The development and evaluation of software for teaching reading at primary level

Thesis

The development and evaluation of software for teaching reading at primary level

Cathy Lewin B.Sc. (Hons.)

Submitted for the degree of Doctor of Philosophy in Educational Technology at The Open University, Milton Keynes

November 1997

Abstract

This research is an exploratory study and concerns the effectiveness of specific design features in talking book software. Children need to acquire a variety of skills and strategies to become competent readers and these are commonly taught together in the context of meaningful text. Talking book software, replicating real books with additional features such as sound, has the potential to complement this approach. Common features of such software include word pronunciations, hearing sentences read aloud and page turning facilities. These features could facilitate the scaffolding of reading development.

The strengths and weaknesses of commercial software were identified through a pilot study. A survey of practitioners was also conducted to investigate how such software is being used in the classroom and whether talking books could be improved. The outcomes of these studies together with a review of the literature were used to inform an innovative design that was implemented. In addition to common features, the implementation also included sub-syllabic word pronunciations, hints to encourage independent word identification and activities to reinforce specific reading skills.

Two versions of the software were compared using a combination of qualitative and quantitative methodologies including word recognition measures, interviews, observations and case studies. One software version incorporated commonly available features only and the other was a full implementation of the innovative design. Each version of the software was used daily for a period of four weeks by 16 children, aged 6 to 7 years, in the naturalistic environment of their classrooms.

The findings of the study were complex due to variations in learner preferences and reading abilities. Nevertheless, it seems that electronic books can complement teaching approaches in infant classrooms and can positively affect both cognitive and affective learning outcomes. It is evident that children of lower reading ability can benefit from common features alone, such as word pronunciations. Those children using the enhanced software who had already acquired a limited sight vocabulary may benefit from additional features such as reinforcement activities. However, these features were not perceived as being fun by the children in this study. Rather, they were seen as educational tasks. The children who had made the greatest progress in reading development prior to the study made more use of the complex features such as hints to assist them in the decoding of unknown words. This study provided evidence to support the theories of scaffolding and the benefits that can be achieved by aiding learners in the zone of proximal development.

Acknowledgements

For exemplary supervision, sound advice, providing constructive feedback on papers and drafts of this thesis (often at short notice), offering emotional as well as academic support, and being there whenever I needed them: Dr. Ann Jones and Dr. Pat Fung. I couldn't have done it without you, and my deepest and most sincere thanks are not enough.

For commenting on a draft of this thesis and being critical friends: Dr. Eileen Scanlon, Dr. Teresa del Soldato and Erica Morris. An academic life is hectic and your time was much appreciated. For proof reading at midnight: Stuart Thornley and Roger Lewin. For expert statistical advice and patience: Karen Vines.

For inspiration, encouragement and advice on matters ranging from pedagogy and software design to conferences and international ports of call: Jim Dowton, Dr. Teresa del Soldato, Prof. Colin Harrison, Dr. Larry Miller, Prof. David Reinking, Anna Walsh, Janet and Charlie Wells. I must also thank all the staff and children who participated in this research and made it such an enjoyable task. For all your support and providing a convivial working environment: Palitha Edirisingha and Martin Oliver. I would also like to thank Erica Morris, again, for her friendship, support and encouragement, and for taking me up a mountain when I most needed a break from it all.

For technical and administrative support, and always responding promptly to my unending questions and requests: Pauline Adams, Pat Cross, Jackie Harris, Ben Hawkridge, Di Mason, Dave Perry, Hansa Solanki, Matthew Stratfold, Chris Valentine, Olwyn Wilson. For helping me to prepare materials and administer the survey: Christine Vince and Stuart Thornley.

Thank you to the Institute of Educational Technology for giving me the honour of holding the first IET studentship. Thank you to the National Council of Educational Technology for awarding me a student bursary, which enabled me to visit a number of institutions in the USA, and introduced me to an active research community. For providing access to such useful resources and information: Ginn and Company Ltd. and also Bill Bonham at Sherston Software Ltd.

A big thank you to my family and friends who have supported me throughout my studentship. Special thanks to Iris and Ralph Lewin, for providing a home from home for my children, and also to Christine Vince for being there when I needed you. Thanks to my Mum and Dad for all your love, encouragement and support, and for bringing me up to believe that anything is possible. Thank you to Roger for everything.

Lastly, with much love to Jack and Poppy, for the time we couldn't spend together.

Contents

Figures

Chapter 2

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Tables

Chapter 4

Chapter 5

Chapter 7

1. Introduction

1.1 The research problem

This research concerns an exploratory study of the use of computers to support children learning to read, thus supplementing literacy instruction in the classroom. Early reading instruction software was commonly designed to provide drill and practice in basic skills, such as letter recognition and word identification (Balajthy, 1996). Yet it is increasingly recognised that the skills and strategies required to become a successful reader are best learned in the context of meaningful text rather than by focusing on words in isolation (e.g. Ehri, 1991; Harrison, 1992). Such an approach could be supported with the use of talking books, which have developed as a result of recent advances in computer-based speech technologies. Talking book software contains the full colour illustrations and text that can be found in traditional books, but with the added benefit of digitised speech enabling pronunciations for words and sentences to be obtained. Although the formats of commercially available titles vary, common features are considered to be whole word pronunciations, speech feedback for sentences and page turning facilities. This type of software can complement current teaching approaches by providing opportunities for practice through reading meaningful stories. Talking book software also has the potential to scaffold reading development (Wood, Bruner & Ross, 1976).

The range of electronic books is growing rapidly (Balajthy, 1996) and they are becoming increasingly popular in British classrooms (e.g. Collins et al., 1996). Yet little research has been conducted on the use of talking books to support early years reading instruction, particularly in the UK. It is vital to demonstrate that such software is theoretically grounded and to evaluate its effectiveness in classrooms. The designs of electronic books should be investigated to ensure that the benefits of technology are fully exploited. Research on electronic books has also taken little account of affective issues such as motivation.

The main aim of this research was to identify whether and how the design of talking book software might be improved. Are there pedagogical weaknesses in current designs and how can they be addressed? Will an adaptive approach, which is modelled on human tutoring

techniques and facilitating individualised instruction, be more effective? Can this type of resource enhance motivation and enable those children who might be reluctant readers or lacking in confidence to overcome their affective limitations? These issues were examined by implementing an innovative design for talking books. The outcomes can be used to inform future commercial software designs. Furthermore, educators will be better equipped to select the most appropriate formats from the increasing range of electronic titles.

A further aim was to develop an effective evaluation method for reading instruction software. Reinking and Bridwell-Bowles (1991) summarise the methodological limitations of previous research in this field and highlight the difficulties that may be encountered in attempts to overcome these restrictions. For example, opportunities for additional practice or novelty effects of computer use can have confounding effects on outcomes of experimental studies. They suggest that qualitative and quantitative methodologies should be employed and that future research should consider both affective and cognitive issues. Such methods have already been used for evaluating other types of educational software (e.g. Jones et al., 1996) in which affective issues have been considered together with cognitive issues. Yet this approach has not hitherto been applied to the evaluation of talking book software. Other research on the evaluation of educational software suggests that the context in which software is used is another important factor and that it should support and complement teaching aims and approaches (e.g. Crook, 1994). Drawing together the findings from this literature, a framework for evaluation was developed and was used to investigate the effectiveness of the implementation developed as part of this research.

1.2 Outline of thesis

The second section of this chapter presents an overview of the organisation and structure of the research in this thesis.

1.2.1 Chapter 2

This chapter summarises both the theoretical background and related research, through reviewing relevant literature in the fields of reading, scaffolding and computer assisted learning (CAL). Although CAL can refer to a wide range of uses of technology to support

learning, it is used in this thesis to describe educational software packages such as tutorials or simulations.

Firstly, the theory of learning to read is presented, describing the skills and strategies that beginning readers are thought to require and presenting the theoretical models of how children might learn to read. The teaching approaches that complement these models are also summarised.

The second section presents summaries of related issues in learning to read that are relevant to this research. The role of phonics is discussed, together with the importance of motivation, self-confidence and attitudes to reading. A brief description of why children might experience difficulties in learning to read is also given.

Thirdly, the literature on the concept of scaffolding instruction is reviewed in relation to the role of the computer in the classroom. The zone of proximal development is described together with techniques for scaffolding in human tutoring. This section concludes by discussing studies of scaffolding and educational software.

Fourthly, the literature on reading instruction software and benefits of CAL that are relevant to reading instruction software are reviewed. The electronic book format is justified and a critique of related research is presented. The section concludes by summarising the relevant benefits of CAL: multisensory stimulation, practice, independence, enhancing motivation, interactivity and adaptivity. In addition, current research on the use of speech recognition is described. It may be many years though before it becomes both financially and technically feasible to incorporate in commercial implementations that are suitable for schools.

Finally, the conclusions from the literature review are presented. It is argued that reading instruction is most effective when all required skills and strategies are taught in the context of meaningful text. Teachers may provide appropriate direct instruction whilst hearing an individual child read a book or sharing a story with a group of children. Sight vocabularies are more commonly developed through exposure to real books rather than by practising words in isolation. In addition, educational software has the potential to scaffold learning.

Therefore, electronic books have the potential to complement classroom practice. Little research has been conducted in this field, particularly in the United Kingdom, and little account has been taken of aspects such as affective considerations or which design features are most beneficial. This research will focus on electronic books and extend recent research in this field.

1.2.2 Chapter 3

The third chapter describes the methodological approaches undertaken in the pilot study described in chapter 4 and the empirical study reported in chapter 7. As a result of reviewing literature on the evaluation of educational software, a framework for evaluation was developed in order to investigate the use of talking book software in primary classrooms. It is argued, in common with Crook (1994), that the context in which educational software is used is as important as its content. The evaluation process can be improved by conducting such studies in naturalistic environments and eliciting the opinions of both the users and the teachers. In addition, it is desirable that the use of CAL has positive effects on both cognitive and affective outcomes. Thus the framework is divided into four main areas: pedagogical, cognitive, affective and technical. Pedagogical issues relate to how the software supports and complements teaching practice. Cognitive issues concern learning outcomes such as gains in sight vocabularies and the acquisition of required skills and strategies. Affective issues include the effects of such a resource on levels of motivation and self-confidence as well as attitudes to the learning task. Technical issues ensure that the resource is value-added and that the potential benefits of CAL are maximised.

The research methodologies and instruments used in the evaluation are also described. These include: observational techniques, norm referenced tests, criterion referenced tests, interviews, and questionnaires. It is argued that a combination of quantitative and qualitative methodologies should be used such that the results can be generalised to different contexts and that differences between individuals can be explored and described.

1.2.3 Chapter 4

Chapter 4 reports on a pilot study in which commercial talking book software was evaluated with nine children aged 8 to 9 years who were experiencing difficulties in learning to read. The purpose of the study was to identify the strengths and weaknesses of current talking book designs. The basic format, in which each page is replicated with page turning facilities and speech feedback is available for both whole words and complete sentences, was considered to be a major strength. In addition, highlighting the words on the screen as they are spoken appeared to be beneficial and the use of colour illustrations and animations was considered to enhance motivation. The weaknesses that were identified included:

- minimal interactivity
- whole word feedback only
- no mechanism for detecting and correcting errors
- no coaching in the use of reading cues and strategies

In addition, the study identified four styles of interacting with the software. Firstly, children were observed to access pronunciations immediately without attempting to identify the word independently. Secondly, children were observed to make effective use of the software using the pronunciation to either confirm a guess or as a last resort. Thirdly, some children made mistakes that they were unaware of and read for meaning. Fourthly, lower reading ability children who had minimal sight vocabularies were more dependent on computer support and selected nearly every word, reading aloud with the computer. All children exhibited a combination of these behaviours and some of them clearly had a preferred style of interaction.

1.2.4 Chapter 5

An innovative design, which incorporated both the strengths of commercial talking books and addressed the weaknesses identified in the pilot study, is presented in chapter 5. The design was also informed by the theoretical background and related research. The five main modules of the software are described in detail: management module, book module, diagnostic module, activity module and maintenance module. The new features incorporated in the design included multiple feedback modes and activities. Feedback at the

level of onset (the beginning sound in a word) and rime (the remaining sound in a word) was provided to develop phonological awareness. A hint mechanism was incorporated to reduce the dependency on immediate word pronunciations and provide coaching in the use of reading cues, such as the initial letter and the illustration. Additionally, a number of activities were designed to provide coaching in the use of reading cues and to further improve interactivity. Technically, the only way to ensure that spoken errors are detected and corrected in reading instruction software is by incorporating speech technology. However, it was beyond the scope of this research to include such an innovation. It was considered that instructing the users to select sentence level feedback after attempting to read the text themselves would overcome this weakness.

It was also important to consider the extent of commercial software use in British schools and how practitioners integrate this resource within the curriculum. A survey was conducted to investigate these issues with the additional objective of ascertaining if teachers considered that there could be any improvements to current commercial designs. A number of features which practitioners would like to see in future implementations were identified. These included a greater variety of supporting activities to reinforce reading skills, incorporating hints for decoding unknown words and feedback for individual words at the level of onset and rime. Thus, it was considered that the new features incorporated in the design matched the requirements of teachers.

1.2.5 Chapter 6

Chapter 6 describes how the software design was implemented. The environment in which the software was created is summarised and an account of the development approach undertaken is presented under the sub-headings of: prototype, design modifications and the procedure for generating new books. A report is given of three stages of formative evaluation and their outcomes, covering preliminary findings, practitioners' reviews and informal classroom trials. Finally, a number of interaction examples are described to illustrate how the software might be used and how some of the features operate. These include initiating the software, the functions available on each page, the activities and the

maintenance module, which enables the software to be configured to meet an individual's needs.

1.2.6 Chapter 7

This chapter comprises a report on an exploratory study in which the implemented software was evaluated. The aims of the study were:

- to explore the relationship between the software and the classroom environment
- to investigate the effectiveness of the enhanced design
- to explore the effects of the software on motivation and self-confidence
- to verify that the resource is value-added (harnessing technology to maximum advantage)

Firstly, the method undertaken is described. The participants in the study came from three diverse schools and are described in detail. One of the design objectives was that reading software should complement classroom practice and therefore the teaching approaches of the participating practitioners are reviewed. Finally, the procedure followed is specified. Two versions of the software were compared in the evaluation. In the basic version, the texts were replicated on the computer screen with page turning facilities and speech feedback for both whole words and complete sentences. The enhanced version additionally included multiple feedback modes and activities as summarised above. The evaluation involved children using the software daily over a period of four weeks. A combination of quantitative and qualitative methodologies was employed with a matched pairs design. The framework for evaluation was used to structure the study design and analyse the data obtained.

The results are presented in the context of the framework for evaluation. The findings relating to each area of the framework are stated. In addition, five case studies are described in order to explore the effects of the implementation on children who had been identified as either lacking in confidence or motivation. These case studies also give a detailed account of how individuals interact with the software.

The overall results of this evaluation show that both formats of the software are beneficial. The format of the software was well received by both the practitioners and the children who were involved in the study. As was anticipated, talking book software complements current

teaching practices which generally adopt a holistic approach to literacy instruction. The attempts to emulate human tutoring behaviour in the computerised texts were not effective and rarely resulted in the desired outcomes. Activities have the potential to extend the time spent using the software and hence the amount of reading practice undertaken. However, they were perceived by the children as being educational tasks rather than as being fun. These tasks might not enhance motivation as much as was anticipated during the software design phase, but were clearly beneficial.

For the lower reading ability children the basic format was as effective as the enhanced version of the software in developing the sight recognition of key vocabulary. These children made more use of the word pronunciation feature whether they used the enhanced or basic software versions. The enhanced version of the software offered more advantage for those children who had already acquired a limited sight vocabulary. The different features introduced in this implementation provided varying levels of support, some being more useful to readers of higher reading ability rather than those with limited sight vocabularies. The evidence suggests that scaffolding in educational software can be effective and that children make use of these features according to their individual requirements.

1.2.7 Chapter 8

The last chapter of this thesis presents the conclusions of the research. The relevance of this work is justified and the main contributions are highlighted. The implications of this research for practitioners, developers and researchers in the field are also presented. The limitations of the research are specified referring to general restrictions, the constraints of the empirical design and identifying the unanswered questions. Finally, suggestions are made for future related research. These are categorised as modifications to the software design, improvements to the empirical design and alternative applications of the software such as integrating the resource in nursery classrooms.

2. Theoretical background and related research

2.1 Introduction

The research problem stated in chapter 1 concerns the use of electronic books in the classroom to provide supplementary reading practice and raises two questions. Firstly, is this type of educational software appropriate in relation to both classroom practice and current educational theories including learning to read and scaffolding? Secondly, are there any additional features that could improve commercial designs? In the light of these research questions, this chapter reviews the theories on learning to read, related issues and scaffolding in instruction.

Initially, current theories on how beginning readers learn to read are presented. All readers, whether new or proficient, need to utilise a variety of skills and strategies to facilitate rapid word identification and comprehension and these are described in this section. The reading process can be represented by three models (bottom-up, top-down and integrated), highlighting differing theoretical perspectives, and these are then reviewed. The teaching approaches that have been influenced by each of these models are also summarised.

Next, a number of related issues in the literature on learning to read that are relevant to this research are reviewed. These include the role of:

- phonics
- affective factors (motivation, self-confidence and attitudes to reading)
- computer-based resources for children experiencing difficulties in learning to read

Thirdly, the theoretical concept of scaffolding, or providing support and guidance to a learner, is briefly reviewed. The zone of proximal development and scaffolding techniques are described. The relevance of these concepts to educational software design is discussed.

The fourth section of this chapter discusses the contribution that CAL can make in the field of literacy instruction, highlighting the suitability of the electronic book format. A critique of research evaluating the use of CAL to supplement [illegible]tion is presented. These studies have been categorised according to the type of software being investigated:

- those focusing on words presented in isolation
- those using continuous text but without illustrations
- those concerning electronic books

The issues arising from these studies are summarised, including the benefits of CAL and effective design features.

2.2 The theory of learning to read

The process of reading can be broadly defined as understanding the function of symbols and constructing meaning. A fluent reader is defined as being someone who has achieved automaticity (immediate word recognition) in a large vocabulary, but who is also equipped with a number of skills in order to decode unknown words.

It is recognised by most educationalists that no one method of teaching reading has ever been found to succeed with all children (e.g. Smith, 1985; Adams, 1990; Medwell, 1992). This implies that children learn to read in different ways according to their individual learning styles. Children also bring with them a variety of experiences and different levels of knowledge of books when they start formal education. Some may have had little exposure to printed materials. Others may have enjoyed having a variety of stories and nursery rhymes read to them, and may even have begun to recognise some words.

In the next section the theories of how children learn to read are summarised. Firstly, the skills and strategies which beginning readers need to develop are discussed. Secondly, three theoretical models are presented which represent how these skills and strategies might be used: top-down, bottom-up and integrated models. Finally, the teaching approaches which support these models are briefly described, highlighting current practice.

2.2.1 Required skills and strategies

The child who is beginning to learn to read may have already acquired a number of skills such as spoken language, shape recognition and phonological awareness (Smith, 1985; Goswami & Bryant, 1990; Medwell, 1992). Children also need an understanding of the concepts of print. They need to be aware of letters, words, the direction of text, and the

general structure of books. These are not prerequisites for beginning reading and can be developed through the early stages of reading tuition (Wray & Medwell, 1990).

It is acknowledged that a variety of strategies and information must be utilised when attempting to decode unknown words (Goswami and Bryant, 1990; Medwell, 1992; Clay, 1993a). Some practitioners place more emphasis on extracting meaning from the text and may not give direct instruction on word identification strategies such as the use of phonics. However, a child cannot rely on contextual cues alone (Adams, 1990). Early readers must be encouraged to use all available information when encountering previously unseen words, including: visual, phonic, semantic and syntactical cues.

- **Visual cues** can include letter recognition and whole word recognition.
- **Semantic cues** are those which relate to the meaning of the text. This can include: previous knowledge of the subject of the book, assessing whether a word makes sense in context and using the supporting illustrations.
- **Syntactical cues** are those which relate to the structure and grammar of the written text and language.
- **Phonic cues** are those relating to sounds (phonemes) which letters or clusters of letters may have. Initially, children may be aware of the starting sound or onset, and the remaining sound or rime in a monosyllabic word (for example, 'str' + 'ing'). This can develop further into the use of analogies (relating an unknown word to a known word with a similar word ending or cluster of letters).

In order to make sense of what is being read, a beginning reader needs to utilise information from all possible sources by employing as many strategies as possible (Clay, 1993a). Clay (1993b) further advocates that experienced teachers should observe beginning readers carefully and take note of the processes and strategies being used, and will therefore be able to identify those which are under-utilised or misunderstood. Tasks should then be designed to ensure that they reinforce processes that need strengthening and support those strategies that have already been mastered.

2.2.2 The influences of bottom-up models on reading instruction

The bottom-up models are based on building up a complete picture from simple and small components by processing information serially. Gough (1972) proposes that in word recognition the individual letters are recognised initially, then grouped together in syllables or spelling patterns before the whole word is recognised. LaBerge and Samuels (1974)

assert that word recognition becomes automatic for fluent readers, enabling processing capabilities to be devoted to comprehension or the identification of previously unseen words. They considered that known words are recognised holistically and previously unseen words are identified through their component letters (Terry, Samuels & LaBerge, 1976).

The bottom-up models take little account of the role of context in the identification of unknown text. Rather, they place importance in letter recognition and assume that the reading process starts by paying close attention to the words on a page. Teaching methods influenced by bottom-up models promote the development of letter recognition and letter-phoneme correspondences, followed by word recognition and stringing words together to form sentences.

2.2.3 The influences of top-down models on reading instruction

Proponents of top-down models argue that the reading process starts with prior knowledge. They suggest that the print on the page plays a minimal role in confirming the readers' hypotheses about the meaning of the message being conveyed. Goodman (1976) states that "reading is more than precise, sequential identification" (p497). Rather, he perceives that the reading process is a "psycholinguistic guessing game" (p498), utilising semantic, syntactic and phonological knowledge. Accuracy is not considered to be necessary and an effective reader is thought to be able to extract meaning whilst minimising the effort given to guessing unknown words.

Smith (1985) suggests that learning the alphabet or specific words is not necessary, nor is the teaching of phonic rules:

> To expect any readers, and especially beginning readers, to learn and rely upon phonics is to distract them with involved and unreliable procedures which are in fact largely unnecessary. Not only does the fluency in reading demand very little in the way of prior knowledge of spelling-to-sound correspondences, but the practice of reading itself provides the implicit understanding of those correspondences that readers require.
>
> (Smith, 1985, p50)

Smith advocates that reading fluency should be developed in the context of real books, through regular practice and with an emphasis on reading for meaning (comprehending the intended message) rather than pursuing accuracy. With frequent exposure to attractive and

inviting stories, and through an understanding of the meaning conveyed, the children increase their knowledge of individual words and letters.

The influence of top-down models on reading instruction has been a greater emphasis on reading real stories and extracting meaning, and a move away from code based instruction of the past. Top-down models have been criticised for their vagueness in explaining the reading process, yet it is acknowledged that some children can discover the principles of the alphabet without direct instruction (e.g. Rayner & Pollatsek, 1989).

2.2.4 The influence of interactive models on reading instruction

It is now widely believed that the process of reading incorporates aspects of both the bottom-up and top-down models, thus being more accurately represented by an interactive model (Rumelhart, 1977; Stanovich, 1980; Rayner & Pollatsek, 1989).

Stanovich (1980) proposes that fluent readers mainly use a bottom-up approach to reading, supporting the commonly held theory that word recognition is generally automatic. He suggests that when a previously unseen word is encountered an expert reader may use phonological recoding skills or the context. He also suggests that less experienced readers are more reliant on contextual cues to compensate for their minimal sight vocabularies.

Children should be taught to use a variety of strategies to assist in word recognition, but set in the context of real and meaningful books. These strategies include letter recognition skills and an awareness of phonemes, as well as prior knowledge and an understanding of the rules of language. Children need a top-down approach to provide a context for learning and the bottom-up approach to provide them with valuable decoding skills (Figure 2.1).

