without older siblings ($M=19.28$; $SE=3.28$) than participants with older siblings ($M=11.24$; $SE=1.54$).

More specifically, JP school 1 improved between September ($M=0$; $SE=0$) and November ($M=1$; $SE=0.46$), and from November to June ($M=15.25$; $SE=4.63$), as did JP school 2 (September $M=0.8$; $SE=0.55$; November $M=4.6$; $SE=1.05$; June $M=20.73$; $SE=2.98$). The THRASS school did not improve between September ($M=0.08$; $SE=0.08$) and November ($M=0.42$; $SE=0.34$), but did between November and June ($M=8.75$; $SE=1.95$). The sibling effect was due to JP school 2 receiving higher ratings for children without older siblings compared to those with older siblings in June (respectively $M=26.75$; $SE=4.34$ and $M=13.85$; $SE=2.11$).

A repeated measures ANOVA and post-hoc *t*-tests on the Miskin non-word reading scores showed that scores improved from September ($M=0$; $SE=0$) to November ($M=0.97$; $SE=0.39$), and from November to June ($M=6.54$; $SE=1.86$) (see Figure 1), and that participants without older siblings showed higher reading scores ($M=10.67$; $SE=3.31$) than participants with older siblings ($M=2.18$; $SE=0.78$).

More specifically, JP school 2 showed sustained improvement, the THRASS school showed only improvement in June, and JP school 1 did not show any improvement at all. In June the JP school 1 performed at the same level as both other schools, although JP school 2's performance was better compared to the THRASS school. The sibling effect could not be explained any further (see Figure 2).

Figure 1: Mean and standard error scores on both reading tasks (Burt=word reading test; Miskin=non-word reading test) in the three time periods (September and November 2006 and June 2007) for all three schools (JP1=JP 1 school; THRASS=the THRASS school; JP2=JP 2 school).

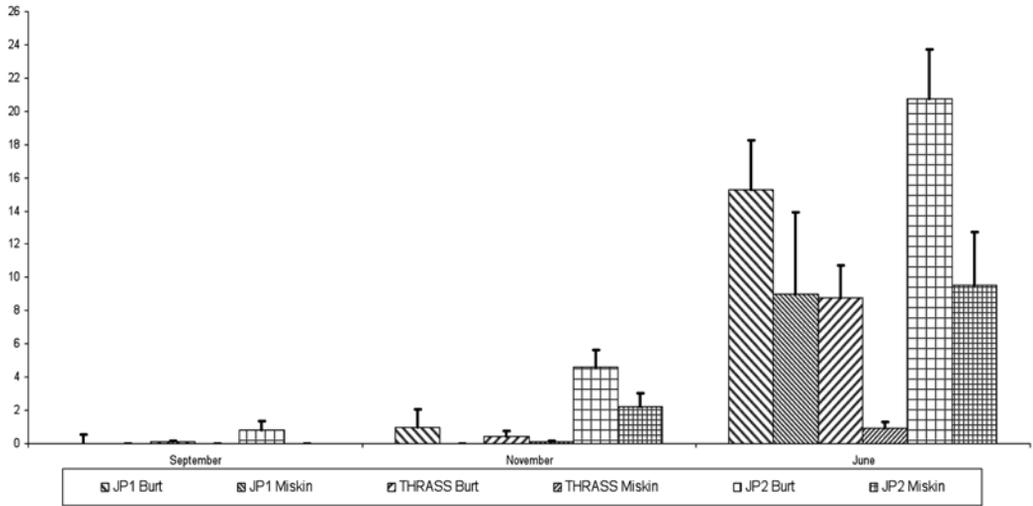


Figure 2: Means and standard errors for participants without older siblings and participants with older siblings on the Miskin non-word reading test in the three time periods (September and November 2006 and June 2007).