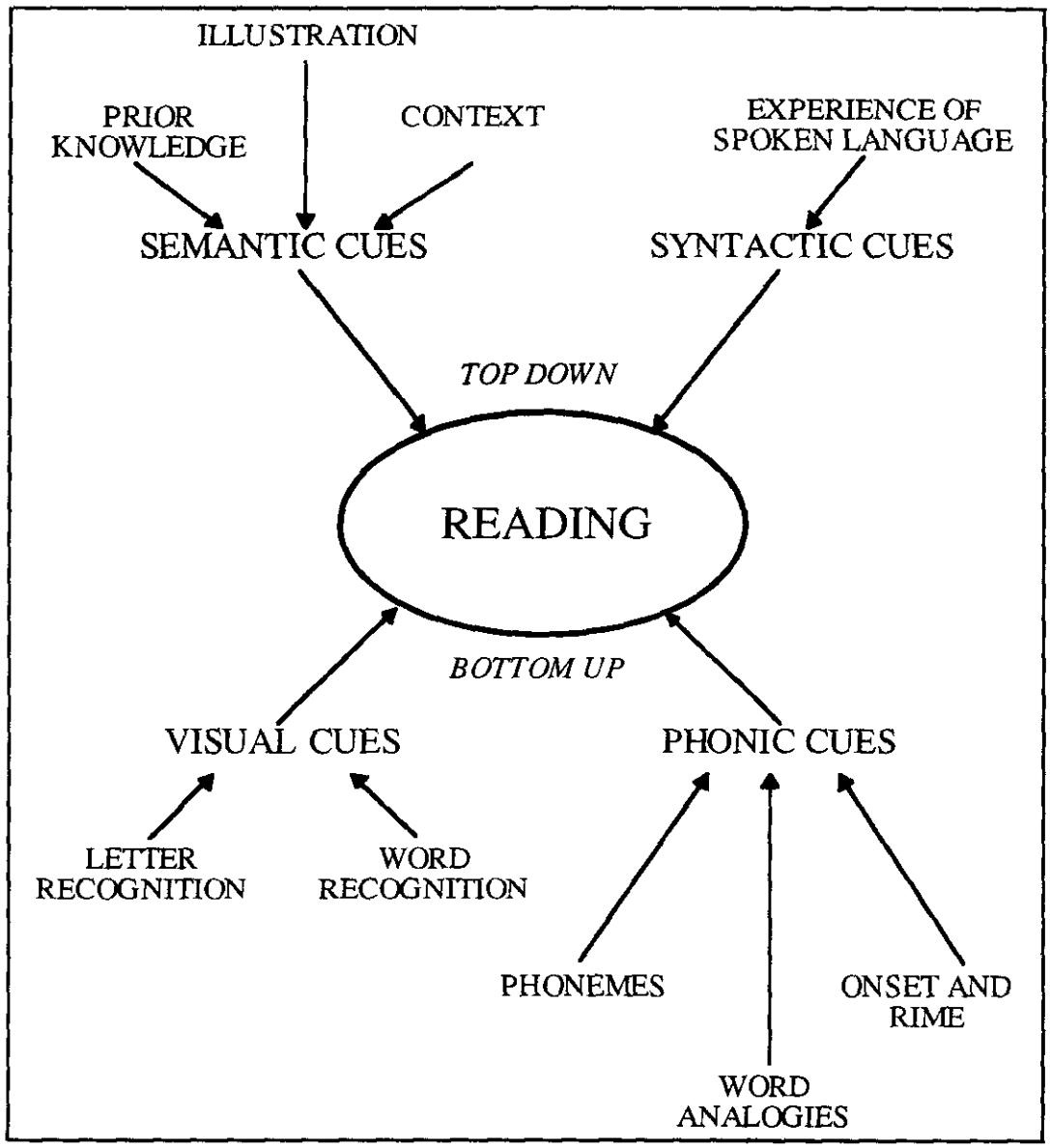


Figure 2.1: Information that a beginning reader might use to identify an unknown word

2.2.5 Teaching approaches

After the second world war, the 'look and say' approach to teaching reading, which was based on the bottom-up model, became very popular. Many reading schemes were developed to support this method and some are still popular. Children learn the words contained in each book through the use of flash cards, before attempting to read the texts. The texts can contain very limited vocabulary, which can be stilted and repetitive, and have been criticised for using unnatural language (e.g. Hudson, 1992). Furthermore, this approach may result in difficulties for children with poor visual memories (Meek, 1982)

who find it difficult to learn words in isolation. Moreover, children may not be provided with the skills for decoding previously unseen words.

Phonic reading schemes, which initially focus on the smallest units of sound and gradually build up to common word families and are also based on the bottom-up model, are similarly open to criticism. The relationship between letters and sounds is not always straightforward and most words in the English language are not phonetically regular. Furthermore, children with poor auditory discrimination find it difficult to acquire phonic skills.

More recently, many educators have chosen to use 'real books' for reading practice, claiming they are more stimulating and interesting. This approach has been influenced by the top-down models of Goodman (1976) and Smith (1971) with their emphasis on reading for meaning. The term 'real books' is used to describe texts that have been written for children's enjoyment. It does not refer to books that have been commissioned for the purpose of teaching reading. The emphasis is on exposure to high quality literature and learning to read by reading. The use of real books is commonly associated with the apprenticeship method of teaching reading. Firstly, a book is read to a child (demonstration). Secondly, the child is encouraged to read with the teacher (joint action). At the third stage, the child reads the book to the teacher but has assistance with any problems that are encountered (supported action). Finally, the child may reach an acceptable standard at a particular level and can read alone and unsupported (individual action). The cycle may be re-iterated with books at a more advanced level.

Other revolutionary methods for teaching reading were devised and became popular. For instance, the 'Breakthrough to Literacy Scheme' (Mackay, Thompson & Schaub, 1978), which is an example of the language experience method. In this scheme, children dictated their own stories and then copied them out, slowly building up their own personalised dictionaries and vocabularies. This scheme formed the basis for the whole language approach to teaching reading, where reading and writing are taught in a holistic fashion.

Publishers of reading schemes have realised the importance of high quality literature and meaningful text and are striving to produce materials that are pedagogically sound. The Ginn

All Aboard reading scheme (Fawcett, 1994) has been designed to support the integrated or interactive model and is organised into three strands. Each of the strands consists of teaching materials that focus on one of the following aspects of learning to read:

- sight recognition
- phonological awareness
- the use of non-fiction to extract information and develop the use of research skills

The scheme also includes a range of workbooks and other support materials that are intended to reinforce learning and provide tuition in the bottom-up skills such as letter and sound recognition. The publishers' intention is to provide a set of resources that can be used in many ways and can be tailored to suit an individual's needs.

Some methods have been devised to address the problems encountered by those who experience difficulty learning to read. Clay (1993a) has developed and delivered many training courses on her Reading Recovery Scheme, designed to help teachers detect and assist those struggling readers at an early age and give them intensive one-to-one tuition. She believes that children need to monitor their own reading and be given guidance on how to detect and self-correct any errors made.

Today, many teachers adopt a mixture of methods, activities and even a mixture of reading schemes and 'real books', and adapt the materials used to support individual learners' requirements (Wray & Medwell, 1990). Medwell (1992) notes that different children require different approaches to learning to read and therefore flexible teaching methods must be employed. Many practitioners also follow some of the principles of the whole language approach, integrating writing, reading and related skills such as prediction. However, they supplement these experiences further by providing direct instruction such as phonics, recognising that these skills are necessary in order to become a successful reader.

One of the most important elements of reading instruction is practice: "a child will need to read extensively in order to become fluent" (Harrison, 1992, p24). Early readers need exposure to a wide range of literature in order to develop their sight vocabulary. It is a common finding that the level of parental support in the home environment can have

implications for the child's reading attainment in the classroom (Topping, 1985; Topping, 1992). The extra practice that children receive through reading with parents at home can be extremely effective.

2.2.6 Summary

- In order to become fluent readers, children need to acquire and utilise a variety of strategies as well as build up a wide range of sight vocabulary.
- Children learn to read in different ways and tuition should be individualised to ensure successful progress is achieved.
- Instructional approaches based on the integrated model, teaching the bottom-up skills such as phonics in the context of meaningful text, are seen to be flexible and effective
- Providing opportunities for additional practice is necessary and can be invaluable in a busy classroom.

2.3 Related issues in learning to read

The effectiveness of reading instruction can be dependent on many related issues, which are now discussed. Firstly, recent research has provided evidence to suggest that developing an awareness of onset (beginning sound of syllable) and rime (remaining sound of syllable) is the most appropriate level of phonics instruction for early readers. Secondly, the importance of practice for beginning readers can be affected by three influential factors: motivation, self-confidence and attitudes to reading. Finally, software developed for teaching reading may have the potential to assist those children who are experiencing difficulties, as well as those who progress without problems.

2.3.1 The role of phonics

Goswami and Bryant (1990) have conducted substantial research into the links between phonological awareness and reading. They comment that previous research has shown that young children experience great difficulty in tasks requiring them to detect and manipulate phonemes. Yet, they can detect rhyme and alliteration or initial sounds before they begin to learn to read.

A skill that can be developed in young readers is the use of word analogies, sometimes referred to as word families. Children may use orthographic sequences of known words (spelling patterns) to decode unknown words containing the same set of letters. Word

families share sounds that are denoted by particular sequences of letters. The simple word families share two letter orthographs which are phonetically regular such as 'at' ('hat', 'mat', 'cat', 'sat'). More complex letter sequences represent irregular word families such as 'ought' ('bought', 'brought', 'thought').

Goswami and Bryant (1990) conclude that children's pre-school exposure to nursery rhymes and stories encourages a sensitivity to alliteration (onset) and rhyme (rime). Some children, on beginning school, may bring with them a degree of phonological awareness. They consider that this is a contributing factor to success when beginning to learn to read. An onset is the initial sound in a monosyllabic word such as 'b', 'br' or 'th' and the rime is the remaining sound in the word such as 'ought'. Every syllable has a rime, and may or may not have an onset (e.g. the word 'and' has no onset). The units of sound that children become sensitive to are thought to progress through stages (Goswami, 1996; Adams, 1990):

- syllables
- onset and rime
- phonemes

Early readers do not develop a sensitivity to phonemes until the later stages of reading development. They are unable to complete tasks involving phoneme detection and manipulation successfully (Adams, 1990). Therefore, phonemic analysis and synthesis (sounding out words and blending the sounds together) may not be an appropriate skill for beginning readers.

Many educationalists have argued against direct phonics instruction in early reading. Smith (1985) is scathing of those who incorporate the teaching of phonics in the early stages. He acknowledges its importance but asserts that the experience of reading eventually leads to an understanding of the correlation of sounds and letters. He does suggest that phonic skills are best utilised when attempting to decode unknown words by comparing them with similar known words, or by using word analogies. Southgate (1984) supports the statements made by Smith regarding the stage at which phonics should be taught, that is later rather than sooner. The English language is not phonetically regular and restricting teaching texts to those words that are regular does not produce interesting material. Also, she suggests that

early phonic training causes readers to slow down at each word, thus inhibiting fluent reading.

Adams (1990), however, points out that previous research has shown that instructional methods based on systematic phonics instruction have been more successful than other methods. Although in these studies the children have been slow to learn initially, they later develop better word recognition and spelling skills than those children who were taught with other instructional methods.

Recognising the initial sound in any word can give an early reader valuable information for cross-checking or decoding (Clay, 1993a). Providing related activities to highlight the basic letter-sound correspondences is commonplace in the classroom. Work on simple word families or spelling patterns can start very early on, without slowing progress in reading or writing and confusing children (Adams, 1990). This then, must constitute 'early phonic training' as indeed could the activities that may have been experienced in the home environment such as singing nursery rhymes (Goswami & Bryant, 1990).

It is suggested that the teaching of phonics does have a place in the early stages of reading tuition. However, the level of teaching should be appropriate for the current developmental stage of each child. It has been shown that instruction on the individual sounds in a word is not necessarily appropriate for beginning readers (Adams, 1990). However, research indicates that phonological awareness at the level of onset and rime is a valuable skill for such children to have and is cognitively accessible (Goswami & Bryant, 1990).

2.3.2 Motivation, self-confidence and attitudes to reading

Research on motivation and learning (e.g. Keller, 1987) suggests that a learner needs to be challenged to enhance and maintain motivational levels. However, the level of challenge should not be too high because it may be frustrating. Furthermore, it should not be too low because the learner may perceive that others have a low expectation of their ability. In addition, learners need both opportunities for progression and feedback to monitor improvement. Motivation can also be enhanced through other means such as interest and enjoyment.

In the field of reading instruction, motivation is commonly associated with the term 'engagement':

> Engaged readers are fluent and flexible. ... Engaged readers also possess cognitive, motivational, and social characteristics that fuel an inclination to read often and widely. ... Engaged readers use reading to satisfy curiosity and as a tool to accomplish tasks.
>
> Afflerbach, 1996, p192

Looking at the pages of a book does not necessarily constitute engaged reading (Almasi, McKeown & Beck, 1996). It is necessary to identify behavioural indicators in order to determine whether or not a reader is engaged, and hence motivated. Observable signs of engaged behaviour (Almasi et al., 1996; Turner, 1995) can be categorised as follows:

- displaying amazement, awe, appreciation of humour
- relating content to personal experience (commenting on text)
- effective strategy use (willing to expend effort on evaluation and monitoring)
- repeating words
- explaining relevance of picture to text
- sounding out a word
- reading text again to bolster comprehension
- predicting what comes next
- interpreting text to a peer
- persisting - remains on task despite difficulty (as opposed to offering excuses for failure)
- using phonetic, contextual and picture cues
- finger pointing to words or illustrations
- obvious eye and head movements (switching from text to illustration)
- recognising miscues and attempting to solve them via another strategy

Research on motivation and learning such as that conducted by Keller (1987) has considered the effects of confidence in relation to the effectiveness of learning. An over-confident learner may gloss over details, whereas a child lacking in confidence may become less motivated through a fear of failure. One of the important turning points in beginning to read is that at which a self-confidence in ability develops and suddenly both the reasons behind and benefits of reading become apparent. Children will often declare that they cannot read, when they already have a limited sight vocabulary. For them, they know that they cannot pick up any book and understand it. At some point they realise that they have the skills to try

to read a book, and the knowledge to attempt to decode unknown word. It is an unofficially recognised state, which is difficult to measure, assess and predict using tools and predefined tests, but easier for an experienced teacher to recognise from a child's behaviour in reading:

> It is not difficult to see when a child's reading is improving. He stops reading word by word, he 'chunks' the meaning in phrases, his eyes turn confidently at the end of the line, and his intonation declares his grasp of meaning. If he stumbles, he retraces his steps confidently.
>
> (Meek, 1982, p82)

Keller (1987) cites several strategies for confidence building. These include:

- ensuring that the learners know the learning requirements
- providing positive and considerate responses to learners' interactions
- creating opportunities for learner success
- giving a sense of personal control by providing corrective or attributional feedback

A reading teacher might offer instant feedback whilst a child is in the early stages of learning to read or even read the entire book to the child first. This contributes to the child's fluency and success in reading, and hence to the building of their confidence. Young children respond very positively to praise and encouragement.

Children's attitudes to reading are as equally important as cognitive ability (Pumfrey, 1979) and over time poor attitude can adversely affect reading attainment. McKenna, Kear and Ellsworth (1995), reporting on a national study on reading attitudes in the USA, comment further on the relationship between attitude and progress in reading:

> growth in ability is linked [in turn, normatively] to one's perception of the value of reading within a particular social context. If its perceived value is low, the development of reading ability will be constrained and beliefs about the outcomes of reading (namely, that the result is frustration) will tend to confirm a normative belief that reading has little value to begin with.
>
> (McKenna et al., 1995, p939)

The authors also state that this relationship becomes more negative over time for poor readers. So experiencing failure in reading adversely affects self-confidence and attitudes to reading. These children spend less time reading for pleasure, get less practice and make less progress.

In the light of this research, the development of motivation, self-confidence and positive attitudes to reading must be seen as important elements the teaching and learning of reading.

2.3.3 Problems in learning to read

Children who experience difficulties in learning to read are affected in other ways (Bryant & Bradley, 1985). For example, their access to all forms of written information is severely restricted, which can effect progress in other subject areas. The skills required in the process of reading are learned through regular practice. Yet, those children with problems read at a slower rate and will not achieve as much as a fluent reader can in a limited amount of time. These children may become more aware of their failure as time progresses and may suffer emotionally.

Slower readers can be more reliant upon reading cues such as illustrations (Stanovich, 1980) and should be encouraged to use them in order to make sense of the text (Hudson, 1992). These children may benefit from increased exposure to reading materials and will require additional practice. Goswami and Bryant (1990) suggest that an insensitivity to phonological awareness is another contributing factor to the difficulties experienced by some early readers. They found that training in rhyme and alliteration, and also in the sound-letter correspondences, can have positive effects.

It can be concluded that struggling readers could make more progress with phonic tuition at the level of onset and rime, and may benefit greatly from opportunities to gain supplementary practice.

2.3.4 Summary

- All readers need phonic skills, but early readers experience difficulty detecting individual sounds although they can be sensitive to alliteration (onset) and rhyme (rime).
- Motivation can be enhanced by providing an appropriate level of support and feedback. Self-confidence is a major factor in self-motivation and a number of strategies for developing confidence may be implemented. Attitudes to reading are equally important and poor attitudes may adversely effect progress.
- It is possible that children who are experiencing difficulties in learning to read would benefit from phonological training, particularly to develop an awareness of onset and rime, and the opportunity to gain additional practice.

The use of computer assisted learning (CAL) to support the teaching of reading in the classroom has the potential to be an effective additional resource. The benefits that might be afforded by the use of technology are now discussed.

2.4 Scaffolding and the role of the computer

Constructivist theories have been influential in the field of education in recent years. In relation to the use of computers in the classroom, the concept of scaffolding is particularly appropriate. Vygotsky (1978) described the zone of proximal development (zpd) as being the difference between what a child can achieve without assistance compared to the degree of success they experience in a task when they receive guidance from a teacher or experienced peer. The extent of the zpd can differ from one child to the next. That is each individual may achieve varying levels of success with the same amount of support. From an alternative perspective, each individual may require different levels of support in order to progress to the next developmental stage.

According to Wood, Bruner and Ross (1976) the tutor should provide support with the elements of a task that are not within a child's capabilities, thus enabling the learner to concentrate on that which they can already achieve successfully. They further suggest this form of tutoring can lead to "the development of task competence by the learner at a pace that would far outstrip his unassisted efforts" (Wood et al, 1976, p. 90). They define six techniques to support scaffolding that provide varying degrees of support:

- Recruitment - promoting an in interest in the task in hand.
- Reduction in the degrees of freedom - simplifying the task.
- Direct maintenance - keeping the learner on task.
- Marking critical features - making the learner aware of the important features of the task.
- Frustration control.
- Demonstration - providing a model of the solution and encouraging the learner to imitate it.

It has been suggested that educational technology can provide scaffolding for a learner, through the pupil-computer dialogue that may arise (e.g. Solomon, Globerson & Guterman, 1989). In many cases, the social interaction can be heightened through collaboration at the computer because children commonly work in pairs or small groups. Crook (1991) argues that this level of interaction during computer based work in the classroom provides the main form of scaffolding support. However, he acknowledges that there is potential for the

"refinement of pupil-computer dialogue" (p. 85) to incorporate the concepts of the zpd and scaffolding techniques.

2.5 Computer assisted learning and reading instruction

A wide variety of literacy software now exists for the British primary classroom, most of which makes full use of the recent innovations in sound and graphics technology. This includes: drill and practice software; talking word processors and talking book software.

Much of the drill and practice software tends to focus on the low level skills required for reading such as letter recognition and word recognition (Balajthy, 1996). It is acknowledged (e.g. Underwood & Underwood, 1990) that there is a place for such software in literacy instruction. However, Reinking (1988-89) comments that a common misconception among software developers is that reading skills should be taught in isolation. This approach does not necessarily complement current classroom practice.

Integrated Learning Systems (ILS) extend the drill and practice model, providing a greater range of course material and a management system that facilitates individualised programmes of work. Although learning gains in numeracy have already been reported, "the value of ILS in the development of basic literacy remains open" (Underwood & Brown, 1997, p192).

Talking word processors are now commonly available and can be used for a variety of activities to support both reading and the related skill of writing (Parham, 1988). Recent research in the UK, for example, the Somerset Talking Computer Project (Bolton, 1995), has highlighted the benefits of such resources, such as making the child aware of mistakes they have made and enhancing motivation.

The number of electronic book titles available has grown considerably in recent years (Balajthy, 1996). This type of software replicates paper books on the screen, with the additional features of speech feedback, usually digitised, and sometimes simple animation. If a beginning reader lacks confidence and requires a lot of assistance, the entire book can be read aloud at a pace dictated by the child. Alternatively, the child may attempt to read the majority of the book, but is able to select individual words to be decoded or confirmed.

Some software has dictionary facilities, enabling a child to look up the meaning of unrecognised words (Parham, 1993). Talking book software is now being produced to support existing reading schemes. Also, there are non-fiction titles available as well as reference materials such as encyclopaedias.

Drill and practice implementations that focus on isolated skills do not closely complement current teaching practices. Talking word processors have a valuable role to play in literacy instruction, but are more appropriate for developing writing and spelling skills than for reading skills. However, talking book software can complement integrated teaching approaches in which the required skills and strategies are taught in the context of real and meaningful text (section 2.2).

Research on reading instruction software is now reviewed and the potential benefits of CAL in this field are identified, in order to determine how the pedagogical soundness of such formats can be improved.

2.5.1 The use of CAL in the field of reading instruction

A selection of studies conducted by educational researchers, which demonstrates the diversity of such investigations, is now discussed. Studies have been designed both to assess accuracy and fluency as measured by word recognition skills, and also comprehension skills. Some have focused upon early readers and others concern those experiencing difficulties in learning to read. The effects of differing forms of computer generated feedback such as whole-word, syllabic and sub-syllabic pronunciations have also been investigated.

Relevant studies have looked at a variety of formats of software. They can be roughly categorised into the following and will be discussed in this order:

- software in which words are presented in isolation and thus out of context
- early computerised texts with little or no illustrations
- electronic books

Words presented in isolation

Early research in reading instruction software concerned drill and practice software in which words were presented in isolation. This approach does not necessarily complement recent classroom practice, but because little research in this field exists overall, it is important to consider such studies.

Roth and Beck (1987) evaluated the effectiveness of two programs for improving word recognition skills and whether or not this led to improvements in reading comprehension. The software consisted of activities that focused a child's attention on the sounds and letter patterns contained in words. It was a long-term study taking place over 8 months. The participants were low to average achievement readers from low socio-economic backgrounds. Statistically significant results were obtained, suggesting that word recognition skills were improved, but more so for the lowest reading ability group.

It is acknowledged that cost-effectiveness in relation to other teaching methods had not been evaluated because the control group had not been involved in equivalent, workbook based activities (Roth & Beck, 1987). Also, the control group were in a self-contained class in a separate school from the intervention group and the possible differences in teaching effectiveness were not addressed. The researchers commented that the participants were motivated by the use of technology and it facilitated greater opportunities for supplementary practice.

Van Daal and Reitsma (1990) also focused on those children experiencing difficulties in learning to read. The software presented words in isolation with two feedback options: whole word and segmented. The segmented feedback given varied according to the word being presented. Sometimes the feedback was at the level of onset and rime (e.g. /c/ /at/) and, for other words, post-vowel segmentation was given (e.g. /ca/ /t/). The study compared the effects of each form of feedback with a control group and identified no difference between feedback modes on words practised with the software. They both positively affected word recognition accuracy when compared to the control group. However, in recognition tests of words that were not practised during the intervention period, there was a

positive trend in favour of the segmented feedback. This implies that segmented feedback equips children better for decoding previously unseen words.

The groups had non-equivalent practice times, although there was no correlation between the time spent using the software and learning outcomes. The researchers observed that children were not selective in their requests for help. They requested hard words equally as often as they selected easy words. It was suggested that poor readers find it difficult to identify when they need help (van Daal & Reitsma, 1990).

Finally, another Dutch team (Spaai, Ellerman and Reitsma, 1991) looked at the effects of whole word and segmented feedback with beginning readers. The segmented feedback in this study was given as individual phonemes, which complemented the approach to phonics instruction in Holland at the time. They found that this form of feedback was not effective with beginning readers and concluded that whole word feedback was beneficial. However, the study involved limited instruction (a total of 24 minutes per child) and has little to say about the prolonged effects of such a resource.

Early computerised texts

Reitsma (1988) compared the effects on word recognition of reading to an adult, reading whilst listening to audio-tapes and reading with computer supported speech feedback. The study focused on beginning readers. The children in each condition read five short stories, especially written for the study, over a period of five days. The materials were designed to provide training on 20 previously unseen words. Those children who read to an adult and those who read with the computer made significant gains in reading fluency of these target words when compared to the other condition. Furthermore, the children of lower reading ability made greater gains than those of higher reading ability. No attempt was made to determine the transferability of these skills other than through the standardised reading assessments, which did not change significantly for any group from pretest to posttest.

Reitsma (1988) acknowledged that the results of the condition in which the children read with the computer may have been influenced by the presence of the observer and the requirement to read aloud. He also noted that this group did not make requests for speech

assistance very often and frequently made reading errors, which were not corrected. However, the gains made were equivalent to the condition in which children read with an adult.

Wise et al. (1989) investigated talking book software that gave speech feedback at different levels: whole-word, syllabic or sub-syllabic (onset and rime) level and a combination of sub-syllabic then syllabic feedback. Children requiring remedial tuition but with average intelligence were selected to form the study group. Three groups were assigned to each of the feedback conditions and used the software daily for 25 minutes during the period of one term. The computerised texts contained no graphics and were reproduced from books and stories recommended by teachers and parents. A control group spent the equivalent amount of time on non-computer-based remedial reading activities.

The gains made on a timed word recognition test were statistically significant for all feedback conditions when compared to the control condition. Importantly, the greatest gains (88% on average) concerned words that had not been requested for feedback. The researchers attributed this to the computer feedback positively affecting phonological decoding skills. This is further supported by the gains made in non-word tests, which indicate that segmented feedback is more effective than whole word feedback in developing these skills. There was an attempt to assess attitude changes using pre-test and post-test questionnaires with both participants and teaching staff, but only for the children in the experimental conditions. The researchers commented that there had been a positive response to the software and that the teachers and children perceived that attitudes to reading had improved.

Electronic books

More recent studies have looked at computerised texts replicating book titles from reading schemes and also from established children's literature. The most relevant projects concern the effects of such resources on word recognition. Additionally, a number of studies have focused on the effects of electronic books on comprehension and they are summarised here for completeness. It is also important to consider the context in which such software is used as well as the effect on learning outcomes. Accordingly, the relevant findings from a project

concerning classroom integration of commercially available electronic books are also summarised.

Davidson, Coles, Noyes and Terrell (1991) evaluated talking book software based on an existing reading scheme in a small-scale study. They were interested in the quality of digitised speech in comparison to the teacher's voice and also, the effect of word pronunciations on sight vocabularies. The study was set in the naturalistic environment of the classroom, which meant that the researchers had less control, particularly over how often the software was used by individuals and for how long. It should also be noted that the control group did not receive any equivalent supplementary practice. The digitised speech was considered to be only slightly less intelligible than the teacher's voice and therefore not a problem. The children who used the talking books when compared to the control group made significantly greater gains as measured by a standardised reading test. It was considered that the regular practice and novelty value, increased motivation and hence a greater interest in reading, could have had an effect on the outcome of the evaluation. In common with the findings of Wise et al. (1989), the greatest gains made in the sight recognition tests were in words that were not practised in the texts.

In a further report (Davidson, Elcock, & Noyes, 1996) describing a similar but larger study, they reported further evidence to support the efficacy of using computers to provide supplementary practice. They identified a significant positive correlation between the amount of practice received by an individual and the gains made. They also commented that no significant relationship was obtained between the amount of speech feedback given and the gains made in word recognition tests. They suggested that recording interaction data for the purpose of informing the teachers of progress may have little benefit.

Miller, Blackstock and Miller (1994) conducted case studies of four children engaged in repeated readings of both commercially available talking books and hard copy story books. These researchers had problems matching students to books that were of a suitable level, due to the limited range of talking books that were available at the time. Requests for help were recorded for both types of media, and for the purpose of comparison it was assumed that a

request for help meant that the word selected was unknown. The number of requests for help decreased on subsequent readings but more so for the electronic books than for the traditional texts. However, indepth analyses of the case study data revealed that student interaction data should be interpreted with caution. A child may read a word incorrectly but believe it to be correct or avoid requesting help for other reasons. Alternatively, they may use the feedback facility to confirm a prediction. Thus, if a child's use of talking books is not monitored occasionally a misleading picture of their progress may emerge.

McKenna, Reinking, Labbo and Watkins (1996) also looked at beginning readers. They found that electronic books positively affected sight word gains when compared to a control group who participated in similar tasks, such as reading to their teacher and alone. They also found that gains made in sight words were not related to the requests for help that had been made. They commented on the styles of interaction, which they termed under-accessing (not requesting help when needed) and over-accessing (requesting help for words that are already known). Furthermore, this study also raised issues concerning reading ability level. Gains were only observed in those children who had already acquired a limited sight vocabulary. Finally, an attempt was made to provide coaching in phonemic awareness in the software design, but the children were not successful in decoding pseudo-words with similar rimes at posttest.

A number of studies have focused on the effects of electronic books on reading comprehension. Matthew (1997) reported significant differences in comprehension as measured through story retelling for 8 and 9 year olds in favour of reading electronic texts when compared to traditional texts. Greenlee-Moore and Smith (1996) also obtained similar results, but only with passages that were long and difficult, and not with those that were short and simple. An earlier similar study (Feldman & Fish, 1991) found no differential effects in comprehension levels for CD-ROM electronic books and traditional texts when the participants were of secondary school age. Standish (1992) also found no difference in improvements in comprehension when comparing the use of electronic texts to that of traditional books. However, she concluded that this was due to methodological limitations in

her study design and suggested that further research should be conducted with more participants and over a longer period of time.

A further study concerning classroom integration of commercial talking books was conducted by DeJean, Miller and Olson (1997). The research objectives were to assess teachers' attitudes and pedagogical changes when talking books were incorporated into the classroom. They were not concerned with the effects of such a resource on children's learning. Two important issues were raised which are relevant to this study.

Firstly, Broederbund's Living Books Series (a popular range of electronic books), which incorporate many hidden animations in each page illustration, were seen both by the teacher and the children to be "entertaining and enjoyable". However, they were not integrated into normal classroom activities because they were not perceived to be of academic value by either the children or the teacher. Underwood and Underwood (1996) report similar observations on the format of Living Books, stating that the children spent most of their time activating the hot spots and very little time reading the accompanying text.

Secondly, in order to maximise opportunities to use computer-based resources, the teacher in the study elected to organise the children in pairs when reading talking books. The researchers observed that power struggles arose between some pairs, not only over the use of the mouse, but also in the way in which the books were read. This resulted in the dominant member of the pair satisfying their own needs at the expense of the other member. Evidently, the use of talking books in pairs might not be the most effective approach to adopt.

2.5.2 Implications of recent evaluations

There are many issues arising from the above studies. There appears to be a relationship between the reading abilities of children and the difficulty level of the computerised reading software (and particularly texts), which can affect the amount of progress made. It would be useful to establish what features are beneficial at each stage of early reading acquisition. Research has shown that segmented feedback positively affects phonological decoding skills, except when the feedback is given as individual phonemes (Spaai et al., 1991). This

is consistent with the literature on phonics instruction (e.g. Goswami & Bryant, 1990). Yet, little research has been conducted on this form of feedback in electronic books. Reading software may be used in a variety of ways and styles of interaction that have been observed so far are under-accessing and over-accessing. These styles of interaction may not be beneficial and could be addressed in future software implementations.

2.5.3 The affordances of CAL

The beneficial features of CAL for reading instruction are now discussed, based on the outcomes of the evaluation studies and a review of literature on the potential benefits educational software.

Multimedia and multisensory stimulation

Most of the recent software designed to assist in the teaching of reading exploits the use of sound technology. Instant speech feedback can be given for individual letters, whole words, sentences and complete stories. This may be either synthesized speech, which is computer generated, or digitised speech in which human speech is recorded and converted into digital format.

The issues concerning instant speech feedback are complex and research evidence has led to conflicting implications. Reinking (1988-89) identified a number of misconceptions with regard to the development of language arts software. He stated that immediate feedback may not always be beneficial for the learner, although the example cited concerned the assessment of student responses to comprehension questions. However, in a later review of research on literacy and technology (Reinking, 1995), he comments that accessing an unknown word in order to obtain its pronunciation can have a positive effect on learning. Balajthy (1989) considers that the provision of instant feedback or ongoing monitoring and diagnosis in language arts software can be far more effective than traditional classroom methods where immediate feedback is not always given. Many studies suggest that instantly accessing word pronunciations does not necessarily improve sight word recognition of the vocabulary selected (Davidson et al., 1991; Wise et al., 1989). Rather, the findings indicate that gains are achieved in words not accessed or even contained within the computerised texts.

Computer software also has the ability to exploit simple visual effects, such as highlighting the word being spoken. This feature draws the reader's attention to the relationship between the spoken word and the written form. Furthermore, it can reinforce concepts such as directionality of print and the use of spaces to delimit words in a sentence (Medwell, 1996).

One feature generally associated with multimedia software is animation. Sometimes animations are provided either as a reward for completing a task or are incidental to the educational aims of the resource. Some researchers have cautioned against the use of animations as rewards, which can distract users from the main elements of a task (Sewell, 1990; Costanzo, 1989; Balajthy, 1989). They state that a reliance on computer graphics can be a serious impediment to educational goals and limit the time spent on task. Furthermore, animated illustrations in electronic texts may decrease the amount of reading that occurs. This can happen because children spend more time interacting with the animations and pay less attention to the words (Reinking, 1988-89; Underwood & Underwood, 1996). Nevertheless, animation can be beneficial as a supplement to instruction (Milheim, 1993). In electronic books, for example, animations can be used to illustrate the meaning of a word (Reinking, 1995), enhancing the display of text and ensuring that the use of technology is value-added.

Practice

The use of the computer lends itself to reading practice (Reitsma, 1988; Davidson et al., 1991; Wise et al., 1989). Through practice a child may widen their vocabulary and hence improve fluency and comprehension. Davidson (1994) concludes from her studies that the use of computer assisted reading software as a supplement to normal classroom activities offers a cost-effective additional resource for the classroom. It has been noted that in research into this field so far:

> A consistent finding from the investigations of reading curricula is that brief, but regular, computer based reading lessons can enhance reading achievement. The results of these investigations, however, are based most often on the use of computer-based activities that supplement rather than replace conventional reading instruction.
>
> (Reinking & Bridwell-Bowles, 1991, p316)

By harnessing motivational features (discussed below), users of software designed for supporting the teaching of reading are encouraged to spend longer practising the required

skills. Many researchers assert that practice is required for developing automaticity of sub-skills, such as word recognition in the case of reading, and advocate computer use (Underwood & Underwood, 1990; Balajthy, 1989).

Independence

A computer-based resource may be perceived as being infinitely positive and patient, providing a private study environment that is under the child's control (Parham, 1993; Wepner & Feeley, 1993; Gardner, 1997). Schofield, Eurich-Fulcer and Britt (1994) found that requesting help and making mistakes in a computer-based learning environment were perceived by the students as being less obvious to both their peers and their teacher. The students were therefore less embarrassed. Furthermore, the students were able to proceed at their own pace.

Independence is a desirable feature, but the use of computers by children in the classroom should still be monitored to ensure that benefits are obtained (Miller & Olson, 1995; Hall & Rhodes, 1986).

Motivation

Many researchers have observed that using a computer-based educational resource can be motivational (e.g. Gardner et al., 1994; Schofield et al., 1994). However, as Balajthy (1989, p86) points out, the use of computers alone cannot "be considered a magical motivational cure". Careful consideration should be given to software design to ensure that effective motivational features are incorporated.

Malone and Lepper (1987), following an investigation of games software, produced a taxonomy of intrinsic motivational factors for software, which includes challenge and curiosity. Challenge can be stimulated by introducing uncertainty and by providing performance feedback. They comment further that feedback is more intrinsically motivating when it is frequent, clear, constructive and encouraging. They raise the issue that learners may be affected differently by specific motivational features:

> For example, increasing the difficulty level of an activity or introducing competition may enhance interest for children with high ability and self-confidence, yet undermine the motivation of children low in ability and self-esteem.
>
> (Malone & Lepper, 1987, p246)

Harnessing the motivational aspects of computers can contribute to increased concentration and what appears to be longer time spent on task (Schofield et al., 1994; Hawkridge & Vincent, 1992). This phenomenon, though, is difficult to measure accurately (Issroff, Jones & Scanlon, 1994). In one study of the use of microcomputers in schools it was noted with a caveat that:

> Children were observed to maintain concentration over longer periods of time than might have been expected when working on the same topic with a different medium. However, teachers need to maintain a check to ensure concentration is sustained and that the activity remains relevant to the educational aims.
>
> (Hall & Rhodes, 1986, p5)

Issroff (1991) stresses that educational software should incorporate motivational features, harnessing them to improve the efficacy of computers in the classroom. It can be concluded that the motivational benefits, which can be offered by computer-based resources, should be fully exploited. However, features that may enhance the enjoyment of software, such as animations, should not detract from the pedagogical aims of the resource.

Interactivity and adaptivity

Interactivity, whereby information is exchanged between the software and the learner, has many potential benefits. It is generally considered to be a major contributory factor to the overall effectiveness of computer-based learning environments (Bork, 1995; Milheim, 1995-96). Not only can a learner receive immediate feedback to their actions, but by monitoring student progress and failures, computer instruction may be individualised. Active involvement with the reading process can greatly enhance the development of reading strategies and the level of comprehension. Furthermore, teaching staff can be provided with both formative and summative assessments in order to facilitate effective classroom management of users (Perfetti, 1983; Reinking, 1995).

Adaptive instructional systems model students' behaviour, such as success and failure, in order to set tasks of appropriate difficulty. This level of adaptivity can provide some form of 'intelligence' without necessitating the incorporation of sophisticated techniques more

usually associated with true ITS systems, such as instructional planning and knowledge representation. It is feasible to develop adaptive CAL software that can be individualised and react to a learner's needs:

> With the appropriate software, the computer can analyse student responses and make instant decisions concerning accuracy, style of response, nature of errors and can react according to the needs of individual students on the basis of the responses provided by the students. In other words, the power of technology is such that it has the potential to act as an individual tutor, responding to different learner characteristics, thereby providing genuine individualised instruction.
>
> (Sewell, 1990, p5)

Speech recognition

Speech recognition technology has advanced greatly in recent years and is already being incorporated into commercial software such as word processors. Although it is beyond the scope of this current research to investigate the benefits of this feature, it may have a great impact on future developments in the field of reading software. The recent research in this area is briefly summarised.

The potential of such advanced technology in the field of computer assisted reading software has already been recognised (Mostow, Roth & Hauptmann, 1994). However, it is accepted that it may be several years before the financial investment required can be justified (Kantrov, 1991). Furthermore, it is unlikely that speech recognition systems will be 100% accurate within the short term (Lange, 1993).

Wetzel (1991) summarises the requirements of an effective speech recognition system for teaching reading. It should:

- recognise the speech of a variety of individuals
- require little pre-training
- accept continuous speech
- be reasonably accurate
- screen background noise

Despite accuracy problems, Mostow et al. (1994) point out that such errors are not necessarily detrimental. They have adapted the interface in their implementation such that only corrective or confirmatory feedback is given. For example, the software never informs a child that a word has been read incorrectly because it may be a speech recognition error.

Rather, it asks the learner to read the word again. The researchers and developers have recognised the limitations of current technology and addressed them by adapting the pedagogical strategies employed.

It seems from the evidence so far that speech technology could have a lot to offer educational software of the future. It will be particularly useful for computer assisted reading software, enabling such a resource to 'listen' to its user, assess reader accuracy, and offer more appropriate and individualised interventions.

2.5.4 Summary

- The use of sound and graphics is important in order to emulate reading real books *with a tutor.*
- Motivational advantages that can be provided with such technology must be exploited. Simple animations can enhance the motivational aspects of the software without undermining the pedagogical intentions and provide an element of fun.
- It is argued that software designed to support the teaching of reading must enable a child to participate in unsupervised practice and receive training in the skills required to become fluent.
- Providing a private and patient learning environment could be extremely beneficial. Also, enabling a student to progress at their own pace is desirable.
- Children learn to read in different ways, and normal classroom tuition is generally individualised. The software should aim to be adaptive in order to emulate this human tutoring technique.
- Speech recognition has the potential to improve software applications for teaching reading in the future.

2.6 Conclusions

CAL should be underpinned by a theoretical framework in order to ensure pedagogical soundness. Therefore, the first three sections of this chapter have reviewed the theories of learning to read, related issues and scaffolding techniques. The conclusions reached, in common with Harrison (1992), are that neither the top-down or bottom-up reading instruction approaches are satisfactory on their own; rather the evidence points to an integrated approach. It has been argued that reading should not be taught by focusing on specific skills in isolation, but that instruction should be set in the context of real and meaningful text. Additionally, phonics instruction is ir ..tive (Adams, 1990) and the debatable issues are when and at what level to provide it ..e strongest arguments currently are that phonics should be taught at the level of onset and rime and this is supported by

evidence from Goswami and Bryant's studies in this area (1990). Specific training at this level can be particularly beneficial for children experiencing difficulties in learning to read. A reading scheme that complements the integrated approach and incorporates phonological awareness at the level of onset and rime has been identified: the Ginn All Aboard scheme (Fawcett, 1994).

An alternative educational theory deriving form constructivism was also discussed. The tutoring technique commonly referred to as scaffolding (e.g. Wood et al., 1976) enables a learner to develop and progress beyond the level of achievement that they can experience without such assistance. It was suggested that pupil-computer dialogue within educational software may provide this to some extent.

The fourth section of this chapter reviewed the role of CAL in the field of reading instruction. Firstly, types of software designed to support literacy instruction that are already in use in classrooms were reviewed. It was argued, in accordance with Reinking (1988-89), that electronic books are more appropriate for supporting reading instruction than drill and practice programs, which generally focus on isolated skills.

Next, the related research on the use of CAL for supporting the teaching of reading was reviewed. The conclusions are that computer assisted reading software can make a valuable contribution to the teaching of reading. Furthermore, it appears to have been particularly beneficial for those experiencing difficulties in learning to read (Wise et al., 1989) and readers who have only just started to develop their sight vocabularies (Reitsma, 1988). There is some evidence to suggest that providing segmented feedback can improve phonological skills, but its effectiveness had not been investigated in recent electronic book formats. There has been little research in the use of talking book software in British schools, even though its use is becoming more and more popular. Few researchers in this field have taken account of affective factors such as motivation, self-confidence and attitudes to reading. No one has investigated the efficacy of incorporating coaching in the use of reading strategies and cues in this software format.

The affordances of CAL were also reviewed, concluding that, although a number of technological features have already been incorporated successfully in commercial electronic books, there are still improvements that can be made. In particular, the adaptivity of such software is currently limited and there are many additional adaptive features that could be incorporated in order to emulate a supportive tutor. Thus, the use of electronic books could be tailored to support individuals' needs.

These conclusions have influenced the design of the electronic books developed for this research (chapter 5 and chapter 6). These are based on the Ginn All Aboard reading scheme, but also incorporate adaptive features, coaching in the use of reading strategies and cues, and phonological feedback at the level of onset and rime. The design was also informed by a pilot study (chapter 4) conducted to identify the strengths and weaknesses of current commercial electronic book designs. Related research in this area (e.g. Davidson, 1994) has shown that changes in reading outcomes may be affected by many different variables, which can be hard or impossible to control. These findings will be used to inform the research methodology that will be adopted and is discussed in the next chapter.

3. Research methodology and instruments

3.1 Introduction

The theoretical background concerning theories of reading and the use of computer assisted learning together with research in the field of language arts software were reviewed in chapter 2. This chapter describes the evaluation methodology employed in this research and how it was implemented in the pilot study (chapter 4) and main empirical study (chapter 7).

Firstly, literature relating to the evaluation of educational software is reviewed and summarised because the research project concerned the use of software in the classroom. The process of ascertaining the effectiveness of computer assisted learning (CAL) can be complex and it is imperative to consider the outcomes of previous research and past experiences in this field. The primary issues that must be considered when conducting an evaluation are highlighted. Arguments are given for considering the content of educational software and the context in which it is used. These aspects are discussed and a rationale for adopting a holistic evaluation approach is presented.

Secondly, the outcomes of this review are used to inform a framework for evaluation of computer assisted reading software and four domains are defined: pedagogical, cognitive, affective and technical. Pedagogical aspects of the framework concern the context in which educational software may be used and its suitability to support current teaching approaches and practices. The cognitive domain relates to learning outcomes and the appropriateness of the content of the software. Although learning outcomes are linked to pedagogy, they are categorised separately for convenience. Affective issues are associated with motivation, confidence and positive attitudes and their effects on learning outcomes. Aspects of the technical domain refer to the content of the software and are intended to ensure that the use of such resources is value-added.

Thirdly, the research methodology that was employed within this framework is described in more detail, including the study design and instruments used for data collection. A

combination of qualitative and quantitative techniques was used, drawing upon the methodologies employed in related studies and their strengths and limitations.

3.2 Evaluating computer assisted learning: a literature review

It is important to demonstrate that CAL is underpinned by a theoretical framework within an educational context and that learning outcomes can be improved. It is accepted by designers, researchers and practitioners that evaluation is imperative, but lack of available funding and time can mean that it is not as thorough as it should be (Balajthy, 1989; Blease, 1986). Blease states that too few packages, commercial or otherwise, are subjected to trials in classroom or naturalistic settings. Information needs to be provided to highlight what contributions could be made by such software and whether or not it can meet its aims as a teaching resource. It needs to be demonstrated that the software 'works' in its intended context. The issues that need to be considered include:

- when should an evaluation take place
- how should an evaluation be conducted
- who should be involved in the process
- what should be evaluated

When should an evaluation take place? Evaluation should occur at several stages of the educational software development life cycle. Formative evaluation takes place during the development phase and its purpose is to ensure that the content and structure of the materials are as effective as possible. Field testing in a naturalistic environment is a valuable exercise because feedback from students and teaching staff can be extremely informative regarding the intended use and context. Summative evaluation is carried out when the software is complete and is designed to establish whether or not the software meets its aims and objectives. One issue, which can be associated with many educational software programs, is that of when to test for the effects of software use on a learner. Certainly it is necessary to conduct tests immediately after software use because a reliance on delayed posttests alone would lead to inconclusive results. Delayed posttests may be used to determine whether any improvements are sustained but will be affected by the additional learning that will have inevitably taken place.

How should an evaluation be conducted? Substantial investment in educational software evaluation has occurred in the United States (Heller, 1991). Many forms and check-lists have been developed, which are based upon a variety of criteria, including pedagogy and social attributes. Squires and McDougal (1994) state that the check-list approach may be suitable for assessing the attributes of the software itself or the content, although they are critical of their use overall. The use of check-lists can be restrictive and can become quickly out of date as rapid technological changes occur. A more comprehensive approach, although being more challenging, is one that also considers classroom management, curriculum integration, pedagogy and learning in order to provide a holistic view (Squires & McDougal, 1994).

Who should be involved in the process? Evaluations should involve educators, researchers and the learners, as well as the software designers, all of who have a valuable role to play (Squires & McDougal, 1994). The educators or teachers may have an indepth understanding of the intended learning objectives, but more importantly may strongly influence the way in which the software will be used. They could adapt their teaching methods in order to integrate the use of software effectively, or supplement the use of such tools with personal interaction. For example, Draper, Brown, Henderson and McAteer (1996) observed undergraduate students who were using simulation software. In this study, the teacher ensured that the conceptual points had been understood by engaging the students in a tutorial dialogue upon completion of the software exercises. The presence of a researcher or teacher can result in optimistic conclusions being drawn (e.g. Crook, 1994). It is more realistic to consider these human interactions together with the software use and acknowledge that both make a valuable contribution to the learning outcomes that may be achieved.

What should be evaluated? Hammond (1995) suggests that early research and evaluation of CD-ROM software in schools has focused on the program (the content) and not the learner, and that many software reviews are pedagogically weak and uncritical personal testimonials. He concludes that the focus should be upon the learners and not the technology. Reiser and Kegelmann (1994) advocate that instructional software evaluation

must include observations of learners interacting with such a resource. Moreover, it has been argued (Crook, 1994; Venezky, 1983) that the use of software in a classroom must be considered in relation to the broader teaching and learning environment (the context). Nevertheless, the quality of the content and design cannot be ignored.

An important question to be asked of any educational software is whether or not learning outcomes can be achieved and how effectively (Jones et al., 1996; Draper et al., 1996). Many researchers and educational bodies want to know how effective learning is with a particular resource in relation to other teaching methods (e.g. Mason, 1995). Comparative studies with innovative technologies can be awkward as it can be difficult finding equivalent non-computer-based tasks. It has also been argued (Jones et al., 1996) that the magnitudes of learning outcomes and the causal links between this and the software used are hard to establish. Establishing the effects of the use of educational software can be problematic and results may depend on many factors that can influence learning such as students' motivation and prior experiences.

Given the difficulties and issues raised above, it is proposed that both formative and summative evaluations must be conducted and should involve practitioners, software designers, researchers and the learners themselves. An evaluation must be holistic and consider both the content of educational software and the context in which it is used in relation to the effects on learning outcomes. In order that such research is rigorous, it may be necessary to use a variety of research methods and tools. CAL resources can differ according to their specified aims and intended subject area. A detailed review of the software content and context of use is now presented in relation to the field of reading instruction.

3.2.1 Content

The content of the software should complement the pedagogic intentions of the educator and should have the potential to improve learning outcomes although the difficulties in demonstrating this have already been highlighted (section 3.2). Classroom practice in reading instruction can vary and software that is designed to support one method may not be appropriate for others (chapter 2, section 2.2.5). The current trend in many schools when

teaching reading is to adopt an integrated approach. Software based on the drill and practice pedagogy, which presents a child with isolated words out of context, would not be suitable for such a teaching method.

Educational software should incorporate principles of motivational design (e.g. Malone & Lepper, 1987) such as providing feedback on progress. Benefits relating to the affective domain can be more difficult to measure. Many researchers have developed questionnaires designed to collect data on children's attitudes to computers (Todman & File, 1990) and motivation to read (Guthrie, McGough, & Wigfield, 1992). These approaches could be adopted to ascertain an individual's attitude to reading software, as well as more general attitudes and levels of motivation. This could be triangulated with information from the teacher and observations by the researcher.

The interface of any software should be usable, efficient and easy to learn (Johnson, 1992). Substantial research has already been conducted in the field of human computer interaction (HCI), in which interfaces are studied in detail in order to make improvements. As noted by other researchers from outside the field of HCI (e.g. Laurillard, 1993), the operation of the software should be intuitive (operational transparency). It should not interfere with the educational aims and objectives of the resource. This now achievable and should be incorporated in future software. The design of educational software should build upon this substantial knowledge as well as fully exploiting technological advances and innovations. The incorporation of sound is already becoming a well-used feature, as advances in storage capability more readily enable its use. Speech recognition, although still in developmental stages, may be a major element of interface designs in the near future. It is already being evaluated for its potential to revolutionise software for teaching reading as well as other curicula areas.

3.2.2 Context

The way in which educational software is used in a teaching environment can vary enormously. It can depend on the type of institution and target group, the resources available, the specific teachers and their previous exposure to information technology, and

the intended aim of its use. This can be examined in terms of who will be involved, that is, whether it is to be used by groups or individuals, and how it should be integrated into the curriculum.