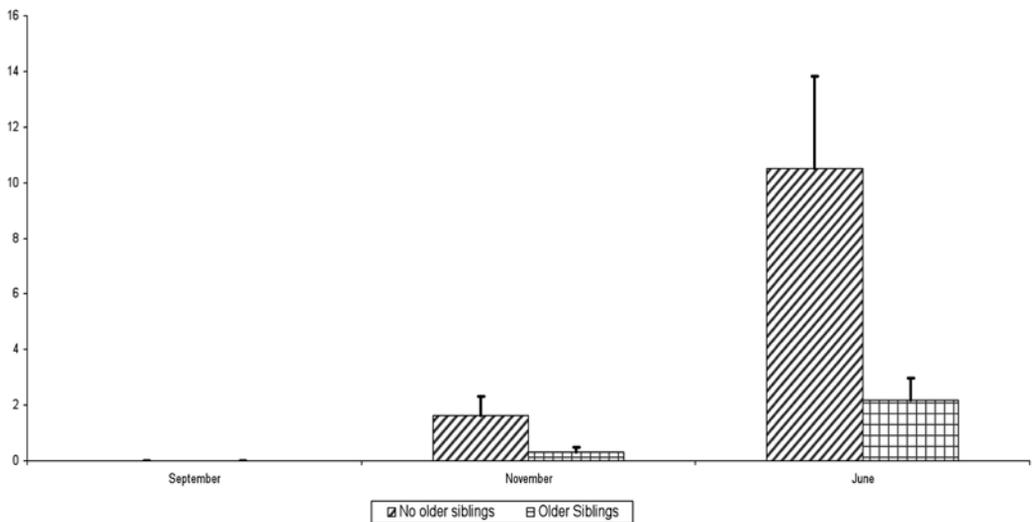
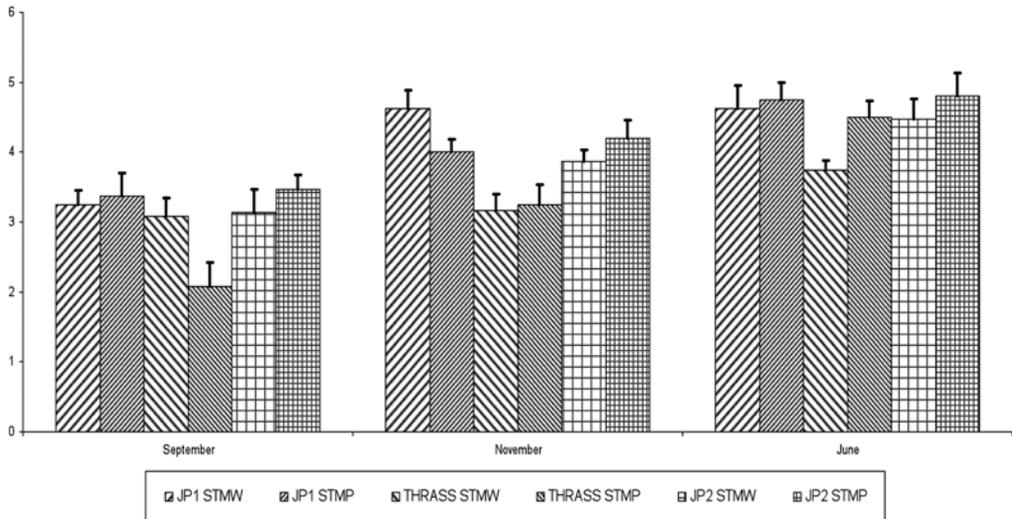


Figure 3: Means and standard errors on both word and phoneme short-term memory tasks for all three schools (JP1=JP 1 school; THRASS=the THRASS school; JP2=JP 2 school) for the three time periods (September and November 2006 and June 2007).



A repeated measures ANOVA and post-hoc *t*-tests on short-term memory scores for words showed that scores increased from September ($M=3.14$; $SE=0.19$) to November ($M=3.8$; $SE=0.15$) and again from November to June ($M=4.26$; $SE=0.15$) (see Figure 3). The absence of any interaction effects showed that this improvement was independent of the pupils' school, gender or sibling structure.

A repeated measures ANOVA and post-hoc *t*-tests on short-term memory scores for phonemes showed an improvement from September ($M=2.97$; $SE=0.19$) to November ($M=3.8$; $SE=0.16$), and from November to June ($M=4.6$; $SD=0.17$) (see Figure 3). Again, the absence of any interaction effects indicates that this improvement was independent of the pupils' school, gender or sibling structure.

Discussion

The results supported our main predictions: reading skills for both words and non-words improved in tandem with short-term memory skills for words and phonemes

during children's first year of synthetic phonics reading instructions. More specifically, by June 2007 the JP school 2 had made greater gains in both word and non-word reading tasks compared to the THRASS school. The JP school 1 did not differ from either school but failed to demonstrate an improvement in non-word reading. Children's improvements in their short-term verbal memory skills could not be linked to the method of instruction they received. These results mean that an anticipated advantage of THRASS instructed children for non-word reading and short-term memory performance was not found.

Our study is limited in that it only considered three different schools, two methods of synthetic phonic reading instruction, and only one year of instruction. The latter limitation may be responsible for not finding the anticipated advantage of THRASS instructed children for non-word reading and short-term memory performance. It could be argued that – due to its nature of instruction – the advantages of the THRASS method are only evident much later than at the end of

year 1. This limitation would be overcome by following children's reading development for more than one year. In order to consider what variables were responsible for our results our study also considered qualitative differences between schools which can be taken on board in a larger scale project comparing different methods of reading instruction. A post-hoc qualitative analysis shows that there are a number of factors in relation to which the three schools that we studied differed (see Table 3 below).

Based on the information in Table 3 it could be argued that the improvement of the children's reading abilities in JP school 2 is related to the time of instruction (early in the morning, when the children are still able to concentrate well on such a difficult activity), and the fact that the children are split into three ability groups (which may offer more tailored support). The way that classroom assistant support was provided might also have impacted on the outcomes in this case study: in JP school 2 the classroom assistant

provided tuition for the lower ability group, in the THRASS school the classroom assistants supervised the children during lessons rather than taking an active role in teaching.

All three schools used Oxford Reading Scheme books in order to support children reading. However, JP school 2 supplied a Phonics Reading Scheme in addition. The advantage of JP school 2 participants in the reading task may have been linked to this additional factor.

Finally, the differences between JP schools 1 and 2 may have been due to class size. Iacovou (2002) demonstrated that there may be an impact of class size on learning how to read in early years, with smaller class sizes being linked to better performance (depending on school size and type).

Our study also demonstrated an advantage of participants without older siblings compared to those with older siblings in relation to reading normal words and non-words (with the former being due to JP school 2). Although these results themselves merit

Table 3: Table of qualitative differences observed between the three schools during the teaching of phonics materials.

| | JP 1 School | THRASS School | JP 2 School |
|---|--|--|--|
| Time of phonics lesson | 9.55 Before play time | 11.15 After play time | 9.15 After morning register |
| Are the children split into ability groups? | No | No | Yes |
| Are there classroom assistants present during lessons? How many? | No | Yes 2 | Yes 1 |
| Role of the classroom assistant | N/A | Supervision of children during lessons | Teaches the lower ability group |
| How many reception classes in the school? | 1 | 2 | 2 |
| Number of children per class | ~30 | ~22 | ~22 |
| Reading schemes used | The Oxford Reading Tree, Collins Big Cat | The Oxford Reading Tree | The Oxford Reading Tree, Ginn, Phonics |

further investigation (for example, they do not support the scaffolding effect that older siblings were found to have on younger siblings; see Wood, 1989), the results do show that the measurement of reading development must control for or take into account contextual variables outside of the school setting.