It is common practice in British primary schools that computers are used by groups of children rather than individuals. It is extremely difficult to measure the benefits achieved by individuals in any collaborative projects whether or not the work is computer-based (Issroff, 1992). Issroff discusses many of the problems inherent in evaluating collaborative learning with computers, pointing out that the computer itself was not designed for group use, and encourages individual behaviour. She comments that the age of the children, issues of gender and friendship groups and the time spent on the activity can all affect the results of such an evaluation. For the purpose of this research individual use only is considered.

Laurillard (1993) suggests that educators should adapt their teaching approaches to make the most effective use of technology. However, current teaching practices and prior experiences may determine the way in which such a resource is integrated (Miller & Olson, 1994). In addition, teachers may not fully understand the designers intended use or how technology may best be used to improve learning. Many teachers have reported that they have little time to participate in information technology training or to evaluate software packages (Dupagne & Krendl, 1992). As with teaching reading, some students require different approaches in order to learn most effectively. The issues relating to the context of delivery can be addressed through training and support and by ensuring that software based resources are flexible and adaptable.

3.2.3 Discussion

Factors to be considered in evaluating educational software, the content of educational software and the context in which it is intended to be used, have been discussed in isolation. However, these issues are inter-related and need to be considered in tandem. The potential benefits to be derived from computer-based resources are dependent on the pedagogical aims of the educator as well as the quality of the software itself. As Laurillard (1993) states:

> The most stunning educational materials ever developed will fail to teach if the context of delivery fails. Conversely, good delivery can retrieve poor materials.
>
> (p220)

The process of evaluation can be problematic. The results of any evaluation may be attributed to many different factors such as the way in which it is used, or the students previous knowledge and experiences. The design of software may be such that it suits particular needs or specific learner styles but cannot be universally employed in a classroom. The content of educational software must be relevant, of an acceptable standard and be guided by specific educational aims, but must also complement classroom practice.

In view of this, it is imperative that a holistic approach be taken to the evaluation of educational software. It is not sufficient to verify the technical adequacy of educational software without considering the needs of both the practitioners and the users. However, it is not simple to advocate one set of criteria or one procedure to follow. Every evaluation will be dependent upon the subject matter, the way in which the software is integrated within the curriculum, and the educational environment and student group with whom the resource is being used. The outcomes of such evaluations will also be influenced by the subjective views of the evaluators and educators.

Rapid advances in technology and the effects of this upon delivery methods and pedagogy imply that evaluation techniques must also evolve. In order to structure this process in relation to the evaluation of computer assisted reading software and ensure that it is rigorous, a framework has been devised and it is described below. This should guarantee that the evaluation is thorough and addresses both the quality of content and suitability to intended context. Furthermore, it should facilitate close collaboration of the educators, software designers and researchers as well as the students themselves.

3.3 A Framework for evaluation

The framework that has been produced classifies the evaluation issues discussed in the previous section into four domains: pedagogical, cognitive, affective and technical. The data collection methods used when applying the framework for evaluation are discussed further

in section 3.4. Here the factors to be evaluated in each domain of the framework are specified in more detail.

3.3.1 Pedagogical

- Teaching approach: the teaching strategies utilised should support those already employed in classroom activities by the teacher. The approach may relate to existing teaching materials; for example, it may be based on an existing reading scheme.
- Reading strategies: the software should facilitate and further develop the use of reading strategies, such as cues to aid decoding. However, it may be designed to provide coaching in specific low-level skills such as word recognition. The software may detect which strategies are under-utilised, hence offering appropriate coaching tailored to the needs of the individual.
- Related skills: there may be support for the closely related skill of writing.
- Suitability: the instructions, tasks, animations and vocabulary should be suitable for the target age range. The tasks should relate to the intended learning outcomes and must be complex enough to be challenging.
- Student autonomy: the child should be able to control the pace of learning, and progression should be allowed.

3.3.2 Cognitive

- Reading age: this should increase and the improvement should be sustained.
- Reading strategies and cues employed: the number of these should be increased and the improvement should be sustained.

3.3.3 Affective

- Motivation: the use of sound and colour should be at a satisfactory level to engage the child in the use of the software without being a distraction. Success should be rewarded in a way that enhances motivation.
- Confidence: the reading confidence and attitude to reading should be improved and be sustained. A positive result may be achieved through use of the software even if there is no measurable improvement in reading ability. If a child becomes more confident and enjoys reading then this may lead to further practice, which could be more beneficial.

3.3.4 Technical

- Value-added: the technical possibilities should be exploited. The use of speech to associate the written word with its sound, to give instructions, instant feedback and encouragement should be maximised. Records of progress should be available.
- Operational transparency: the use of the software should be easy and intuitive.
- Instructional adequacy: instructions should be clear. The child should be able to use the software independently.
- Cost-effective: the CAL should be cost effective in relation to other methods and existing materials.

3.3.5 Summary

The framework that has been presented should not be seen as a check-list. Rather it is a set of guidelines to structure the evaluation of reading software. This should ensure that the context in which it is used is considered in conjunction with both the content and technical adequacy. Each specific evaluation will naturally differ according to the schools, teachers and children who were involved. Use of the framework should ensure that the evaluation is holistic, thorough and methodical. The methodology that was employed in the pilot study (chapter 4) and main empirical evaluation (chapter 7) is presented next.

3.4 Methodology

The methodology and research instruments used in the evaluations conducted in this thesis are now described. Related research was reviewed in chapter 2 (section 2.4.1) and the methodological limitations arising from these studies are briefly summarised first. Next, the general design is presented for both the pilot study and the main evaluation. This is followed by descriptions of the instruments used: both systematic and informal observations, standardised reading tests and criterion referenced tests, interviews and questionnaires, and case studies.

3.4.1 Methodological limitations

Many limitations of research into computer-based resources for teaching reading have been highlighted by Balajthy (1989). He states that researchers have focused on drill and practice software built around the programmed instruction model rather than investigating alternative approaches. Some applications have been inappropriate for their intended task (reading instruction) and little research has been conducted by experts from the field of reading. Rather, the studies have been conducted by computer experts and software designers. He also comments that the Hawthorne or novelty effect can influence the outcomes.

Davidson (1994) further states that current evaluations of reading instruction software, including her own research, have not been entirely satisfactory, but that available alternatives in design methodologies all have their shortcomings. She argues that experimental studies conducted in laboratory settings take no account of the classroom context and integration of

the software. Furthermore, she notes that a purely qualitative approach would not answer ques ons of efficacy. Yet she acknowledges that her results may have been influenced by un rolled confounding variables although she attempted to address limitations outlined above.

Reinking and Bridwell-Bowles (1991) noted that the results of studies concerning reading software may be affected by novelty, conventional instruction, the amount and quality of reading outside school, and previous exposure to computer-based instruction. It seems that limitations and difficulties still exist in this field, as in other areas of education, and that these issues may never be resolved. In order to address these limitations, the design approach undertaken involved a combination of quantitative and qualitative data collection and analysis. Schofield and Anderson (1987, p272) have stated that: "Quantitative and qualitative research methods are not diametrically opposed strategies." The majority of research can be enhanced by the use of multiple data collection strategies (Delamont, 1992). Sometimes referred to as methodological triangulation (Cohen & Manion, 1989), this research strategy can minimise any bias that might occur due to the use of a specific data collection method. The use of multiple methods facilitates richer data collection and a more holistic view of the educational phenomena being studied. The use of qualitative techniques does not always necessitate detailed and time-consuming collection methods readily associated with ethnographic research. Rather, they can serve to illustrate the findings of the quantitative research (Delamont, 1992).

3.4.2 Design

Research instruments such as criterion and norm referenced tests were identified as being necessary for successful evaluations of computer assisted reading software. They were used in two studies in this research: a pilot study and a main empirical study. It was important to verify the validity of the assessments and procedures that were used. They were evaluated as part of the pilot study, which was a single group pretest - posttest design (Campbell & Stanley, 1963) and is reported in Chapter 4. This was an investigative study intended to both inform the software design (chapter 5) and assess the effectiveness of the framework for

evaluation (section 3.3). Therefore, it was not considered that a control group was necessary.

The main empirical study was conducted in a variety of classrooms, to maximise ecological validity by conducting the research in naturalistic environments. The study was a pretest/intervention/posttest design. The use of pretests and posttests enabled the effect of the software on learning outcomes to be measured, such as changes to sight vocabularies. As the number of participants was small (32), pairs of children were matched on reading age, class teacher and gender. This ensured that the children, who had varying abilities, were evenly distributed between the two groups and also, that the effects which might be attributed to the teaching styles of individuals were minimised. The effects of confounding variables, such as parental support at home, were minimised by randomly assigning each member of the pair to one of the two groups.

The pretests and posttests consisted of a number of assessments (informal testing, standardised reading tests and criterion referenced tests) as improvements in reading are difficult to measure (Pumfrey, 1979). In order to study in detail the effects of such a resource on children lacking in motivation and self-confidence regarding their reading ability, a number of case studies were run in the main empirical study. This enabled a richer set of data to be gathered. Further qualitative data, gathered through video-taping and informal observations, were used to support the findings of the quantitative research where appropriate. For example, motivation and attitudes to reading need to be evaluated by both observing user behaviour and eliciting comments and opinions from participants.

The study was a comparative study of two versions of the software, rather than comparing the use of a computer to supplement reading tuition with a more conventional classroom resource. Comparing the use of a computer with an equivalent paper and pencil task, or time with a human tutor introduces different problems (Blease, 1986; Venezky, 1983; Jones et al., 1996). For example, using the computer can enhance a learner's motivation, particularly so if the novelty effect occurs. Reinking and Bridwell-Bowles (1991) comment that "a major methodological limitation affecting existing research [on reading instruction software]"

(p321) is that the control group are not exposed to any equivalent, additional instruction. Thus the results obtained may be confounded by introducing a practice effect.

3.4.3 Observation and the running record

Observation of the use of the computer assisted reading software took place in the naturalistic setting of the classroom. It was intended that this would minimise any anxiety felt by the participants and ensure that ecological validity was high. The observer was not hidden, although it is acknowledged that the presence of a researcher may have affected the results. The children may have tried harder and concentrated on the task more than if they were working entirely independently (Ball, 1993). Each observation session was composed of two parts: systematic observation through use of a Running Record (Clay, 1993b), and informal observation recorded in field notes.

A Running Record provides a standard method of annotating a child's reading behaviour whilst performing a natural classroom task, that of reading aloud (Clay, 1993b). It is a protocol for recording errors made (miscues), self-corrections and other behaviours such as repeating a word or phrase. Systematic observation can be subject to criticism. For example, McIntyre and Macleod (1978) comment that the use of predetermined behaviour categories can be restrictive. The Running Record concerns an individual's reading behaviour and is more descriptive and flexible. By recording all miscues it is possible to infer what kinds of reading cues are being utilised and also, areas of weakness. For example, if a child reads a line of text inaccurately but it still makes sense contextually, then one can assume that they have used 'meaning'. For the purpose of recording audio data and as a requirement for taking a Running Record, the children were instructed to read aloud.

The wealth of data collected through observation contributed to all aspects of the evaluation framework. Taking a Running Record during observations ensured that a formal analysis of the utilisation of reading strategies could occur, providing an indepth picture of a child's reading behaviour (Goodman, 1973). This enabled both pedagogical and cognitive aspects of the framework to be evaluated. Informal ob[illegible]ons relating to the affective domain, such as concentration and evidence of distraction, were recorded. Technical aspects, such as

the ease of use of the software and the frequency of requests for help or signs of frustration, were also noted.

3.4.4 Norm referenced tests

Standardised reading tests were used as one of the instruments for evaluating learning outcomes. Any tests are subject to measurement errors. Reading test measurements can be affected by: illness, attitude to the test, motivation, recent exposure to test vocabulary and the success rate of guessing an unknown word (Pumfrey, 1985). Very little information on such aspects as a child's specific reading difficulties can be obtained from such tests and they may be considered as only an indicator of current performance. Clay (1993b) argues against using test scores alone:

> By comparison with the observation of learners at work, test scores are mere approximations or estimates, at times misrepresenting individual progress in learning, and at times presenting results stripped of the very information that is required for designing or evaluating sound instruction.
>
> (Clay, 1993b, p1)

Standardised reading tests are, however, commonly used to assess the reading abilities of a class for many reasons, mainly to quantify and record each child's progress. They are quick to administer and allow comparison of the group's scores with the national norms. Normally, a teacher would use standardised reading tests in conjunction both with other diagnostic tests and the rich data they gather from informally observing a child reading on a regular basis. This allows the reading abilities of each individual to be constantly monitored and the instruction to be tailored accordingly. A sentence based reading test, which facilitates the use of reading cues such as meaning to aid in the decoding process, was employed as the main instrument of measurement. A word based reading test was also used to assess sight vocabularies of the children.

3.4.5 Criterion referenced tests

These types of tests measure particular reading skills and competencies that have or have not been acquired or met. They are more suitable for eliciting information for diagnosing weaknesses and prescribing corrective measures, but give no information enabling comparisons with national reading standards to take place (Pumfrey, 1985). As with standardised tests, test scores from criterion referenced tests are also subject to error.

However, they do enable comparisons of skill levels to be made between participants. The tests were used to assess knowledge of commonly occurring words, and phonic skills such as blending and decoding consonant-vowel-consonant words through to multisyllabic words.

3.4.6 Interviews and questionnaires

Interviewing children

Interviewing was used to address the affective aspects of the evaluation framework. As well as being time consuming, interviewing children can be difficult for a number of reasons (Breakwell, 1995). They may:

- give an answer that they think the interviewer wants
- give a 'yes' answer, without considering the question
- answer 'don't know' for a number of reasons, such as not understanding the question
- be easily distracted
- discuss the interview with other participants who have not yet been interviewed

Nevertheless, it is considered to be the most effective data collection method for the young children who participated in the study. For example, a low reading age prohibits the use of questionnaires.

The interviews were structured using a variety of closed question styles, as well as open ended questions (appendix B). In addition, picture cards containing diagrammatic representations for the participants' feelings (happy and sad faces) were used to give young children a tangible medium to focus their attention upon (Murphy, Issroff, Scanlon, Hodgson & Whitelegg, 1995). This enabled them to rate their attitudes towards issues such as reading and computer use. The interviews took place in the children's classroom to minimise any stress.

The children were interviewed prior to using the software to:

- familiarise the child with the researcher
- assess their attitude to reading, preferences for books and reading partners and gauge the time spent on reading activities in the home environment

- assess their existing familiarity with computers and feelings towards their use (anxieties, gender stereotyping)

The children were interviewed after using the software to:

- assess whether or not attitudes towards reading and time spent reading at home had changed
- assess whether or not attitudes towards computer use have changed
- ascertain their feelings about the software in relation to reading with a teacher, a parent or guardian, or alone (what they liked and did not like about each method)

Questionnaires

The use of questionnaires that have been well designed (i.e. unambiguous questions, simple, well structured, short) provides a low cost method of data collection and can be thorough enough to support hypothesis testing (Fife-Schaw, 1995). They are most useful for eliciting information from a large number of people, or where it is more convenient than trying to arrange an interview. However, response rates can be low and the analysis of such data can be subject to misinterpretation.

Aspects of pedagogy and affective issues were investigated by asking all teaching staff to complete a simple questionnaire. The issues of confidence and reading preferences of individual children were assessed. Another important aspect is the teacher's feelings about the use of the software and classroom management. The issues of user independence, distraction caused by computer use, and contribution to classroom management must be considered as well as applicability to the intended pedagogical goals. This was supplemented by informal interviews. A questionnaire was also used for conducting a large-scale survey (chapter 5, section 5.4), designed to verify that the software design (chapter 5) in this thesis met the requirements of practitioners.

Case studies

Affective issues that relate to a child's feelings and experiences can be researched by case studies (Cohen & Manion, 1989). Walker (1993) highlights the disadvantages of this approach. For example, case studies can be partial accounts, and the reliability and validity of such data can be questionable. He suggests that case studies should not be conducted in isolation, but supported by other research methods. By examining observational notes,

considering informal comments from both teachers and parents, and including standard classroom documentation, such as assessment records, a richer set of data can be utilised. Evidence that a particular behaviour occurs can also be insightful and used to inform future research, or even be used to contradict a theory (Coolican, 1994). Furthermore, it can be argued that the generalisability of such qualitative methods can be increased by "choosing study sites on the basis of typicality and conducting multisite studies" (Schofield, 1993, p109).

3.4.7 Summary

It has been argued in section 3.4 that adopting quantitative and qualitative data collection methods is not only more informative but necessary in order to assess both cognitive and affective issues in relation to learning. A matched pairs approach in a pretest - posttest control group design was used to minimise sample bias. Each group used a different version of the software for the same amount of time during the intervention period. This ensured that the novelty and practice effects were controlled in this research and that it was a valid comparison of equivalent tasks.

The research instruments have been discussed, briefly highlighting the advantages and disadvantages of each. A Running Record was taken enabling inferences to be made about an individual's reading behaviour and diagnosis regarding areas of weakness to take place. Standardised tests can only be considered as approximate measures, but they are quick and easy to administer and enable scores to be compared to national norms. The use of criterion referenced tests enabled individuals to be assessed on word recognition and phonic skills. Affective issues were researched by interviewing all the children taking part in this study and were examined in depth by conducting case studies of specific individuals.

3.5 Conclusion

In this chapter, the literature on evaluating computer assisted learning has been reviewed. In common with Crook (1994), it has been argued that a holistic approach, in which both the software content and the context of its use are considered together, is imperative. Furthermore, evaluations of educational software can be improved by considering issues

such as classroom management, curriculum integration, pedagogy and learning (Squires & McDougall, 1994). Also, the evaluation should involve teachers and children, as well as the researcher. A framework has been developed to structure the evaluation of reading instruction software. The framework was classified into four domains (pedagogical, cognitive, affective and technical) in order to ensure that a holistic approach was undertaken.

Finally, the methodology and research instruments were discussed. In accordance with Delamont (1992), it is believed that incorporating quantitative and qualitative techniques will enrich the findings of this research. Qualitative research was considered to be necessary in order to analyse the affective aspects of the framework for evaluation effectively. It has been argued that a pretest - posttest design with matched pairs and a comparison group is the most appropriate structure for the evaluation of the software implemented in this research. A variety of research instruments that were used in both the quantitative and qualitative lines of inquiry have been discussed and justified.

The next chapter reports on a pilot study designed to identify the strengths and weaknesses of commercially available talking book software. Furthermore, the effectiveness of the research instruments used in the evaluation was investigated such that any required adjustments could be made prior to the main empirical study (chapter 7).

4. Investigating commercial talking book software

4.1 Introduction

This chapter describes a pilot study to evaluate existing software designed to support reading instruction in British primary schools. The study had two aims: to identify strengths and weaknesses of such software and to ascertain the most effective way to evaluate it. The background information together with a description of the design, procedure and results are described below. There follows a discussion of the results relating both to the software design and the effectiveness of the research instruments that were used. Finally, the conclusions of the study are summarised.

4.2 The pilot study

4.2.1 Software

The talking book software used was from the Sherston Talking Stories (Bonham & Bonham, 1993) range (see appendix A for complete list) and consisted of:

- 12 books of the Oxford Reading Tree scheme at level 2
- 6 books of the Oxford Reading Tree scheme at level 3
- 6 books written for Sherston, Naughty Stories Volume 1

The Oxford Reading Tree (Hunt, 1986) reading scheme adopts the top-down approach to instruction (Smith, 1971) and emphasises meaning over accuracy. The text in the Oxford Reading Tree books closely matches the illustrations such that unknown words can be decoded using semantic cues and the knowledge of language structure. The Naughty Stories were written specifically for the Talking Stories range (Bonham & Bonham, 1993) and the hard copy books were also provided. Only whole word pronunciations could be accessed and, when they were, the written form was simultaneously highlighted on the screen, which can support visual recognition. Any individual words that were selected were recorded in a log for teacher assessment purposes. This could be a useful feature if the child is to use the software unsupervised. Many of the talking books also included simple animation. Exercises were included at the end of each talking book and related to three or four key words, which were presented both in isolation and in context.

This talking book format was considered to be suitable for supporting the 'whole language approach' to teaching reading in which all required skills are taught holistically in the context of meaningful stories (Medwell, 1992). There was no coaching in phonic skills in these talking books because of the approach undertaken within the reading scheme.

4.2.2 The framework for evaluation

The framework used was that described in chapter 3, classifying the domains for evaluation into pedagogical, cognitive, affective and technical areas. The pedagogical issues concern the context in which the software is used, and whether or not it complements the specific aims and approach undertaken to reading instruction by an individual teacher. This was assessed through interviewing the teaching staff involved and by observation of normal classroom practice. The cognitive issues relate to the possible learning outcomes such as an increase in utilisation of specific reading cues or a sustained improvement in reading ability. This was assessed through standardised tests and by observation of the use of the software. The affective factors include harnessing the motivational elements of computer use, and both improving self-confidence and enjoyment of reading. This was assessed by interviewing teaching staff and asking each child about their attitude to reading both before and after using the software. The technical elements are those which ensure that the technology is being employed to its maximum advantage, such as incorporating sound and colour, as well as assessing usability issues. This was assessed by evaluating the content of the software and by observing its use.

4.2.3 Participants

The school that participated in the study had a catchment area including children from a wide range of socio-economic and ethnic backgrounds. It was a middle school with children from year 4 (8 - 9 year olds) to year 7 (11 - 12 year olds) of the National Curriculum. A group of 15 children were identified as having reading abilities considerably below that expected for their age range by their class teachers. Nine of these were selected to participate in the study from the three classes in year 4. These children were the poorest readers of the group, and were identified as being at a suitable level to benefit from the stage of the reading scheme that the software covered. The children selected, who have been given pseudonyms, were:

- three children with English as a Second Language (ESL), one girl (Suraiya) and two boys (Karim, Kunal)
- one boy (Leonard) with a visual problem (cross referenced eyes)
- one boy (Tony) with Emotional and Behavioural Difficulties (EBD)
- two boys (James, Trevor) with behavioural and learning difficulties
- one boy (Aaron) and one girl (Laura) with learning difficulties

Their ages ranged from 8 years 9 months to 9 years 2 months at the beginning of the study. The mean age was 8 years 11 months. The pilot study was conducted during the second half of the summer term.

4.2.4 Design and procedure

The study was a non-comparative exploration of the use of talking book software with children whose reading ages were considerably below their chronological age. Although the intervention took place in a special room, outside the naturalistic classroom setting, the children were already used to leaving the classroom for additional tuition with a welfare assistant, generally in the school library. The study included qualitative research, which looked at issues of confidence and motivation through observations and interviews. It also incorporated quantitative research, in which improvements in reading ages, phonic knowledge and the use of reading strategies were evaluated using both norm referenced and criterion referenced tests.

Initially, all children were pretested. They were interviewed and asked a number of pre-defined questions (appendix B.1) on their attitudes to:

- books and reading
- computer use at home and at school
- reading with the computer as compared to reading either alone, with a friend, with a carer, or to a teacher

The children were then asked to undertake a battery of criterion referenced tests that were based upon the type generally contained in a local education authority reading assessment pack (appendix C). These tests were used to measure sight vocabulary of common words, alphabet knowledge and phonic ability. The phonic assessments ranged from the ability to decode consonant-vowel-consonant words through to multi-syllabic words. Lastly, two

norm referenced tests were carried out: the Salford Sentence Reading Test (Bookbinder, 1976), and the Burt Word Reading Test (1974 revision) (Scottish Council for Research in Education, 1976).

The children were already using a variety of books in the classroom. These ranged from reading schemes designed for all beginning readers (e.g. The Oxford Reading Tree scheme, Hunt, 1986) to those designed for children with problems. Six of the children were receiving additional support from a welfare assistant: James, Karim, Kunal, Tony, Trevor and Laura. Two of the children were involved in the Somerset Talking Computer Project (Bolton, 1995) during the same period (Laura, Aaron). For this project, the children spent 20 minutes a day typing phonically graded sentences into a talking word processor with adult support. It is likely that the additional practice provided by both the welfare support and participation in the Somerset Talking Computer Project contributed to any gains made in learning outcomes. It was not possible to identify nine children with reading problems who were not receiving additional support.

It was assumed that most children in year 4 (ages 8 to 9) would be independent readers and, indeed, the majority of the children in each class had reached an acceptable standard. The staff did not teach reading skills to the class as a whole, but encouraged the children to read together or independently. This occurred either when they had finished their given classroom tasks or at designated 'reading' times. Welfare assistants and other adult helpers regularly visited each classroom to listen to the children read. The children were expected to read every word, and those words they could not decode were noted on their individual reading records.

The children then spent 20 minutes a day, four days a week, for a period of four weeks, using the software described above. All the children were familiar with the books in the Oxford Reading Tree range, although not in the talking book format, and some of them had read the texts many times before. They were given a five minute introduction on how to use the software. They were instructed to attempt to read the words on their own first and shown how to get feedback if they required it. They were also asked not to activate the animation on

each page until they read the text. The children worked alone at the computer, in the presence of an observer.

They were asked to read two different books during each session, and were asked to repeat the books if they had not read them successfully. The next session either started with the last book read or the next book at that level, depending on the confidence level and accuracy level that each child was observed to have. This intervention was an attempt to emulate the actions of a teacher. After completing each book, in order to assess the child's comprehension, each child was asked:

- to explain the meaning of three key words from the text
- to answer three questions on the story content
- to summarise the story in their own words

Each session was recorded on video or audio tape. An observation sheet was used to note any problems encountered and points salient to issues of motivation and confidence. A Running Record (Clay, 1993a) (chapter 3, section 3.4.3) was taken for each book read, in which all errors and self corrections were documented together with the child's use of computer feedback for both cross-checking and decoding.