No effects of gender on reading ability were found. This finding is consistent with the literature focusing on synthetic phonics tuition (e.g. McGuinness, 1998, 2004, 2005) which suggests that the synthetic method removes the variability found in reading scores between males and females. An alternative explanation might be that the low number of participants tested, combined with a greater male reading ability variance compared to a smaller female reading ability variance (as found for 15-year-old children, see Machin & Pekkarinen, 2008) has led to the absence of any male-female differences (for example, because males performed better in the sample compared to the population). Contrasting JP and THRASS in a more comprehensive study would make it possible to determine which of these explanations is correct.

We propose that further research should be conducted with a larger experimental sample and over a period of at least three years (the latter based on the way the THRASS method is defined). However, we believe that our study has demonstrated that verbal short-term memory capacity can be used as a variable in tandem with measuring reading performance, and that further research should take into account such variables as 'the use of ability groups', 'the time of tuition', 'the presence and role of classroom assistants', 'the use of reading support (including the reading schemes that are used)', 'family structure', and 'class size'.

The authors

Carol Callinan & Emile van der Zee

Department of Psychology,
University of Lincoln, UK.

Correspondence

Carol Callinan

Department of Psychology,
University of Lincoln,
Brayford Campus,
Lincoln LN6 7TS, UK.

E-mail: ccallinan@lincoln.ac.uk

References

- Brooks, G., Pugh, A. & Schagen, I. (1996). *Reading performance at 9*. Slough: NFER.
- Burt, C. (1974). *The Burt Reading Test* (rev ed.). Retrieved 16 May 2006 from: www.rrf.org.uk/Burtreadingtestonweb.pdf
- Davies, A. (2006). *Teaching THRASS: Whole-picture keyword phonics*. Chester: THRASS (UK) Ltd.
- Davies, A. & Ritchie, D. (1998). *THRASS: Teacher's manual*. Chester: THRASS (UK) Ltd.
- Dufva, M., Niemi, P. & Voeten, M.J.M. (2001). The role of phonological memory, word recognition, and comprehension skills in reading development: From pre-school to grade 2. *Reading and Writing, 14*, 91–117.
- Flesch, R. (1955). *Why Johnny can't read*. New York: Harper & Brothers.
- Gathercole, S.E., Tiffany, C., Briscoe, J., Thorn, A. & the ALSPAC team (2004). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry, 45*, 1–14.
- Griffiths, Y.M. & Snowling, M.J. (2002). Predictors of exception word and non-word reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology, 94*, 34–43.
- Goswami, U. (2005). Synthetic phonics and learning to read: A cross-language perspective. *Educational Psychology in Practice, 21*, 273–283.
- Hepplewhite, D (2005). Criteria for evaluating a phonics programme. Retrieved 12 May 2006 from: <http://syntheticphonics.com/pdf%20files/Criteria%20for%20evaluating%20a%20phonics%20programme.pdf>
- Iacovou, M. (2002). Class size in the early years: Is smaller really better? *Education Economics, 10*(3), 261–290.
- Johnston, R.S. & Watson, J. (2005). The effects of synthetic phonics teaching on reading and spelling attainment. *The Scottish Executive Education Department Report*. www.scotland.gov.uk/library5/education/ins17-00.asp
- Johnston, R.S. & Watson, J. (1997). Developing reading, spelling and phonemic awareness skills in primary school children. *Reading, 31*, 37–40.
- Lloyd, S. (1992). *The phonics handbook*. Chigwell, UK: Jolly Learning.
- Machin, S. & Pekkarinen, T. (2008). Global sex differences in test score variability. *Science, 322*, 5906, 1331–1332.
- McGuinness, D. (2005). *Language development and learning to read: The scientific study of how language development affects reading skill*. Cambridge, MA: MIT Press.
- McGuinness, D. (2004). *Early reading instruction: What science really tells us about how to teach reading*. Cambridge, MA: MIT Press.
- McGuinness, D. (1998). *Why children can't read and what we can do about it*. London: Penguin Books.
- Miskin, R. (2006). The Miskin Non-word Test. Retrieved 12 May 2006 from: www.rrf.org.uk/Ruth%20Miskin%20Nonsense%20Word%20Test.pdf
- Primary National Strategy (2006). *Primary Framework for Literacy and Mathematics: Supporting Guidance for Headteachers and Chairs of Governors*. Nottingham: DfES Publications.
- Rose, J. (2006). *Independent review of the teaching of early reading*. Nottingham: DfES Publications.
- Rogoff, B. (1998). Cognition as a collaborative process. In W. Damon, D. Kuhn & R.S. Siegler, *Handbook of child psychology: Vol. 2. Cognition, Perception and Language* (5th ed.). New York: Wiley.
- Stuart, M., Dixon, M., Masterson, J. & Gray, B. (2003). Children's early reading vocabulary: Description and word frequency lists. *British Journal of Educational Psychology, 73*, 585–598.
- Tulviste, P (1991). *Cultural-historical development of verbal thinking: A psychological study*. Commack, NY: Nova Science Publishers.
- Watson, J.E. & Johnston, R.S. (1998). Accelerated reading attainment: The effectiveness of synthetic phonics. *Interchange, 57*. Edinburgh: The Scottish Office.
- Wood, D. J. (1989). Social interaction as tutoring. In M.H. Bornstein & J.S. Bruner (Eds.), *Interaction in human development*. Hillsdale, NJ: Erlbaum.
- Ziegler, J.C. & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3–29.