After the period of intervention, the children repeated all the pretests and the teachers were given a questionnaire (appendix D) which collected information via open-ended questions on:

- the social background and characteristics of the child
- an assessment of the child's reading ability and problems
- whether the child had been given additional instruction that may have affected the outcome
- any obvious changes in the child's attitude to and confidence in reading

4.3 Results

The results were derived from pretest and posttest data, the observation records kept for each child and the questionnaires completed by the teaching staff. This data was analysed in the context of the framework for evaluation (chapter 3, section 3.3) and the results are presented under the headings of pedagogical, cognitive, affective and technical issues.

4.3.1 Pedagogical issues

The following results relate to the pedagogical issues outlined in the framework for evaluation. These issues are concerned with the suitability of the software for the intended context and the soundness of the educational theories upon which the software is based.

The software has already been described (section 4.2.1). It was considered to be suitable for the context in which it was used because the text books upon which the software was based were already being used in the classrooms involved. The software was clearly related to the existing teaching materials.

Four interaction styles were observed when the children were reading talking books in the Oxford Reading Tree range.

The first style was that of obtaining feedback instantly, without any obvious attempt to decode the unknown word, for which there could be many reasons. The child adopting this behaviour could be confident in their own ability to read, but lazy. Alternatively, they could lack confidence in their own ability and be fearful of making mistakes. They could be either totally unaware of the utilisation of reading strategies or aware of them, but lack confidence about their use. Alternatively, the child could perceive instant feedback as the most efficient way of learning and it could be their preferred learning style.

The second style identified was using the software more effectively, attempting to decode an unknown word by employing a variety of strategies and requesting computer feedback if unsuccessful. The child exhibiting this behaviour could be confident in their use of strategies and in their ability to read.

The third style was reading for meaning, when word for word accuracy was not achieved. However, the substituted word still made sense and the child was unaware of their error, seeking no help from the computer. This type of behaviour could indicate confidence in decoding accuracy, and/or a desire to find out what happens next in the story. Children exhibiting this behaviour could be holistic in their approach to learning to read. Conversely, this type of behaviour could indicate that the learner is demotivated, uninterested in the story,

and not concerned about their accuracy. Some children, for a variety of reasons such as a desire to maintain independence, are reluctant to ask for help, even if they need it. Fewer errors than had actually been made were recorde[illegible]e software generated logs used for the purpose of teacher assessment. Despite the errors, the children were able to successfully answer comprehension questions relating to the story.

The final style was reading 'with' the computer by selecting each word as it was attempted. If the child was confident of being able to read a specific word it was spoken clearly before the pronunciation was requested, but otherwise it was selected simultaneously.

Each child who participated in the study exhibited more than one style of interaction throughout the observations. In addition, they were all observed cross-checking a guess by selecting whole word pronunciations. Seven of the 9 participants clearly preferred one specific style but were sometimes observed using one or more other styles of interaction. The remaining two children used a combination of styles in equal proportions. Table 4.1 summarises how frequently the interaction styles were observed, with the shaded cells indicating the primary style (exhibited in 50% of the observations). Two children, Tony and James, read 'with' the computer, and were observed to use the other three styles but the data recorded was not accurate enough to include in the analysis. For example, it was difficult to determine if feedback for a word was given before the child pronounced it (instant feedback) or if the order was reversed and confirmation of a guess was being sought.

Table 4.1: Percentage of interaction style examples observed

Participant	Reading for meaning	Effective use	Instant Access
Laura	57%	16%	24%
Karim	35%	25%	40%
Suraiya	56%	27%	17%
Leonard	68%	15%	17%
Kunal	38%	11%	51%
Trevor	39%	20%	41%
Andrew	74%	4%	22%
Tony	N/A	N/A	N/A
James	N/A	N/A	N/A

Four children appeared to prefer reading for meaning, one child was inclined to opt for instant feedback, and the remaining two children were observed using a variety of interaction styles. Two of the children read with the computer and data on interaction styles was not collected. No child was observed using one interaction style to the exclusion of all others.

The illustrations in the original hard copy books spanned both the left and right hand pages of the open texts. Yet the illustrations had been replicated on the computer screen as one complete page as depicted in Figure 4.1. All of the children were observed to miss text on pages where there were two sentences on the computer screen, one at the top of the page and the other at the bottom. In these cases, they omitted the top line, whichever side of the screen it appeared upon. The children were only alerted to this mistake if they used the 'Ear' button (requesting that the entire sentence be spoken) as a cross-checking facility. On pages containing two sentences, selection of the 'Ear' button activates speech for one sentence, and a second 'Ear' button becomes visible on the screen for the second sentence.

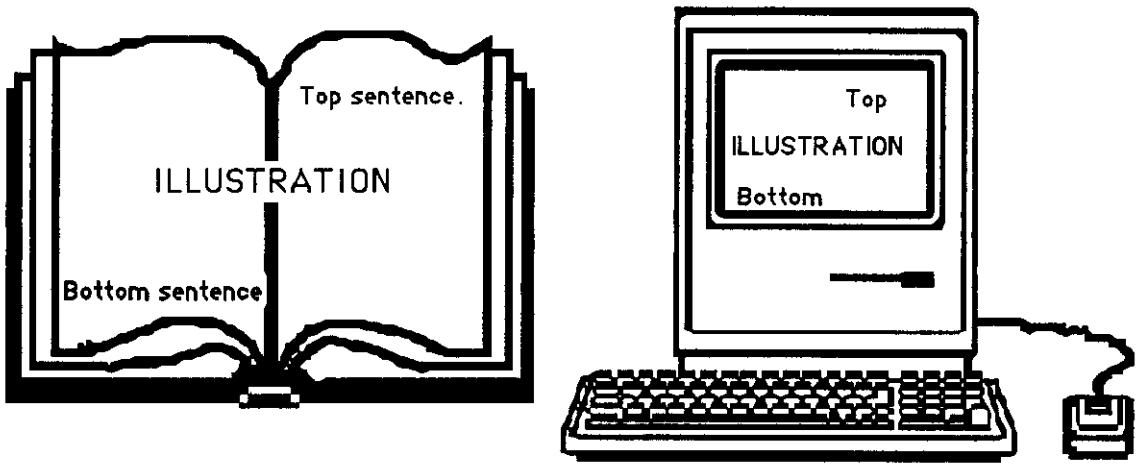


Figure 4.1: Two pages in book become one page on screen

This design approach also created another problem, as illustrated in Figure 4.1 above. Children are taught to read text page by page, top to bottom and left to right. Thus, in the hard copy book a child will see two pages and read the left hand page first and then the right hand page. When it is replicated on the computer screen, the child will see one page only and may read the top sentence (originally the second page) and then the bottom sentence

(originally the first page). However, in the specific occurrences of this that were observed, the meaning and logical flow were not disrupted for the children. In these instances, the order in which the text was read did not seem to matter.

Those children who made use of the 'Ear' button facility on these pages, when the left-hand 'bottom' sentence was spoken first and then the right-hand 'top' sentence, did find it slightly confusing. This is a design feature that should be investigated further and addressed in future implementations.

Most of the children used the computer feedback for cross-checking purposes, again in different ways. Some children requested feedback for individual words to verify accuracy and some children opted to have the entire sentence spoken to them after their own attempt.

The interactivity of the software was minimal and there was no support for related skills such as writing, prediction or comprehension. The children were able to control the pace at which they read all the books, but progression from book to book was controlled by the observer. The level of difficulty and vocabulary used in the Oxford Reading Tree Talking Stories was suitable for the children involved in the pilot study. The Naughty Stories were written for Sherston Software and about 50% of the vocabulary was unknown to the children who read them. The books were designed to be read to children aged 5 to 7, rather than as material to be read independently. The children found the level of difficulty frustrating and did not enjoy these books as much. Those children in the study group who successfully completed the Oxford Reading Tree Stage 3 Talking Stories went on to read the naughty stories but exhibited a different kind of behaviour to those noted above. They became more reliant upon instant feedback from the computer, and also opted to have each sentence spoken to them before they attempted to read it for themselves.

4.3.2 Cognitive issues

The cognitive skills to be assessed within the evaluation framework are:

- gains in reading ages as measured by standardised reading tests
- increases in common word recognition and phonic knowledge as measured by criterion referenced tests
- improvements in the employment of reading strategies

Table 4.2 summarises the results of both the Salford Sentence Reading Test and the Burt Word Reading Test, and the assessment of common words contained in the criterion referenced tests (appendix C). The number of children successfully completing each test is indicated, as James did not reach the standard required to record any results in the Salford Sentence Reading Test. The mean gain is given in months for the two standardised tests and in number of words for the common word knowledge. The study group was composed of the children with the greatest problems in learning to read and so the range of results is also given to demonstrate their diversity.

Table 4.2: Reading ages and Common 100 Words Test (mean gain, standard deviation and range)

Test	n	mean gain	SD	range
Salford Sentence Reading Test (months)	8	1.6	2.6	-4 to 4
Burt Word Reading Test (months)	9	5.7	3.5	0 to 12
Common 100 Words Test	9	8.5	9.0	-2 to 24

Without any intervention, after an elapsed period of one month, we might expect a mean gain of one month in a standardised reading test score. Of course, it is important to recognise that reading does not develop smoothly in this way: children may make no apparent progress for a while and then 'leap' forward two or three months. Nevertheless, the notion of a month's gain over a chronological month is a useful, if rough, baseline against which to measure any improvement when an intervention is introduced. The overall results were therefore very positive. Three of the nine children showed improvements of more than one month in both standardised reading tests as well as improvements in their knowledge of common words. The individual results for the children in the study are given in appendix E.

The criterion referenced tests included assessments of phonic knowledge. Although the computerised texts were based on the top-down models of reading processes (e.g. Smith, 1971), it is possible that these skills may develop without direct instruction through frequent exposure to words (e.g. Rayner & Pollatsek, 1989). Two children did appear to make more use of phonic synthesis during the posttest than they did during the pretest. The strategies

employed for decoding unknown words did not change overall during the intervention stage. Four children used the full range of available strategies and cues, whereas one child clearly read for meaning and made little use of phonic strategies.

4.3.3 Affective issues

The following results relate to the affective issues considered within the framework for evaluation. The specific interests are motivation and confidence.

The children were very motivated by the use of this technology. They were very keen to participate in the study and talked enthusiastically to their teachers about what they had done. Several of the children repeatedly asked whether they would be first in the next session or would rush to take their turn if they realised that someone from their class would be next. The children enjoyed the special and individual attention they felt that they were receiving. Although some children did appear to fidget occasionally, they did generally keep their eyes on the screen. Recording time on and off task was not considered to be a valid measure. It is difficult to draw conclusions as to whether or not looking away from the screen implies that the child is indeed off task (Issroff et al., 1994).

Two children, Laura and Aaron, were involved in the Somerset Talking Computer project (Bolton, 1995) at the same time that they were participating in this research. They commented at posttest interview that the talking book software was better than the Somerset Talking Computer project because there were more pictures and it was more fun.

The children found the simple animations both humorous and rewarding, although Suraiya chose not to activate them. Tony, the child with emotional and behavioural difficulties, paid more attention to the animations. He discovered early in the intervention period that the animations could be activated by clicking the mouse on various 'hot spots' on the screen. Tony was extremely demotivated by reading and may not have spent so much time on the activity without this incentive. He also seemed to be disappointed that the talking books were the same as those already used in the classroom. He refused to read more than one book during a session or to attempt the key word activities at the end of the book.

The children made effective use of the computer feedback for cross-checking purposes, either for an individual word or for every sentence they had read. James, who had learning difficulties, was extremely motivated by this facility, enthusiastically calling out "Yes!" when using the feedback to confirm that his attempt at a word was correct.

Four of the nine subjects indicated a preference for using the computer for practice, rather than reading alone, reading to a friend or relative, or to the teacher. All the subjects expressed a desire to continue using the computer and the talking book software. Eight of the children did not feel that there were any disadvantages or problems using the computer to practise reading, although Kunal commented that "it shouted". There were several comments relating to the positive aspects of this teaching resource. They felt that it helped them to learn new words and it was enjoyable because of the use of sound and animation. They also felt that the facility to get help when in difficulty was very useful, which clearly contributed to their sense of independence (Chapter 2, Section 2.4.4):

> Makes it so you don't have to ask Miss ... you can just click on the word that is the problem.
>
> It means that you don't have to waste the teacher's time.

Four of the nine subjects felt that their reading had improved and stated that they enjoyed reading more, on completion of the intervention period. Three of the subjects did not feel that their reading ability or enjoyment of reading had changed. When questioned as to why he felt that using the computer had not had any effect on his reading ability, Kunal said "it tells the words easy". He commented further that the facility to request a word pronunciation meant that he did not have to think about or attempt to decode unknown words. Kunal perceived that the level of challenge in the software was low, which may have contributed to his lack of motivation. One child felt that his reading ability had improved but that the enjoyment of reading had remained unchanged. One child did not feel that his reading ability had improved, but he enjoyed reading more "on the computer, but not with people".

Teacher G made the following comments regarding the children's attitude to reading and confidence, stating that they had all improved:

> He has been keen to take part in the study which luckily happened at the end of the summer term, when he was beginning to become frustrated with his lack of ability. It has renewed his interest in books.
>
> He has been very keen to work on the computer, he enjoys the one-to-one approach. He now tries to read information without relying completely on someone to support him.
>
> He is more confident of trying to read, he used to seem frightened of making a fool of himself and would rather stay silent. Now he will have a go.

Teacher D felt unable to state categorically that the intervention had directly caused an improvement in the children's attitude to reading and confidence, but suspected that it had been a contributing factor. He did state that the children had enjoyed it.

Teacher H stated that one child's attitude to reading had not changed, but his confidence in reading had improved. The other two children improved both in attitude to reading and confidence. Teacher H made the following comment:

> He enjoys reading with others. Reads to others and is 'excited' about doing so which is very unusual.

4.3.4 Technical issues

Technical issues relate to harnessing the potential of the technology available such as the use of sound, text and graphics, and the acceptability of the user interface. The resources available should be cost-effective and provide some advantage when compared to alternative classroom resources.

The feedback was clear, although some children had difficulty hearing one or two words correctly (for example 'fight', 'stunt-man', 'full'). It was easy for the children to obtain feedback as required, either as a single word or a sentence. The feedback was given as whole word only. It would be possible for a child to listen to the story in its entirety, play a passive role and not interact with the software at all, although this was not observed. No instructions were given by the software, but with a five minute introduction the children experienced little difficulty learning how to use the software and were able to use the software independently. However, intervention was required to ensure that the children progressed at a suitable pace through the books. No software interaction giving

encouragement or further guidance took place. The children used the software with minimal supervision.

4.4 Discussion

4.4.1 Evaluation of talking book software

As there was no support for phonic coaching, the current feedback mechanism would not be suitable for reading schemes incorporating this teaching approach. For example, recent schemes have been designed to support phonological training at the level of onset (the starting sound of a syllable) and rime (the remaining sound of a syllable). There was no specific coaching in any of the strategies, other than whole word recognition. Those children who were not fully utilising all reading strategies prior to the intervention did not show any signs of improvement at the posttest stage.

The presence of an observer could have affected the interaction styles that were seen. It is possible that some of the children made more effort to decode the words themselves and use the software effectively on more occasions, simply because they knew they were being watched. However, the interaction styles of the children using the software has demonstrated that two specific issues relating to the user interface and overall design of the software need to be considered.

The first concerns accuracy, although this will only be an issue for those teachers who demand it. This issue will not cause problems for those teachers who are committed to reading for meaning where accuracy is not essential (Smith, 1985). This problem can only really be addressed by the addition of speech recognition capabilities, which would record the accuracy of the user. This facility would also address the problem identified in this evaluation whereby children were seen omitting complete lines of text.

The second issue relates to the use of immediate whole word pronunciations in preference to attempting to decode words by any other means. This is a complex issue because, although children may become reliant on the computer feedback, they may still be making effective use of the information provided and improving their word recognition skills. The child who most strongly exhibited this behaviour was seen to be using other strategies on occasions

and, more importantly, he made gains in both standardised reading tests and the common word test. This child did not show an improvement in the utilisation of strategies and the use of error detection and self-correction techniques were not observed.

One of the questions to be answered is how the use of talking book software compares with reading to an adult or experienced peer. According to Clay (1993a) an effective tutor should give the child a way to detect errors and encourage them to attempt to correct them. If a child struggles then rather than being given the whole word, they should be directed to clues to encourage independent self-correction. The software in this evaluation does not emulate an experienced tutor in this way. However, in relation to silent reading and getting help from the teacher as required, the children involved in the study seemed to prefer the computer as it gave them a sense of independence. Using this software could be preferable for children who do not feel at ease reading to adults or those who are fearful of making mistakes in front of their peers (Davidson, 1993). Thus, it could be a suitable tool for those children lacking confidence in their reading ability.

Another issue arising from the study relates to those teachers who regard reading accuracy as a measurement for progression to the next book. Relying upon the reporting system built into the software may give a misleading picture of a child's success. The assessment summary available to teachers indicates which individual words have been selected in order to receive speech feedback. However, in accordance with Miller et al. (1994), a child may select feedback to cross-check a guess and records of words selected may not be a reliable indication of unknown words. Again, this problem could only be rectified by the use of speech recognition technologies, which could monitor and record those words that had not been read correctly. Currently a teacher needs to listen to a child reading the text book independently of the software in order to make an accurate assessment of their progress. As the software stands, a teacher would still need to be responsible for controlling progression through a number of different books as indeed they do already with traditional books.

Many strengths in the current software design have been identified. The digitised feedback is generally clear and serves both as a facility for cross-checking information or as an

alternative decoding strategy when all else fails. This can give a sense of control as well as independence. Using colour, illustrations and simple animations can motivate the children. The use of the software can provide the opportunity for regular practice, at little cost in terms of teacher support. Results of the pilot study have highlighted improvements in reading ages above that which might be expected during the time period. There was also evidence of improvement in both attitude and confidence in relation to reading. However, it is only possible to state that the use of the software has been a contributing factor to these improvements, not the sole cause.

Several weaknesses that should be addressed have also been identified. As the feedback given is whole word only, there is no provision for phonic training although this requirement is dependent on both the teaching method and the particular reading scheme that have been adopted. There is no other support for providing coaching in the employment of other reading strategies, such as the use of meaning and illustrations, and no encouragement is given. It is possible that a child may not read the text accurately, but again the importance of this factor depends upon the teaching method adopted. Finally, it is too easy for children to receive feedback and to adopt a passive role if they choose to do so.

4.4.2 Review of research instruments and framework for evaluation

The framework used in this pilot study ensured that the evaluation design provided a comprehensive amount of data, both quantitative and qualitative. One issue arising from this study was how talking book software should be integrated within the classroom, which may depend on individual approaches to reading instruction. In this study, it was the researcher who decided how the software should be integrated within the classroom based upon information gathered from the teacher, classroom observations and findings from the literature. There is a need for a more methodical approach to establish how the software best fits within the context of the classroom in which it is being evaluated. This could be addressed by conducting an indepth, but informal, interview with teaching staff at the pretest stage. In addition, the framework could be extended to include areas such as parental involvement but this has not been done in this research.

Although measuring motivation can be difficult, indicators for motivation appeared to be easily identifiable in this study. It was however difficult to assess any changes in confidence in reading from the data that was collected. This weakness could be overcome by incorporating qualitative analysis of video data, identifying instances of confident and motivated behaviour.

Any gains made in the standardised reading tests may have been affected by the limited range of new vocabulary introduced in the software. The quantitative measurements would benefit from the addition of a word recognition test relating to the specific vocabulary introduced in the software, such as the key words.

4.5 Conclusion

The children participating in the study greatly enjoyed the experience. They found the use of talking book software fun and perceived it to be a useful resource, which offered the facility to practise reading in an independent fashion. All of the children received some benefit from using the software, whether it be an improvement in standardised test scores or a greater enjoyment and confidence in reading. However, it is not possible to attribute cognitive or affective improvements to the use of the software alone. Rather, the evidence suggests that the use of talking book software to provide supplementary practice can contribute to such outcomes.

Several issues relating to the current design of talking book software arose from the study. A number of strengths that should be incorporated in future designs were highlighted. Weaknesses noted in the current design include the lack of phonic instruction, reinforcement exercises and coaching in the use of reading strategies. Four styles of interaction were identified, one of which indicated a reliance on immediate whole word pronunciations. This could be considered to be an undesirable style of interaction, enabling the child to take a passive role and the route of least effort.

The framework for evaluation (chapter 3, section 3.3) provided a mechanism for the collection of a comprehensive set of data. Basing the evaluation upon the framework enabled it to be holistic thus bringing together issues of pedagogy, content and technology.

However, improvements have been suggested in relation to the design of research instruments used in software evaluations guided by the framework. These are: greater teacher involvement in the approach to classroom integration of talking book software, incorporating qualitative analysis of video data and including a key word recognition test.

A design for talking book software that incorporates the strengths and addresses the weaknesses identified in this pilot study is presented next, in chapter 5. Modifications to the data collection methods used in the evaluation were implemented in the main empirical study in chapter 7.

5. Enhanced talking book software design

5.1 Introduction

The software design of enhanced talking books reported in this chapter was informed both by the findings of a literature survey (chapter 2) and the results of a pilot study (chapter 4). In addition, a survey of practitioners' use of such software was conducted to verify that the proposed design and innovative features were perceived by them as being potentially beneficial.

The aims were to ensure that the design:

- is theoretically underpinned
- complements current classroom practice and incorporates any features identified as being of potential benefit but not yet available in current commercial software
- complements the structure and philosophy underlying the reading scheme upon which the software is based

Firstly, issues informing the system design are briefly reviewed. Theoretical influences are summarised together with the findings of a review of existing commercial software and a pilot study evaluating one range of talking books. This study identified the strengths and weaknesses of such software, together with styles of interaction and affective considerations in relation to tutoring strategies.

Secondly, the design specification is presented. The objectives are specified, and an overview of the design in which the system is subdivided into five modules, each concerned with a particular functional role, is discussed. This is followed by a detailed description of each of the modules.

Thirdly, in order to ensure that the design complements current classroom practice, a survey was conducted of existing users of commercially available talking book software. The design of this survey is described together with the results that were obtained. The conclusions drawn from this and their implications for supporting the system design are summarised.

The chapter concludes with a discussion of how the aims and objectives have been met within the design.

5.2 Issues relating to design

The following section summarises the findings discussed in earlier chapters that informed the design.

5.2.1 Theoretical implications

The conclusions from the review of the current theories of reading were that an integrated approach to instruction is the most flexible and effective teaching method (e.g. Wray & Medwell, 1990). In the integrated approach, both bottom-up and top-down skills are taught together in the context of meaningful and interesting stories. Thus, talking book software is an appropriate format with which to provide computer-based supplementary tuition.

It is accepted that phonics tuition should be incorporated in any reading programme (Adams, 1990). Goswami and Bryant (1990) have shown that phonic training at the level of onset and rime can have positive effects on early readers. Previous research on this form of feedback has also highlighted its potential (e.g. Wise et al., 1989). Clearly feedback at the level of onset and rime should be provided in addition to whole word feedback.

It is not always appropriate to tell a child an unknown word immediately. A teacher may draw a child's attention to the onset of a word, or encourage them to look more carefully at it and then try starting again from the beginning of the sentence. This is substantiated by educationalists, such as Clay (1993a), who advocate that children should be encouraged to use a variety of strategies and cues to decode unknown words independently. A means of replicating these tutoring strategies should be considered in the design.

It may be the case that some of the features, such as word pronunciations, provide a form of scaffolding for children learning to read. A variety of features will be provided and each one may be more appropriate for readers at differing stages of learning to read. Furthermore, children will be able to request as much or as little support as they require.

5.2.2 Strengths to be incorporated

From a review of commercially available software and the findings of the pilot study, the following strengths in the current software design have been identified:

- the basic format of replicating the page of a hard copy book on the screen is intuitive and enables a child to be both independent and in control of their learning
- the facility to obtain word pronunciations serves both as a facility for cross-checking information and as an alternative decoding strategy when all else fails
- when a word on the screen is selected for speech feedback, highlighting the word as it is spoken reinforces the relationship between the sound of a word and its written form
- the facility to hear a complete sentence spoken with expression provides an appropriate model and enables a child to self-check their own attempts at reading a sentence
- the use of colour, illustrations, and simple animations make the software more fun and help to motivate the children

The pilot study results indicated that the talking book format has potential as a supplementary resource for providing additional reading practice. Furthermore, the results indicated that learning outcomes could be improved, although it is not possible to attribute this to the use of the software alone. There was also some evidence to suggest that both children's attitudes to and self-confidence in reading improved.

5.2.3 Weaknesses to be addressed

The following weaknesses of existing talking book software design were identified in the pilot study:

- a child may become reliant on using the computer to decode unknown words and not develop the use of alternative strategies
- a child may have the entire book read to them without attempting to read it themselves
- interactivity can be minimal, with no supporting activities to reinforce the learning
- the feedback tends to be whole word only, particularly in commercially available software
- there is no mechanism for correcting and detecting errors if help is not requested
- talking book software does not always provide coaching in a variety of reading strategies

5.2.4 Affective considerations and styles of interaction

Adults listening to a child reading aloud can provide instruction that is adapted to suit an individual's needs. They can also provide affective support, promoting self-motivation and self-interest in reading by making reading a pleasurable experience and giving guidance, praise and encouragement. Lepper and Chabay (1988) state that there is substantial variation in the amount of tutoring support required by individual learners, some wanting complete independence and some needing constant direction. They strongly believe that computer-based tutors should combine cognitive and affective issues of learning in order to fully replicate an empathetic and supportive human tutor. Lepper, Woolverton, Mumme and Gurtner (1993) take this further and highlight the importance of the affective goals of a human tutor when working with remedial students whose confidence levels are already at a low ebb. They stress that to build confidence levels, success should be maximised and failure minimised. Furthermore, direct negative feedback should be avoided and should be replaced with hints or other indirect feedback.

It is difficult for computer-based tutors to diagnose or even detect the affective or motivational state when a human tutor draws upon so much intangible information such as body language (Lepper et al., 1993). One possible solution is to elicit the information from teachers and from the children at periodic intervals. Currently, without the facility for a computer tutor to hear and see its learner, the information it can obtain is extremely limited.

One outcome of the pilot study reported in chapter 4, section 4.3.1, was that four styles of interaction were observed:

- obtaining feedback instantly, without any obvious attempt to decode the unknown word
- making several attempts to decode an unknown word independently and using the computer feedback as a last resort
- requesting no or few words for feedback, reading for meaning and unaware of errors made
- reading with the computer, selecting every individual word and sometimes attempting to guess it, prior to hearing it

Automatic detection of these four styles of interaction may be difficult with the limited information available, though this could be rectified in the future with the development of

speech recognition technologies. The differences between the styles are subtle, and monitoring the frequency of requests or timing interactions will not contribute greatly towards identifying how a user is interacting with the software. For example, if a child using the talking books made no requests for feedback then the software could interpret this as the interaction style that concerns reading for meaning. However, it could also be the case that the book had been read successfully, by a child adopting the first, second or third interaction style. The discussion of these styles in chapter 4, section 4.3.1, highlights a number of possible reasons as to why each behaviour type may be adopted by a user. Any intervention, such as varying the type of feedback given according to the interaction style detected, could improve the effectiveness of the tuition. Conversely, it could make the process of learning to read even more frustrating. Of the four interaction styles, the only one that is easily identifiable is that of the child who reads with the computer. The information that can be obtained dynamically will not be sufficient to diagnose accurately the feedback requirements for an individual who exhibits any of the other three styles of interaction. Thus, there will be a need for a teacher to provide this information and specify the level of support that is required.

In conclusion, talking book format is theoretically grounded, may offer scaffolding to early readers, can complement classroom practice, and has many strengths that should be incorporated into future designs. Also, a number of weaknesses, or features that could improve the design as it stands have been identified, including strategies for addressing affective issues. Interaction styles have been reviewed in relation to the limited information that a computer tutor can obtain, and this has led to the suggestion that additional information should be provided by a teacher.

5.3 System design

An innovative design for the software is now presented in which the issues raised in section 5.2 have been addressed. The system is referred to as the Computer Assisted Reading Instruction Software (CARIS). The objectives are specified, followed by an overview of the complete system and a detailed description of each of the main modules.

5.3.1 Objectives of design

In order to meet the aims of the design, the important objectives were to:

- incorporate the identified strengths of existing talking book software
- address the identified weaknesses of existing talking book software
- be adaptive to individual users' needs
- emulate a supportive tutor and incorporate affective as well as cognitive teaching strategies

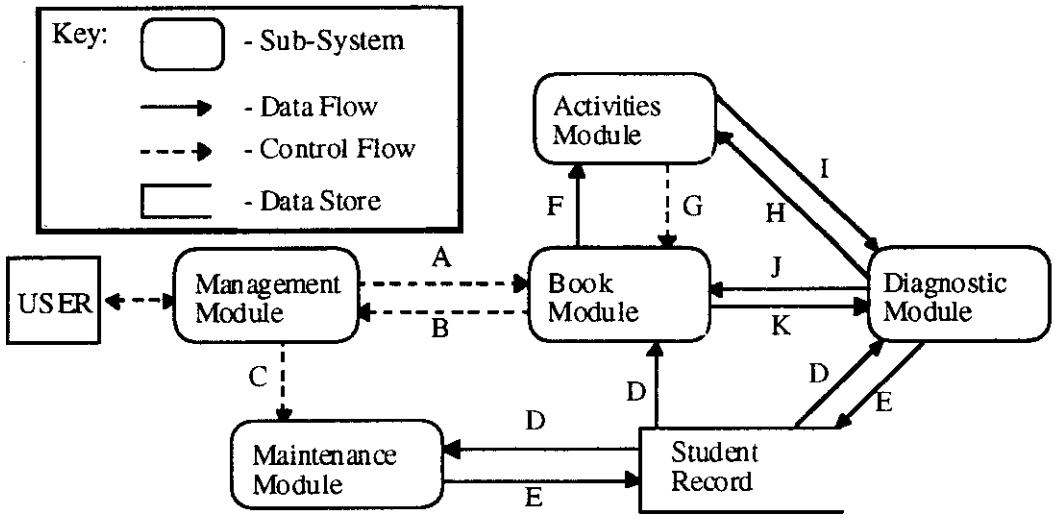
5.3.2 Overview

CARIS is based on the current talking book software format for individual books. In addition, CARIS includes a front end management system to control access to the books and to ensure that each user can log in individually. The talking books are based on the Ginn All Aboard reading scheme (Fawcett, 1994) which supports an integrated approach to teaching reading. This particular scheme is divided into three strands at each level of difficulty, designed to support:

- the acquisition of sight vocabulary of key words
- the development of phonic awareness
- the development of reference skills (non-fiction)

CARIS comprises twelve books from the sight vocabulary strand and eight books from the phonic awareness strand, all of which are taken from Stage 3 of the reading scheme (appendix F). This stage of the reading scheme is intended for those children in reception and year 1 of the National Curriculum who are progressing normally.

Figure 5.1 gives a high level overview of the data flow through the CARIS modules. Five different types of activities provide coaching in the use of specific cues in reading: initial letter or onset, meaning and syntax, key word recognition and illustration. When an activity is invoked, the illustration and text from the current page are used within the task in order to maximise the variety offered. For example, in the key word activity the sentence from the current page could be presented and the child asked to click on one of the words it contains. This would vary according to where the child was in the story.



A: Read new book
B: Exit book
C: Maintain student data
D: Get student record
E: Update student record
F: Current book page for activity
H: Get weak strategy to select activity
I: Increment/decrement strategy models based on activity results
J: Get weak strategy for hint
G: Return to book
K: Increment/decrement strategy models based on hint given and/or word selected

Figure 5.1: High level data flow diagram for CARIS

A diagnostic module models both the use of these specific reading cues and the confidence level of a user. This enables both the style of feedback and type of activity selected to provide coaching in areas of weakness. An additional activity encourages the development of prediction skills. A maintenance module enables the style of interaction to be configured for each individual user, allowing CARIS to be tailored to suit individuals' needs.

A detailed description of each of the modules represented in the diagram is now presented.

5.3.3 Management module

The initial screen contains the title of the software and all credits and acknowledgements. The remaining two screens enable a child to select their own name from a list box, and then to select a book title from a list box. The available books for each child can be specified via the Maintenance Module (Figure 5.2) and altered at any time. This ensures that the order in which the books are read and the number of titles that are available at any one time can be controlled by the teacher, supporting current practice in the classroom. Access to the Management Module is from a pop-down menu box on the first selection screen, Select User Name.

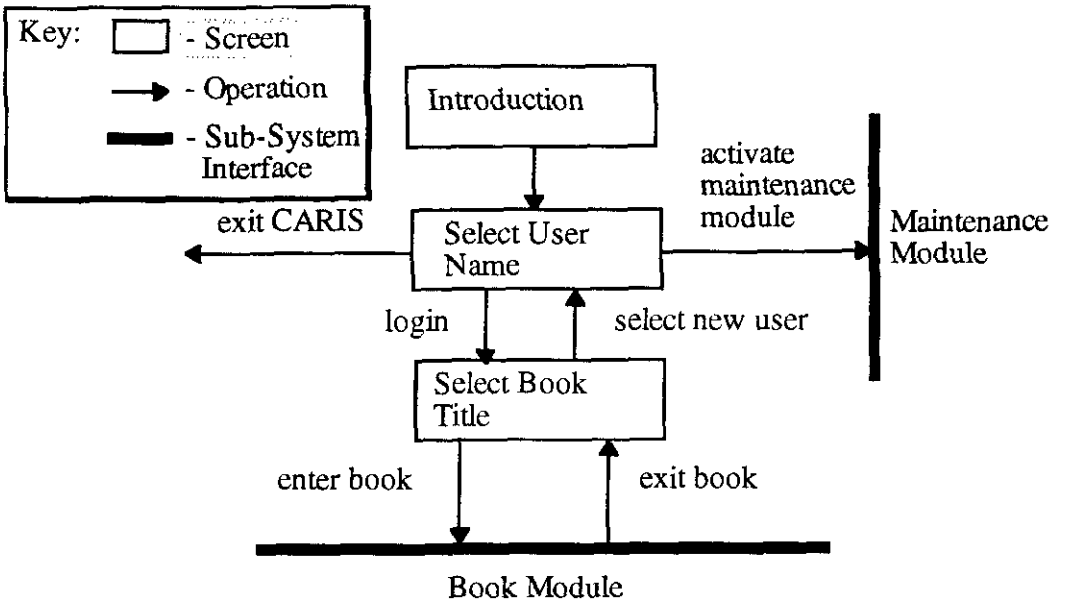


Figure 5.2: Control flow of management module

5.3.4 Book module

In common with commercially available software, each screen (Figure 5.3) replicates the illustration and text on each page of the original hard copy book. Also, left and right arrow buttons allow the user to page backwards and forwards when they wish.

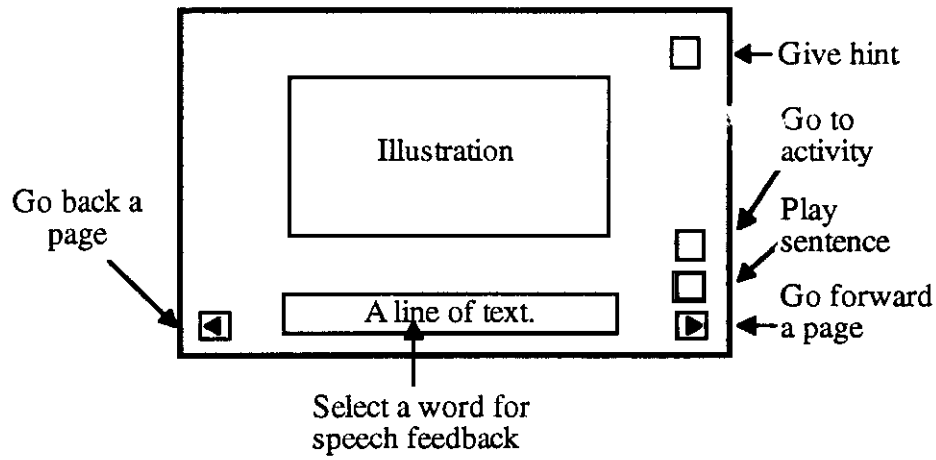


Figure 5.3: Page design

Much current software includes animations on each page, which are activated by clicking on hot spots on the screen. Including animations can enhance both motivational levels and learning outcomes in educational software (Milheim, 1993). However, it has been shown (Scoresby, 1996; Underwood & Underwood, 1996) that such features can be distracting in reading software and can adversely affect the level of comprehension. Some ranges of talking books contain large numbers of animations, many of which bear little or no relation to the story being told. Illustrations in the current version of CARIS are not animated, although it is considered that appropriate use of animations can be beneficial.

The feedback mechanisms

A teacher listening to a young child reading may give plenty of encouragement and praise, particularly for those children who are lacking in confidence. Self-correction techniques are encouraged and this may require a teacher to give one or more hints on how best to approach this. The system provides several different types of feedback:

- **Instant whole word feedback** for individual words selected in the sight recognition strand of the reading scheme.
- **Instant feedback at the level of onset and rime.** If the book is from the phonic awareness strand then feedback is given at the level of onset and rime, but only if it is suitable for the individual. In this case, the onset is spoken whilst the onset is highlighted on the screen. Then the rime is spoken whilst the rime is highlighted on the screen. Finally the whole word is spoken and simultaneously highlighted. This is intended to develop awareness of onset and rime and promote blending skills.

- **Hints** relate to either the illustration, the first letter, the onset, the rime, going back to the beginning of a sentence, or are more general and encourage the child to try again. The hint mechanism is described in more detail below.
- **Sentence feedback** when the complete text on the page is 'spoken' with natural intonation and expression, providing an appropriate model for the beginning reader and a method of self-checking accuracy.
- **Praise** is only relevant within activities because word accuracy cannot be detected during reading of books.
- **Encouragement,** which could be given at any stage, but is more suitable for readers lacking in confidence.
- **Verbal Instruction,** because beginning readers benefit from spoken rather than written instructions, and this can contribute to their sense of independence.

The hint mechanism

A teacher could feel that a child may benefit from receiving a hint to promote independent decoding of unknown words rather than being given the word instantly. The teacher may force this to happen in the software by selecting an option in the Maintenance Module. In this case, the first time any word is selected for feedback on a page, the Diagnostic Module is accessed. It randomly selects a reading cue that has been identified as being under-utilised by the particular user and a hint relating to this is given. For example, if the software detects underuse of meaning and syntax, a child could be given the hint "Try starting again from the beginning of the sentence." A child may however select a word to confirm their own guess or have already tried all possible strategies for decoding and failed. Therefore, the second time that the word is selected its pronunciation is given instantly.

If hints are not specified as being mandatory by the teacher then the child is able to choose between instant feedback and a clue, enhancing student autonomy and control. A separate button on each page of every story enables a child to optionally request a hint rather than obtaining the whole word instantly. Requesting a hint may not be perceived as failing and may assist those children who are fearful of making mistakes.

5.3.5 Diagnostic module

The Diagnostic Module provides support for both the Book Module and the Activities Module by modelling and analysing the use of specific reading cues commonly used in the

decoding of unknown words. In addition, it models the confidence level of the user. It has three key roles:

- to adjust the reading cue and confidence models, as appropriate for each possible interaction type
- to analyse words selected for speech feedback, identifying which cues may have been under-utilised and adjusting the cue models accordingly
- to detect under-utilised reading cues, in order to select a hint or an activity which can provide coaching in an area of weakness

The modelling mechanism and the three key roles are described in more detail below, together with an example of how the diagnostic module operates in practice.

The modelling mechanism

The modelling mechanism is based on the approach taken in del Soldato's (1994) tutoring system design, named MORE, which used a linear scale to model students' confidence and independence levels. Due to the nature of the tasks in MORE, the information that the system could obtain and analyse was much richer than that which can be obtained from CARIS. For example, MORE allowed users to choose which task they undertook and used this information to adjust their model of confidence level depending upon the level of difficulty of the exercise selected. The children using CARIS have no choice over the difficulty of the activities that they are asked to complete. The approach taken in CARIS was much more simplistic.

The independence level refers, in this research, to the type of feedback a student might need or what level of support they might require. It is controlled entirely by the teacher on the basis that it is difficult for the system to detect and analyse it. Prior to using the program a teacher can configure the feedback mechanisms to be employed, via the Maintenance Module described in section 5.3.7, such that individual interaction preferences can be met.

Use of reading cues is modelled on a linear scale for each of the five types specified: initial letter, common word family, meaning and syntax, illustration and key word recognition. Adjustments made to the cue models depend on:

- input from the teacher
- an analysis of the properties of the words selected for feedback

- the results of the activities undertaken

Meaning and syntax are grouped together to simplify the design and implementation of the Diagnostic Module. The differences between the two cue types are difficult to separate without storing information about the word selected and its position in the sentence. They are interrelated and it makes sense to group them together in this way.

The confidence level is also modelled on a linear scale. It is adjusted according to: information provided by the teachers, the amount of help requested by the children, and the results of activities. Clay (1993b) defines that an instructional text should include 90-94% familiar words, which implies that up to 10% of the words in a text will be unknown to the child. It is assumed that the child's teacher would select a book of an appropriate difficulty level. If the number of help requests exceeds 10% of the total number of words in the book, the confidence model is decremented on each further request.

Adjusting the reading cue and confidence models

The possible adjustments for each model type are specified in Table 5.1 and Table 5.2.

Table 5.1: Possible adjustments to reading cue models

Reading Cue	Increment	Decrement
initial letter	hint given for initial letter success at onset activity	word feedback selected failure at onset activity
illustration	hint given for illustration success at illustration activity	word feedback selected for noun failure at illustration activity
common word family	hint given for common word family success at rime activity	word feedback selected for word in common family failure at rime activity
meaning and syntax	hint given for meaning and syntax success at meaning and syntax activity	word feedback selected failure at meaning and syntax activity
key word	hint given for key word success at key word activity	word feedback selected for key word failure at key word activity

Table 5.2: Possible adjustments to confidence model

	Increment	Decrement
Confidence	if succeeds at any activity if completes book without selecting any words (large increment) if completes book selecting less than 10% of words (small increment)	if fails at any activity if selects a word for feedback (not hints) and have previously selected 10% or more words if does not attempt prediction activity

Analysing words selected for speech feedback

The diagnostic module analyses the words selected for feedback and ascertains which reading cues may not be being employed in the decoding process. Each word in every talking book has one or more properties defined for it, the types of which are summarised in Table 5.3. The definitions are based on the following assumptions:

- that a word can be decoded by analogy if it is from a common word family (e.g. 'would', 'should', 'could')
- that children are expected to recognise the key words by sight
- that the illustration can be used to decode the word if it is a noun and appears in the associated picture
- that looking at the initial letter may be used for any word

When a word is selected by a child for speech feedback then for each property associated with the word, the related cue model is decremented. A child may request a hint or use the speech feedback to confirm a guess. Nevertheless, it is assumed that they are still unsure of the selected word and cannot have utilised all possible cues. It is therefore acceptable to adjust the models when any form of feedback is requested. The models will only be adjusted if the word has not already been selected in the current book. This is because one interaction style identified in the pilot study (chapter 4, section 4.3.1) indicated that some children may select most of the words, even if they can read them.

Table 5.3: The relationship between word-property and strategy requirement

Property	Strategies required
Common word family	onset and rime (phonic)
Key word	visual recognition (visual)
Concrete noun	illustration (visual), meaning, syntax
Verb	meaning, syntax
Article (the, a), conjunction (and), pronoun, preposition, adjective, adverb	syntax, meaning

Detecting under-utilised reading cues

Both the Activity Module and the Hint Mechanism access the Diagnostic Module in order to identify an under-utilised reading cue. The activities invoked and the hints provided are then specific to an individual's needs and provide coaching in areas of weakness. More than one reading cue may be identified as being under-utilised. In this case, one of these is selected at random. Alternatively, if no cues are be identified as under-utilised then one is randomly selected from the complete range. This was designed to provide variety and hence contribute to improving motivational levels.

An example of how the diagnostic module operates in practice

Table 5.4 gives a step by step example of model adjustments made during normal use of the software. Adjustments made to models after specific actions have been taken are indicated by shading. The examples given include selecting words for feedback with different properties, requesting a hint for a word, and both successful and unsuccessful attempts at activities. Those models with a value of four or less are considered to represent cues that are under-utilised. The hint type given and the activity types selected are such that coaching is provided in an area of weakness. Fewer than ten words were selected for feedback in this example throughout the whole book. Therefore, the confidence model is incremented because the child is considered to have read the book with an acceptable level of support.

Table 5.4: Example of adjustments to models during normal use

	con-fidence (con)	initial letter (il)	illustra-tion (illus)	common word family (cwf)	meaning and syntax (ms)	key word (kw)
initial default values	7 (high)	5 (medium)	5 (medium)	5 (medium)	5 (medium)	5 (medium)
selects word: cat (il, illus, cwf, ms)	7	**4**	**4**	**4**	**4**	5
selects word: back (il, cwf, ms, kw)	7	**3**	4	**3**	**3**	**4**
activity: illus successful completion (con, illus)	**8**	3	**5**	3	3	4
activity: kw successful completion (con, kw)	**9**	3	5	3	3	**5**
selects hint: book (ms hint given)	9	3	5	3	**4**	5
activity: ms not successful (con, ms)	**8**	3	5	3	**3**	5
activity: ms not successful (con, ms)	**7**	3	5	3	**2**	5
activity: ms successful completion (con, ms)	**8**	3	5	3	**3**	5
finishes book and exits (con)	**9**	3	5	3	3	5

5.3.6 Activity module

The activities are invoked from a button, provided on each page of the book except the title page. There is one activity designed to provide coaching for: meaning and syntax, illustration and onset. In addition, there are two further activities: one to develop sight

recognition of key vocabulary and one designed to develop prediction skills. Due to time constraints the rime activity was not implemented but is described here for completeness. The activities are described in more detail below with simple examples of how they might be used. Concrete examples of activity use are provided in chapter 6, section 6.5.3, which describes the implementation of the software.

The variety of activities offered is maximised whilst ensuring that the implementation was efficient. The functions were designed to be generic such that they can be invoked from any page in any book but incorporate the illustration and text from the page from which the module was activated. Thus, even if the same activity is generated as that which was invoked on a previous page, it will be different with respect to the illustration and words used.

Meaning and syntax

A common activity found in the classroom consists of cutting up a sentence into individual words, shuffling them up and asking a child or group of children to reconstruct the original text. This kind of task encourages the children to think about the structure of a sentence, or syntax, and is also concerned with meaning or ensuring that the reconstructed text makes sense. This kind of activity was considered to be a value-added use of technology because although it can be done manually, it is easier and faster when implemented on a computer. In addition, the accuracy can be assessed immediately and useful feedback can be given instantly. On invoking this activity, the illustration from the page on which the activity button was pressed is shown together with the first sentence from the associated text. The sentence is broken down into individual words, which are shuffled and displayed in a random order. The user is then asked to reconstruct the sentence. The concepts behind this activity could be difficult to grasp for some children. Hence, only three attempts are allowed followed by the correct answer being demonstrated.

For example, the sentence "Mrs Hall looked at his money." could be shuffled and presented to the child as a series of disparate words: 'money', 'Mrs', 'his', 'looked', 'Hall', and 'at'. An illustration, depicting Mrs Hall looking at the money on the table, from the page

containing the sentence would also be displayed. The user would be given three attempts at reconstructing the sentence, by dragging the words from the top of the screen to a grid representing the sentence as indicated in Figure 5.4 below.

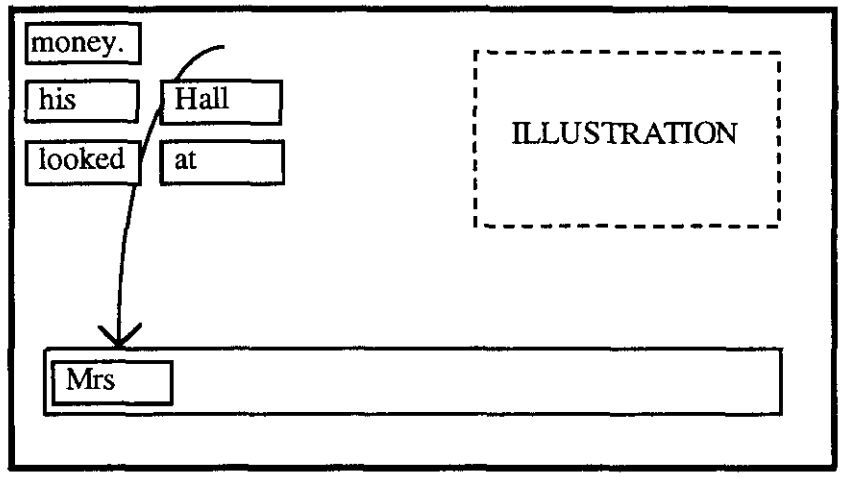


Figure 5.4:Reconstructing the sentence by moving word cards on the screen

Illustration

The concept underlying the design of this activity is to encourage the child to make use of the illustration. Due to the nature of the reading scheme selected, the text on each page is assumed to be pictorially cued. Guessing what happens in the story next by looking at the picture should generate a sentence that is very similar, if not the same, as that written on the page. Thus, it is important to encourage the children to look carefully at the illustrations. CARIS shows the illustration from the page on which the activity button was pressed, and displays a number of nouns that occur in the text and appear in the illustrations. The child is asked to identify which of the nouns are visible in the illustration and which are not. Again, they are only allowed three attempts and then the correct answer is displayed.

For example, in the book entitled "A Book for Jack", the illustration on the first page depicts some school children at a book sale. One character, Sam, is looking at a book containing a picture of a panda. In this activity, the illustration would be displayed and the following nouns would be presented: 'book', 'money', 'Jack', 'Sam', 'panda', and 'pens'. The child

would be asked to select those nouns that were contained in the illustration: 'book', 'Sam', and 'panda'.