Appendix 1

Word Strings

Rehearsal items: cup van
tree pie
pot fish

1 frog
house
tiger

2 door bear
fairy coat
table mouse

3 dog bus hand
bike cat pencil
pony shoes apple

4 dinosaur bed fox ship
banana toad boy doll
fly box mermaid bowl

5 car hat girl bag monster
letter giant eyes duck rocket
owl bath alien sun crocodile

6 horse cake baby pond teeth chicken
aeroplane robot hen moon road pirate
socks key pig ball garden queen

7 stone elephant window nose phone chair spoon
wolf park feet rabbit bread bee teddy
witch cow goblin feather book face teacher

8 bird flower hair lamp dress rope lorry spider
school king money dragon farmer boat donkey ghost
bun watch sand bottle clothes grass legs paper

9 snake egg button gate monkey balloon jam rainbow milk
kite sausage torch sheep policeman mud farm mouth bucket
boots lion beach cheese robin bath brick goat umbrella

Appendix 2

Phoneme Strings

Rehearsal items: /b/ /æ/
/k/ /u:/
/eI/ /f/

1 /r/
/I/
/m/

2 /v/ /^/
/tʃ/ /l/
/w/ /D/

3 // /u/ /d /
/z/ /a:/ /n/
/e/ /θ/ /i:/

4 /j/ /D/ /s/ /I/
/h/ /:/ /d/ /e/
// /æ/ /t/ /aI/

5 /g/ /e/ /ʃ/ /):/ /l/
/ð/ /u/ /w/ /u/ /d /
/aI/ // /e/ /p/ /a:/

6 /i:/ /tʃ/ /eI/ /m/ /D/ /t/
/h/ /u:/ // /I / /z/ /^/
/d / /I/ /d/ /)I/ /k/ /:/

7 /n/ /æ/ // // /w/ /aI/ /s/
/tʃ/ / /^/ /v/ /D/ // /u:/ /k/
/g/ /l/ /f/ /eI/ /d / /a:/ /j/

8 /h/ /eI/ /k/ /I:/ /ð/ /e / /b/ /u/
/D/ // // /j/ /aI/ /h/ /æ/ /a:/
/tʃ/ // // /z/ /aI/ /m/ /D/ /ʃ / /u:/

9 /a:/ // // /u:/ /p/ /^/ /z/ /I/ /f/ /d/
/w/ /æ/ /t/ /v/ /:/ /ʃ / /n/ /I / /d /
/θ/ /):/ // // /I/ /g/ /i:/ /z/ /e/ /j/