Onset

As this type of activity relates to the development of phonic skills, it is only applicable, and thus only available, in those books in the pattern and rhyme strand of the reading scheme. At this particular stage of the reading scheme, it is specified by the publishers that an awareness of the following initial sounds should be developed:

- Blends and digraphs - cl, sw, st, br, cr, ch
- Initial sounds - b, c, d, e, f, j, p, s, t

Accordingly, the activity focuses on these onsets. The child is shown and told the ending of a word and asked to identify the correct onset from a choice of three. As the choice is limited, the activity continues until the onset is selected. When the correct choice is made, the child is given the onset then the remaining sound of the word and finally the whole word. This reinforces awareness of the onset further and develops the skill of blending. This type of activity was not related to a specific page, but each book focuses upon three initial sounds and three blends or digraphs, offering two distinct versions of the exercise. The target word is randomly generated from a limited list of vocabulary that start with one of the onsets and are contained in the books.

For example, the child could be shown the word ending 'air' and given its pronunciation, and then asked to select the correct onset from three blends: 'br', 'cr', and 'ch'. If they were to select an incorrect blend, then they would be prompted to try again. Upon selecting the correct blend, 'ch', the whole word would be displayed on the screen and the onset, the rime and then the whole word feedback would be given.

Rime

As above, this type of activity would only have been available in the pattern and rhyme strand of the reading scheme. The text in this strand of the reading scheme includes a lot of rhyming words, which would have formed the basis for this activity. The child would have

been given a target word and asked to select another one that rhymed from a list of three alternatives. Again, as the choice would have been limited, the activity would have continued until the correct choice was made. Each book would have focused upon three common word families used frequently in the text and the target word would have been randomly generated from a limited list.

For example, the user might have been presented with the target word 'book' together with its pronunciation and asked to select another word that rhymed with it from: 'baby', 'look', and 'made'. If they selected an incorrect word, its pronunciation would have been given and they would have been asked to try again. Upon selecting the correct word, in this case 'look' they would have been given the pronunciation of both this and the target word, 'took'.

Key word

One of the agreements made with the publisher of the reading scheme was that the software developed should complement the structure and philosophy of the materials. As one of the primary aims of the scheme is to develop sight recognition of a limited vocabulary range, termed key words, an activity to support this has been included. A common teaching strategy used by an adult listening to a child reading is to ask the child to identify a particular word contained in a sentence. This can only be achieved with some assistance, normally from an adult or a more experienced peer. This kind of activity was therefore considered to be a value-added use of technology. The illustration and text are shown to the child, who is asked to click on a key word in the sentence. If more than one key word is contained in the text, one is selected at random, which ensures that the variety level is maximised. The activity continues until the correct word is selected because there are only a limited number of words in the text on each page. Speech feedback is given for each word that is clicked on by the user, alerting them to any errors made.

For example, the sentence "He wanted a book about spiders." could be presented to the user, together with its associated illustration. The user could be asked to read the sentence again and click on the word 'about', identified within the reading scheme as being a key word. If the child clicked on the word 'wanted', whole word speech feedback would be

given and the original instruction would be repeated. The activity would be completed when the child clicked on the correct word in the sentence.

Prediction

The skill of prediction is very important in early reading and is something that is generally encouraged by teachers in the classroom. Therefore, this type of activity was included in CARIS. After the title page the child is asked to predict what the story will be about by typing text into a text box (Figure 5.5). At a suitable point, somewhere in the middle of the story, the user is prompted to predict what might happen next. With both prediction screens, the child will not see the screens again once they have paged forward beyond the activity, even if they subsequently page back. This will ensure that the child cannot find out what is going to happen and then return to the screen. Similarly, the screens are only shown the first time that the child enters a book.

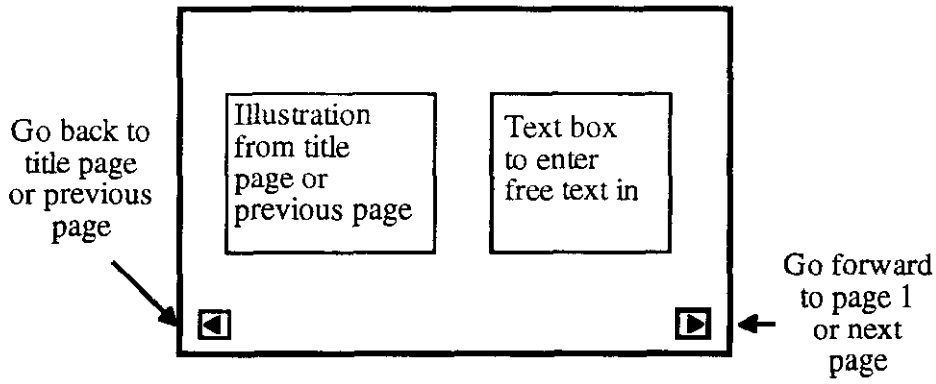


Figure 5.5: Page design for prediction activity

5.3.7 Maintenance module

The information that can be obtained about a user's preferred style of interaction or desired level of independence is limited. Different children will benefit from different teaching strategies and it is important to enable the teacher to specify some of these options (Figure 5.6). This maximises the flexibility and adaptivity of CARIS to suit an individual learner's needs and preferences.

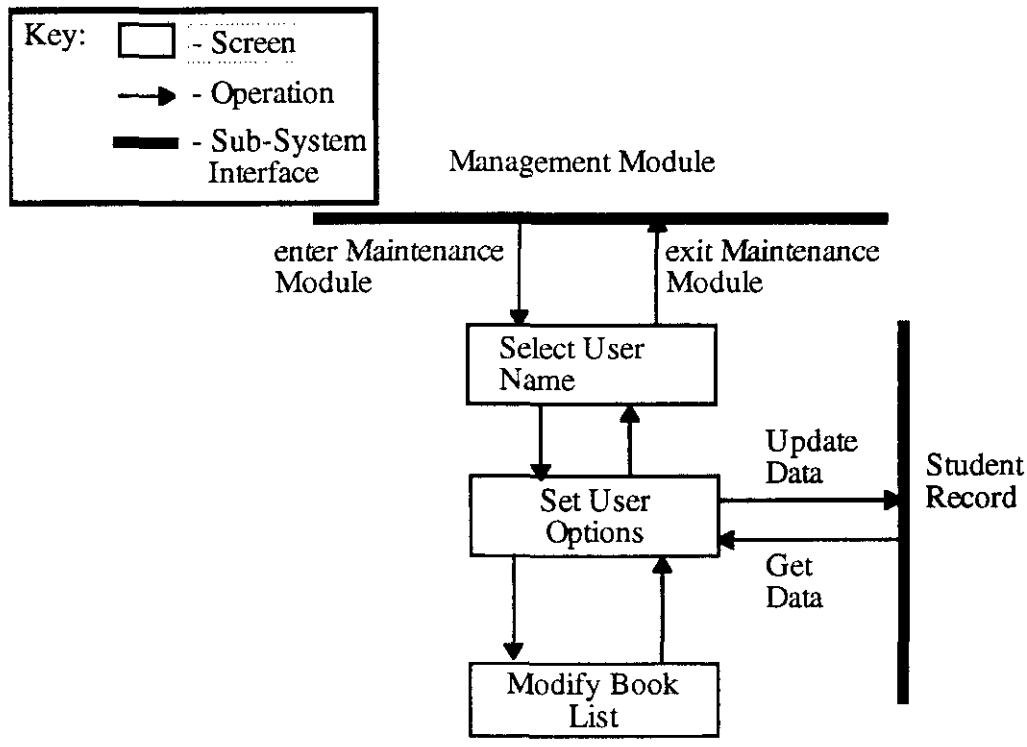


Figure 5.6: Flow of control in maintenance module

The options that can be configured are specified in Table 5.5.

Table 5.5: Range and type of configurable options

Configurable Options	Range	Type
Reading Cue	illustration initial letter common word family meaning and syntax key word	check boxes
Confidence Level	low medium high	radio button
Feedback	whole word onset and rime	radio button
Other Options	free play a hint first sentence read enabled	check boxes

The 'Reading Cue' option enables a teacher to specify which strategies, in their opinion, are being utilised successfully by an individual child and also over-ride the system diagnosis at

any point in time. One or more of the reading cues may be selected. Similarly, the confidence level can be specified initially and adjusted at any point in time. The teacher can specify the confidence level as being low, medium, or high.

The 'Feedback' option enables a teacher to specify whether or not speech feedback at the level of onset and rime is appropriate for an individual user. If the onset and rime option is selected then whole word speech feedback is given in those books from the sight recognition strand of the reading scheme. The books from the pattern and rhyme strand of the reading scheme are designed to support the development of phonological awareness. In these texts, the speech feedback given is as that described above: onset, rime, and finally, the whole word.

The 'Free Play' option disables the innovative features in CARIS, providing a version of the software that replicates the basic functionality of current commercially available software. This enabled a comparative study of two versions of the software to be conducted (chapter 7). The 'Hint First' option enables a teacher to make one hint mandatory for an individual (as described in section 5.3.4) and this can be changed at any point in time. The 'Sentence Read Enabled' option allows a teacher to specify whether or not this feature is available. The default is to enable speech feedback for sentences to be obtained, but if this option is de-selected then the speech feedback button on each page of a book is disabled.

5.4 A survey of practitioners' use of talking books

The purpose of this formative survey was to elicit the opinions of practitioners on the use of talking books in the classroom. Firstly, to verify that the current format of such a resource was perceived to be beneficial. Secondly, to demonstrate that the design presented in section 5.3 was realistic and would complement teaching strategies. One focus of the questionnaire was to determine whether or not practitioners would like to see additional feedback modes, such as segmented feedback or hints, in future implementations.

5.4.1 Survey design

The respondents were asked to identify the range of talking books being used in the classroom and which year groups were using the software. They were also asked to specify

whether or not the whole class were involved or only those children identified as requiring some additional input. They were also asked how the software had been integrated into normal classroom practice with regard to reading tuition. This included the frequency of use, the size of each group assigned to work at one machine, the level of adult support offered during use, and for what reasons talking books were being used. Next, the respondents were asked to specify how often a number of software features were used.

Finally, information was collated on how future talking book design could be enhanced to provide additional functionality. A particular focus of this section was the type of speech feedback that might be required on the selection of an individual word by a child. The respondents were asked to indicate their interest in the provision of feedback both at the level of onset and rime, and the level of syllable and/or sub-syllables. They were also asked if the option to obtain a hint to assist the child in decoding unknown words independently, rather than obtaining the whole word immediately, would be potentially beneficial. An opportunity to express any further comments was included at the end of questionnaire.

Two copies of the questionnaire (appendix G) were sent to 1230 schools from around the UK who were known to have purchased Sherston Software's Talking Stories (Bonham & Bonham, 1993). These schools included a wide variety of mainstream primary and combined schools, as well as a number of special schools serving all age ranges.

5.4.2 Results

The response rate was 30% and a total of 494 completed questionnaires were received, from 371 schools. The year groups using the software ranged from nursery and reception classes through to year 2 in which all children were generally involved. The software was also used by those children in year 3 through to year 11 experiencing difficulties in learning to read, many of who were in special schools. The results are presented in two graphs. Figure 5.7 shows the percentage of classes using talking books by year group. Figure 5.8 shows what percentages of the teachers were using the talking books with children progressing normally

pupils experiencing difficulties in learning to read. The percentages were calculated by totalling the responses from all returned questionnaires (494). The possible responses were

not mutually exclusive. For example, a class may consist of more than one year group. In some cases the IT Co-ordinator had completed the questionnaire on behalf of the whole school and thus, several year groups were selected as were both reader types. The software is most commonly used with children progressing normally in learning to read and in the early years of schooling. It should be noted that the results obtained may have been influenced by the current range of talking books, which are based on texts designed for early readers in reception and year 1 classes.

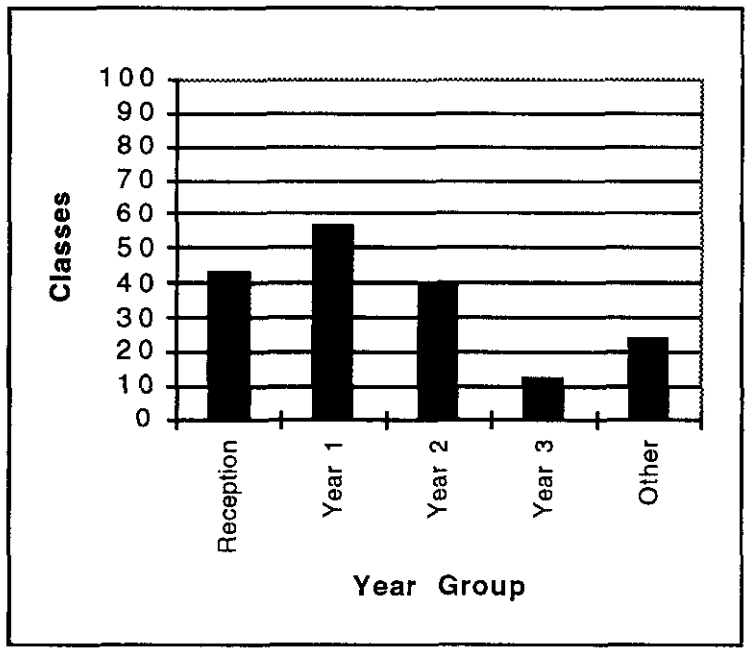


Figure 5.7: Percentage of classes using talking books by year group

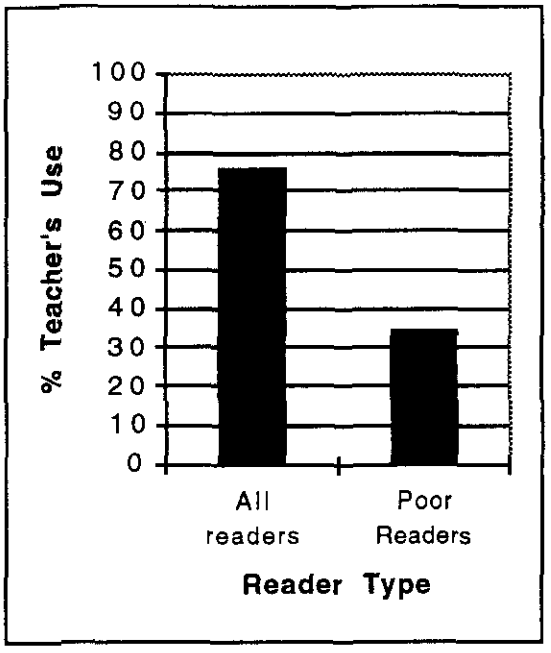


Figure 5.8: Percentage of teacher's use of talking books by reader type

The frequency of use was generally identified as varying depending on individual children's needs. Many respondents indicated that they were using the software with more than one grouping of children from a choice of: one at a time, pairs, and groups of three or more. The majority of practitioners (82%) were using the software mainly with pairs of children, but occasionally with individuals (29%), though this seemed to be with readers experiencing difficulties rather than those progressing normally. A smaller number of respondents (11%) specified that the software was usually used by groups of three or more children at one computer. It was reported that most of the children (87%) were able to use the software independently with minimal adult support, after receiving some initial training. Several of the classes involved older children experiencing difficulties learning to read or those identified as having special needs. In these cases, an adult was generally available to work with these children for the duration of the session. A variety of reasons were given for the use of the software including:

- providing supplementary practice
- contributing towards general reading skills such as familiarity with books
- to boost the confidence of those children with low self-esteem with regard to reading
- to motivate children and expose them to the pleasures of reading

This reasoning was further justified by some of the optional comments received:

> Has encouraged the children to look at or read the books for themselves. You can see their confidence develop as they use this software.
>
> I find the stories are very valuable in the acquisition of reading skills as they provide a different context to practise reading.
>
> I have used Talking Stories a lot over the past couple of years - the children are always highly motivated, regardless of their ability.

The responses to the survey question concerning the perceived ease of the software are indicated in Figure 5.9. It was reported that the majority of children usually or always find the software easy to use (96%).

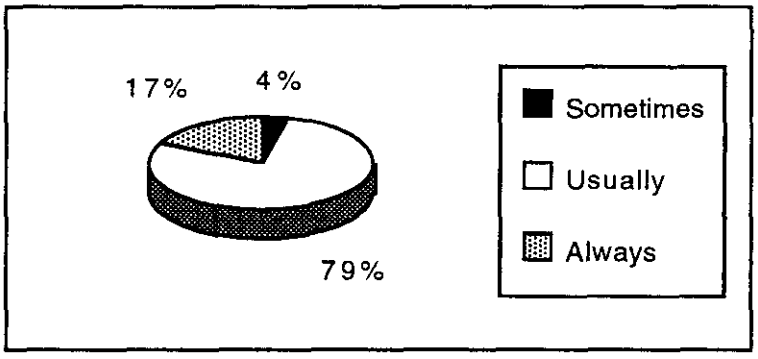


Figure 5.9: Perceived ease of use of talking book software

The respondents were asked to identify how often they used four features generally found in talking book software:

- keeping a log of individual words that have been requested for speech feedback
- having the entire story read aloud to the child (sometimes termed autoplay)
- using an index to move directly to a particular page
- completing the key word exercises at the end of the story, where the child is asked to read a word in isolation and then in the context of the sentence it appeared in

The results obtained from these questions are presented Table 5.6. A total of 88% of respondents indicated that they made no or little use of the facility to log words requested, although 7% found it useful for monitoring children's progress. Miller et al. (1994) found that a child may either read a word incorrectly but believe it to be correct, or avoid requesting help for other reasons. Alternatively, children may use the feedback facility to confirm a prediction. Thus, a misleading picture of their progress may emerge from word logs alone if

a child's use of talking books is not monitored occasionally. This is corroborated by the findings of the pilot study (chapter 4).

Table 5.6: Responses to questions on use of features of talking book software (%)

	Word log	Autoplay	Index	Key word
Null Response	4	4	3	4
Never	53	31	34	19
Sometimes	36	54	46	49
Usually	6	10	12	22
Always	1	1	5	6

Not every practitioner used the facility that read the entire story to the child initially, and those that did so indicated that they only used this feature occasionally. Similar responses were received regarding the use of an index to enable a child to move directly to a specific page of the book. The key word activity was used by the majority of the respondents (77%) either sometimes, usually or always. An optional question concerning additional features or other comments was completed by 131 respondents. A total of 17 of these respondents said that they would like to see the software enhanced with simple comprehension exercises, cloze activities and various support materials to accompany its use.

The final section of the questionnaire revealed that most practitioners (89%) would like the software to offer at least one of three further feedback types: onset and rime, syllable and a hint. Figure 5.10 below shows the percentage of affirmative responses for each individual feedback type. The most popular choices were feedback at the level of onset and rime, and a hint facility, to encourage children to decode unknown words independently rather than being given the word immediately.

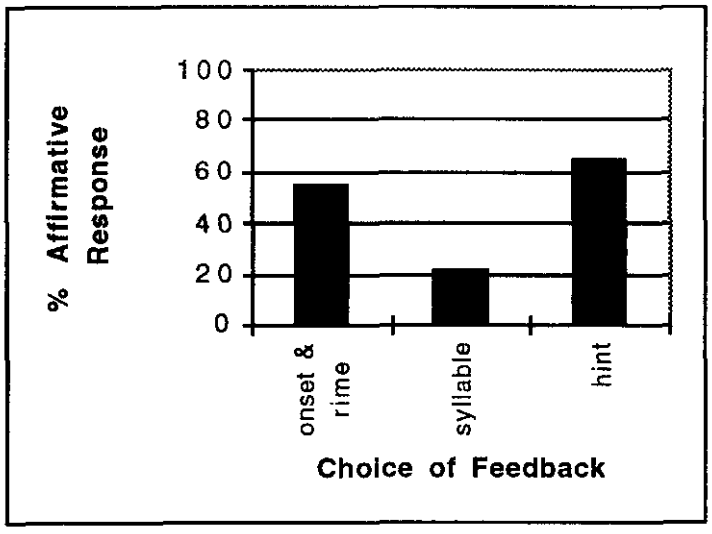


Figure 5.10: Percentage of affirmative responses for additional feedback modes

5.4.3 Conclusions and implications for software design

From the results of the survey, the current format of talking book software appears to have been successful both with early readers who are progressing normally and those older readers who are struggling. However, there are several features which practitioners would like to see in future implementations. A greater variety of supporting activities to reinforce reading skills and proficiencies were identified as being potentially beneficial. It was also indicated that both incorporating hints for decoding unknown words and feedback for individual words at the level of onset and rime were considered to be worthwhile. Thus, the design presented in section 5.3, as well as being theoretically grounded, addresses the issues raised by the practitioners themselves.

5.5 Discussion: Justification of design

A justification of the system design in relation to the objectives is now presented and discussed. The basic format of the talking book software incorporates all the strengths of such a resource, which were identified in the pilot study (chapter 4). The way in which the weaknesses identified in the pilot study have been addressed in this design is described in this section. The adaptive nature of the system and how the affective issues were addressed are also highlighted.

5.5.1 How the weaknesses have been addressed

Below is a comment relating to each point raised in section 5.2.3:

- One style of interaction observed was that of a child becoming reliant on instant feedback. As discussed above, this may not always be detrimental and therefore, the system design is flexible. A child may be allowed to engage in this style of interaction, if it is perceived as being beneficial, or forced to receive a hint first for every word selected.
- The facility to obtain feedback for sentences is configurable such that a child who might opt to have the entire book read to them, and adopt a passive role, can be prevented from doing so.
- A variety of activities have been included to ensure that learning is reinforced and interactivity is maximised.
- Two additional feedback types have been incorporated. Firstly, to provide feedback at the level of onset and rime. Secondly, to enable a child to obtain a hint, which may prompt independent decoding of unknown words.
- It is not currently feasible to provide a mechanism to detect and correct errors made if help is not requested by the user. It is possible that the incorporation of speech recognition technology will address this issue (Mostow et al., 1994). However, by enabling a child to obtain sentence level feedback and by instructing them to select it after attempting to read the text themselves, before proceeding to the next page, this situation may be overcome.
- The activities and hints have been specifically designed to provide coaching in the use of reading cues and prediction skills.

5.5.2 Adaptivity

Information is obtained from a variety of sources, both dynamically whilst the software is in operation and via an interface to allow a teacher to further specify details about individual users. Although this could be further enhanced by obtaining information from the users themselves concerning their self-perceptions of knowledge and skills, it is considered to be inappropriate for the target age.

Control and choice are given to the user where possible, but this can be over-ridden if the teacher considers that this could be detrimental to an individual or abused by them. For example, a hint may be optional but can also be selected as being mandatory and the facility to obtain feedback at the sentence level can be enabled or disabled. This means that all identified styles of interaction can be accommodated, but ensures that inappropriate use of the software, such as adopting a passive role, is avoided. There are models for each of the five specific reading cues. These are modified according to: initial information provided by the teacher, words selected for feedback, and the results of completed activities. This ensures that both the activities and hints provide coaching in areas of weakness. If feedback

at the level of onset and rime is considered to be beneficial for an individual user, then it can be enabled. Thus, the software can be tailored to an individual's specific needs and preferred learning style.

5.5.3 Affective considerations

The motivational aspects of computer assisted learning are harnessed to maximum effect. As concluded in chapter 2, the use of sound and colourful graphics, independence and control, and interactivity can all contribute towards improving motivation and hence extend periods of interest and concentration. The design of the activities is such that interactivity and variety are maximised. The use of sound and graphics are prominent design features, and independence and control, although being configurable for individuals, are also maximised where possible.

The confidence level will be monitored such that the levels of support, praise and encouragement can be altered to meet the needs of each individual. Strategies for building self-confidence such as those cited by Keller (1987) have been incorporated into the software. These include avoiding direct negative feedback and creating opportunities for learner success, which are underlying concepts in the design of the activities.

5.6 Conclusion

An innovative design for talking book software, which is adaptive and incorporates affective as well as cognitive teaching strategies, has been described. Chapter 6 describes and discusses the implementation of this design.

6. Implementation

6.1 Introduction

This chapter discusses how the design presented in chapter 5 was implemented. The development environment is described, including the software tools utilised and the hardware that was required. The development approach undertaken is presented next, in which the initial prototype is described and design modifications are highlighted. This is followed by the results of three formative evaluation stages and the outcomes of these. Finally, several interaction examples are illustrated including: initial start-up, normal use of book, and the completion of each of the activities.

6.2 Development environment

The software was developed on a Macintosh Performa 5300 platform using a multimedia authoring package, Macromedia Director 4.0 (1994). This facilitated the incorporation of graphics, sounds and text. A scripting language, Lingo, for supporting interactivity was also provided.

The illustrations were scanned in from the original artwork for the selected texts (with permission of the publishing company, Ginn & Company Ltd.), using a standard flatbed scanner at a resolution of 72 dpi with 256 colours. This was to ensure that the storage requirements were minimised. The scanned images were touched up as required. For example, some illustrations in the original books had been enhanced with text during the printing process. The fonts used in the computerised texts were selected such that they matched original books as closely as possible.

Sounds were digitally recorded for each word contained in the texts using the researcher's voice. They were recorded directly onto the Macintosh machine with a microphone and Macromedia SoundEdit 16 (1994), at 22.254 kHz. No compression was applied so that the highest possible sound quality was achieved, but at the expense of the greater storage requirements. Each word was recorded without expression. Each sentence was also recorded with expression and intonation and stored as separate word sound files. For those

words contained in the strand of the reading scheme designed to support the development of phonological awareness, the onsets and rimes were also recorded without expression. In addition, recordings were made of instructions, hints, and other feedback. It should be noted that the sounds were recorded with a south-eastern accent, which could be considered to be close to standard English pronunciations. For areas in the UK with strong regional dialects it would be possible to record appropriate sounds, but it would be time-consuming and has not, as yet, been implemented in British commercial talking books.

One objective of the software design was that it should emulate a supportive human tutor and consider affective aspects of teaching reading. Feedback in the activities was therefore given such that direct negative statements were avoided. As a form of encouragement, if a child correctly completes an activity they will be given the feedback:

Hooray....Well done!

If they attempt an activity but are not successful, they may be given feedback such as:

Ahhh! Never mind. Try again!
Not quite. What you can see in the picture is ...
Not quite. The sentence on the page is ...

6.3 Development approach

6.3.1 Prototype

A rapid prototyping approach was undertaken whereby the basic book structure was implemented initially. In the second stage the additional modules and adaptive features were implemented. A book from each of the two selected reading scheme strands (sight recognition and phonological awareness) was produced.

The scripts generated were designed to be generic so that minimal modifications were required upon the development of further books. For example, the script to give a hint when a click is detected on a word is given in Figure 6.1.

The number of the word in the text field (e.g. the first, second, third, and so on, in the sentence) is generated by a Director Lingo function, the mouseWord, and passed into the script. The first sub-function, fGetWord(), obtains the actual word clicked on by using this number. It also accesses a global variable, gCastNum, which holds the number of the 'cast

member' or variable containing the text on the page. For example, if the sentence on the page was "He put the baby book back." and the user clicked on 'baby' then fHintWord() would be called with the value 4. The function fGetWord() would get the sentence from the page and return the fourth word. Thus, the script can be called on any page of any book.

```
--  fHintWord
--  Description: get current word and randomly select hint
--  Parameters:  pCurrentWord - current word number
--  Return Value: NONE
--  Caveats:   No error checking!
--       Child's record will be updated
--       (words requested)

on fHintWord pCurrentWord

 -- Get current word using word number passed in
 set lWord = fGetWord(pCurrentWord)

 -- Before determining type of hint, update child's record.
 -- Even though they haven't requested direct help, its still a word
 -- they do not recognise immediately.
 set lWordType = fUpdateRecord(lWord)

 -- Randomly select a strategy and check if its underutilised, until
 -- we find one that is (20 iterations maximum).
 -- Then get a hint relating to the chosen underutilised strategy or
 -- last randomly generated strategy and speak it.
 set lPosition = fGetWeakStrategy()
 set lHint = fGetHint(lPosition, lWordType)

 -- Log information: which page, which word, which hint
 set gActivityResults = gActivityResults & RETURN & "Giving hint ¬
                        on page:" && the frame && "Word:" && lWord ¬
                        && ", Hint:" && lHint

 -- Speak the hint selected.
 fSound(lHint)
 puppetSound 0

 -- Since we've given specific coaching in one strategy, increment
 -- model
 fIncrement(getPropAt(gInfoList,lPosition), 1)

end fHintWord
```

Figure 6.1: Script which gives hint for selected word on screen

Any Lingo functions that are specific to an individual book have been isolated in one set of scripts. One of the functions is automatically invoked as the book is entered (startMovie) and the others are called as required. This facilitates the rapid generation of further books, minimising the possibility of introducing errors or bugs (see section 6.3.3).

For the purpose of data analysis (chapter 7), many of the scripts contain code to log information relating to an individual's use of the talking books. The information recorded includes:

- the book title
- the time the book was started
- the time the book was completed
- a summary list of which words were selected and how many times
- the reading cue and confidence models on leaving the book, if applicable
- page turning
- words selected on specific pages
- hints given, if applicable
- onset and rime feedback given, if applicable
- results of activities undertaken and number of attempts, if applicable

6.3.2 Design modifications

Onset and rime refer to the beginning and remaining sound of each syllable in a word (Goswami, 1996). Thus, the word 'downstairs' can be decomposed into two syllables: 'down' and 'stairs'. Each of these syllables has its own onset and rime:

'down'	=	'd'	+	'own'
'stairs'	=	'st'	+	'airs'

Due to time constraints, the implementation of this form of feedback was simplified and multi-syllabic words were treated in the same way as those consisting of a single syllable. A word that starts with a vowel (e.g. 'and') has no onset and so the whole word pronunciation is given instantly. Otherwise, the initial sound is given followed by the remaining sound, and finally the whole word as specified in chapter 5, section 5.3.4. However, it is acknowledged that, in the case of multi-syllabic words, the feedback may not assist in developing an awareness of the use of common word families or word analogies. For example, the words 'downstairs' and 'vegetables' would be given as:

'd',	'ownstairs',	'downstairs'
'v',	'egetables',	'vegetables'

6.3.3 Procedure for generating new books

The starting point for generating a new book was to use the basic shell, into which the relevant illustrations and text were inserted. Each sentence was recorded with expression and intonation as appropriate for the story being told. The recorded words were separated and stored in separate sound files. Each book has its own start-up script in which all functions that are required to be tailored for an individual text are contained. This includes:

- a flag indicating which strand of the reading scheme the text was from (gOnsetAndRime - see appendix H, Figure H.1)
- a list of the key words used in the text (gClickList - see appendix H, Figure H.1)
- the nouns to be used in the illustration activity (what can be seen in each picture) (gPictureItems - see appendix H, Figure H.1) and a script identifying which nouns can be seen on each page (fWhatCanISee - see appendix H, Figure H.2)
- a script that returns a list of all words contained in the text and their associated values, used to identify which properties each word has (fGetWordProperties - appendix H, Figure H.3)
- a script that returns a list of pages on which the sentence shuffle activity would not make sense (fPageSentenceOK - appendix H, Figure H.4)

Examples of the scripts from the book entitled "A Book For Jack" are given in appendix H.

6.4 Formative evaluation and outcomes

6.4.1 Initial evaluations

Early versions of the initial prototype were demonstrated at an annual conference held by the Open University Computer Assisted Learning Research Group and also at the Open University Open Day. This enabled a number of colleagues to comment on the design and user interface.

In the early versions, the hint button was represented by a question mark in order to enable children to opt to receive a clue rather than the whole word immediately. However, it was suggested that it may confusing for the users because many software systems use the question mark to represent the general help facility. Several suggestions for improving this icon were made, including representing a teacher. Two pictures of a character from the selected reading scheme, who was a teacher, were identified. As the mechanism was designed to be toggled on or off, two states were required. The first depicted a teacher

looking on, but not intervening (no hint) and the second showed the teacher talking, with a speech bubble containing the word 'Hint!' (see page 119, Figure 6.6).

It was suggested that some of the activities could be further improved by incorporating more feedback. Firstly, in the sentence shuffle activity, two general hints were incorporated to be played if the correct sentence is not reconstructed:

Read the sentence you have made. Does it sound right?

Check that the sentence starts with a capital letter and ends with a full stop.

These were selected to complement human tutoring strategies. Secondly, both the sentence shuffle activity and the illustration activity included the use of separate words, which would be out of context and may be difficult to decode. The facility to obtain speech feedback was therefore also incorporated. If a child 'double clicks' (two mouse clicks made in rapid succession) on a word card then its pronunciation is given instantly. It was acknowledged that this might be difficult for a young child to do, but was considered to be the simplest and most intuitive way of implementing the feature. It was not possible to use a single click to obtain a word pronunciation because the activities use a drag and drop mechanism. It was considered that obtaining the word pronunciation by other means, such as using a pull-down menu, would be too complicated for the target age range.

6.4.2 Practitioner evaluation

Two of the teachers who had agreed to participate in the main study (chapter 7), volunteered to view the software and give some feedback on its suitability for classroom use. They were generally happy with the overall design, but they did offer some constructive advice. In particular, several suggestions were made to improve the hint mechanism in relation to human tutoring strategies. This included rephrasing the hints in a way that emulated the approach these teachers would take in their classroom. For example, a more general hint was suggested:

Have a guess, and then read the sentence again to see if it sounds right.

One of the hints suggests that the child should look at the initial letter. However, it could be confusing if the word for which they are being given a hint starts with a digraph ('th', 'ch', or 'sh'). For example, the sound associated with the digraph 'th' is different from that

associated with the first letter 't' if it is considered on its own. Instead, the child should be asked to look at the initial sound, rather than initial letter, and the digraph should be highlighted if appropriate. This could be extended to include all onsets in future implementations.

They were happy with the activities, but suggested that in the illustration activity the children were presented with fewer words. Also, a common problem experienced by early readers in all aspects of literacy development is understanding the concept of a space between words. This could be emphasised more clearly within the sentence shuffle activity, particularly if the child had to position a 'space' between each word. It was not considered to be feasible to implement this feature within the time constraints, but should be considered in future investigations.

The outcomes of the practitioner evaluation were that the hint mechanism and the illustration activity were adjusted.

6.4.3 Classroom evaluation

One of the schools that had agreed to participate in the main study (chapter 7), also agreed to allow a small usability study to take place. The study took place in the naturalistic setting of the classroom. Five children, from a Year 1 class, were selected by their teacher as being ready to read Stage 3 of the Ginn All Aboard Reading Scheme. Each child was shown how to use the software including turning pages, accessing pronunciations and what each button on the screen did. They were instructed to attempt to read each page, then select the sentence button to hear the text being read aloud and finally, to try an activity before continuing with the story.

The children were unfamiliar with using a mouse for controlling a program. The technology already in place in their classroom was relatively old, and the available input devices were a normal keyboard and a concept keyboard. However, it did not take any of them very long to master the use of the mouse and it appeared not to concern them. Therefore, it was decided that it would not be an issue in a long term study as the children mastered the skills required within one session.

It was noted that children of this age found the action of 'double clicking' very difficult, which was probably due to motor co-ordination skills. 'Double clicking' was implemented within two of the activities, to pronounce individual words. It was decided that the feature would remain, but that any children observed experiencing difficulties with this aspect of the user interface would be offered assistance.

The background noise caused considerable problems for all of the children, even though the computer was located at the back of the classroom. It was clear that the use of headphones would be necessary in the main study. This would have the added benefit of minimising any disruption to the rest of the class.

One child clicked twice on the arrow key by accident and turned two pages, unaware that she had missed part of the story. It was noted that this behaviour could also occur with a paper based book. It would be possible to detect this and alert a child to the problem, but it was not considered to be of high priority and was not implemented.

The last issue arising was that it took the children between 20 and 45 minutes to complete the book as the activities were time consuming, particularly reconstructing a sentence from disparate words. A requirement of the main study was to control the amount of time spent each day using the software. It was decided that the children should be instructed to read the entire book first, before attempting any activities. This would also ensure that the comprehension and flow of the story were not disrupted.

Overall, the children enjoyed using the software and had no problems with the basic operation of CARIS. The interface appeared to be easy to learn and use. The children enjoyed the activities and liked the use of sound, and commented that it was more fun than reading the paper book.

The outcomes of this evaluation were to:

- be aware that 'double clicking' is problematic for this age range and offer assistance with this feature in the main study
- ensure that headphones are available in future
- be aware that turning two pages at once by accident could occur

- adjust the instructions given to ensure that the book is read once before any activities are attempted

6.4.4 Undetected bugs

Two problems remained undetected until their existence was revealed through extensive use of the software in the evaluation (chapter 7). Firstly, due to the way in which the software was generated (a rapid prototyping approach) not all the books were included in the formative evaluation. This meant that a number of problems relating to the illustration activity occurred. This was because the objects that could be seen in each picture had to be specified individually for each book and thus errors could easily be made and were difficult to detect. Secondly, a bug in the Diagnostic Module prevented the software from being truly adaptive as originally intended. The Diagnostic Module was selecting a cue at random instead of identifying one that was represented as being under-utilised by the modelling mechanism. Thus, the hints or activities that were selected did not provide coaching in the cues that were under-utilised for each individual. Rather, all the children received coaching in all the cues.

6.5 Interaction examples

6.5.1 Starting CARIS

On activating CARIS, an initial credit screen is displayed (Figure 6.2).

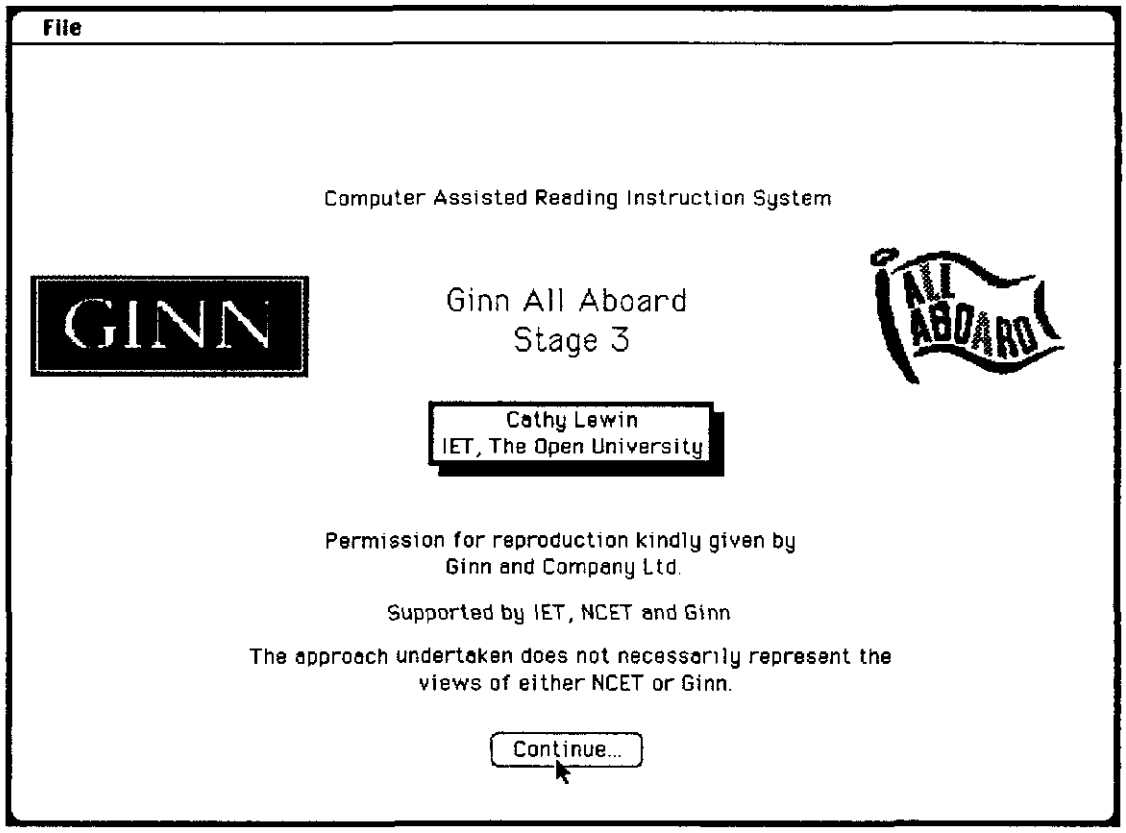


Figure 6.2: Initial credit screen on starting CARIS

Clicking on the 'Continue' button takes the user to the login screen (Figure 6.3), where they can select their name from a scrolling list box. The 'File' pop-down menu is available at all times, from which 'Quit' can always be selected to exit CARIS. The teacher's maintenance module can be accessed via the 'File' pop-down menu on this screen. Double clicking on a name, or selecting it and clicking on the 'OK' button, takes the user to a similar screen (not shown) allowing them to select a book to read from a scrolling list.

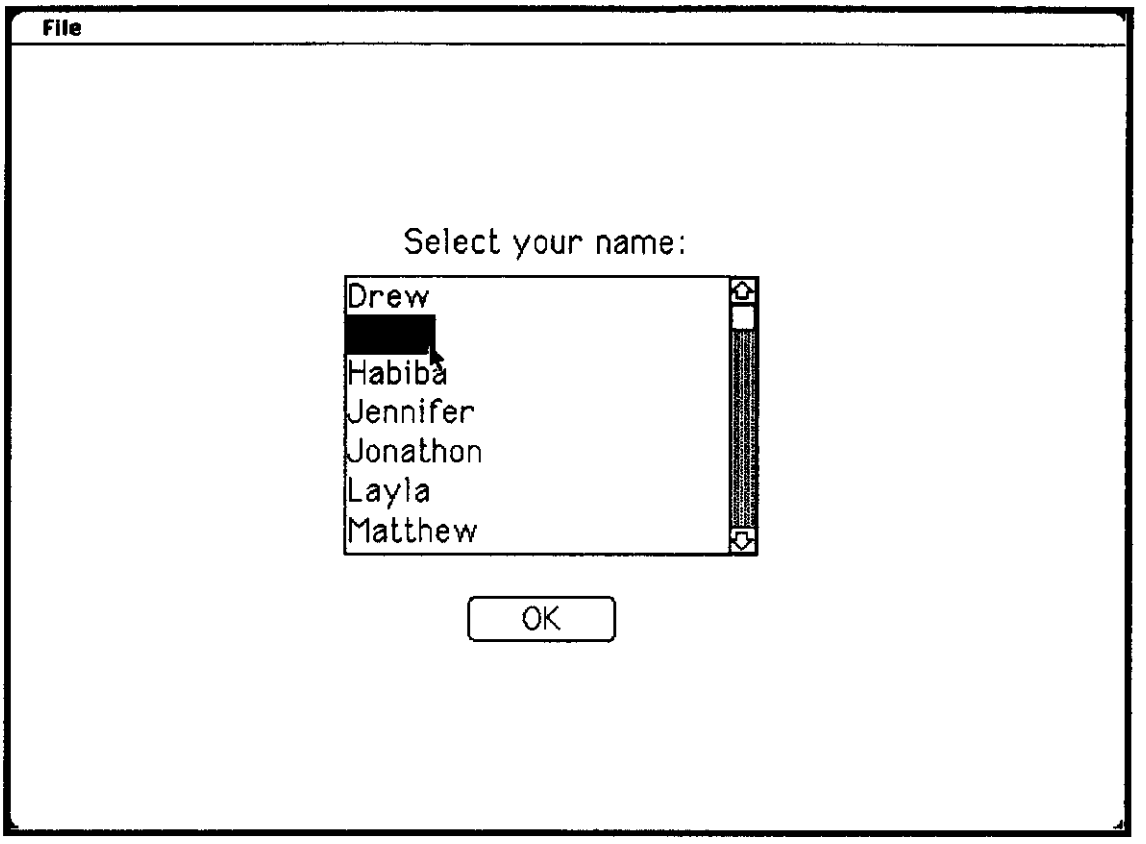


Figure 6.3: Login screen for selecting a name from the list of users

Selecting a book, in the same manner as selecting a user name, results in its title page being displayed (Figure 6.4). Any word can be selected independently on this page. Clicking on the speaker button will read all the text on the screen: the title, by whom the story has been written, and by whom it has been illustrated. The left arrow button is not displayed on the title page because it is only possible to page forwards.

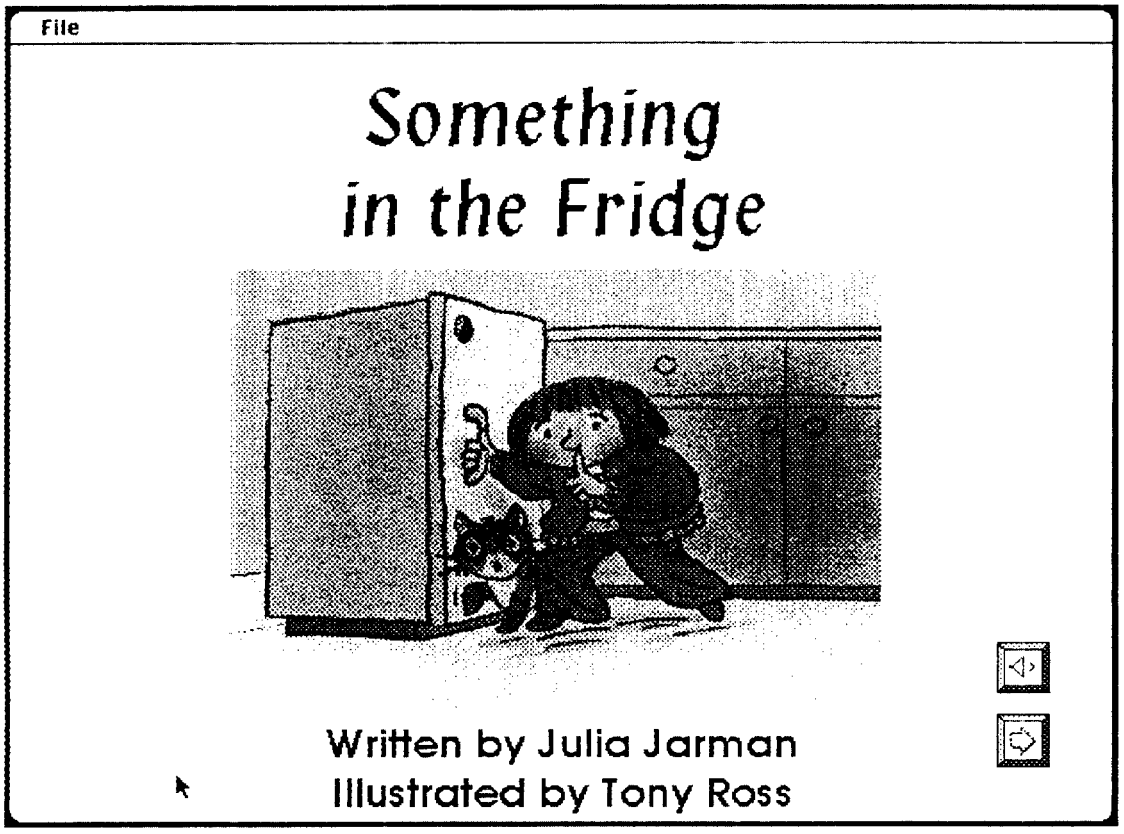


Figure 6.4: Title page of a book

6.5.2 Page and options

Figure 6.5 is an example of a standard page of a story. The user can click on an individual word, page forwards (right arrow button) or backwards (left arrow button), and invoke an activity depicted by a red pencil button. They can also hear the sentence read with expression, where each word is highlighted as it is spoken, by clicking on the speaker button. Clicking on an individual word causes the word to be highlighted and appropriate feedback to be given: whole word, onset and rime, or a hint. Clicking on the 'File' pop-down menu at this point would enable a child to exit the book by selecting 'Quit' and also toggle on and off the verbal instructions for activities. The 'Hint' button is toggled off initially.

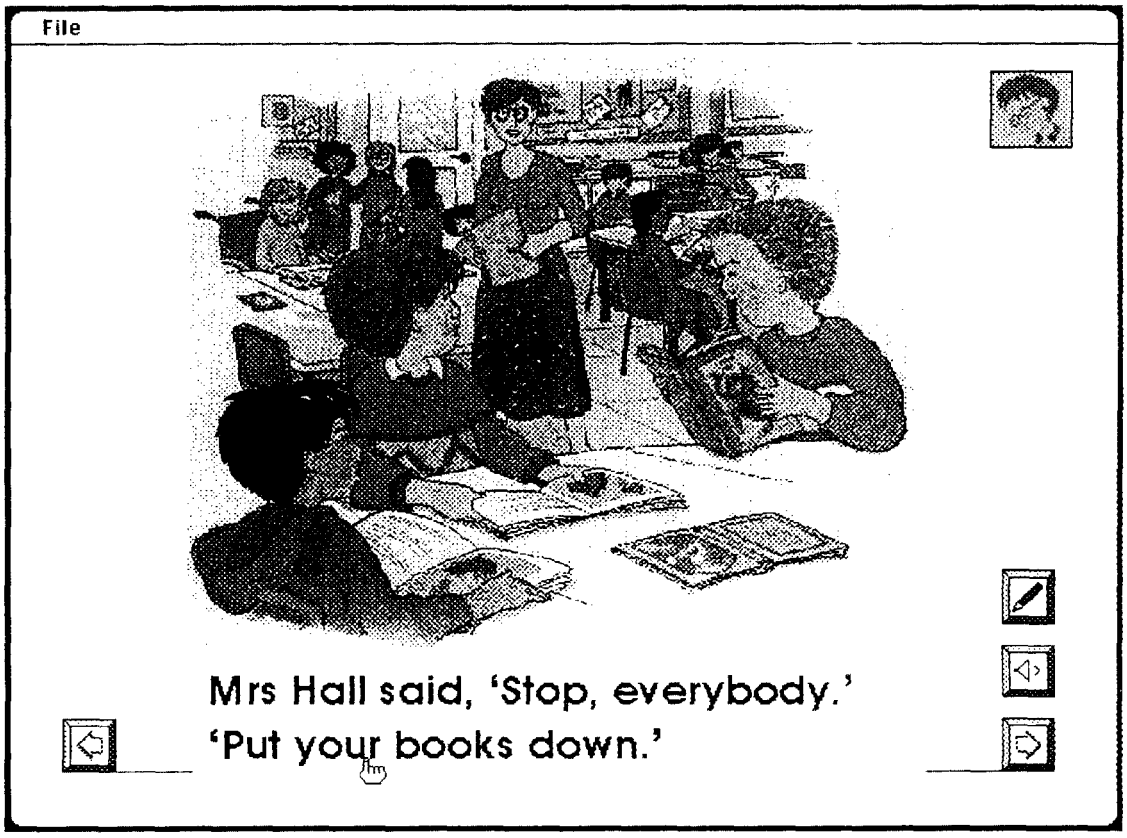


Figure 6.5: Example of standard page in a book

Figure 6.6 shows an example of what happens when a user clicks on the 'Hint' button. In this state (toggled on), a user would receive a hint when they click on an individual word. This would continue to happen until the user clicks on the 'Hint' button again, returning it to its initial state (toggled off).

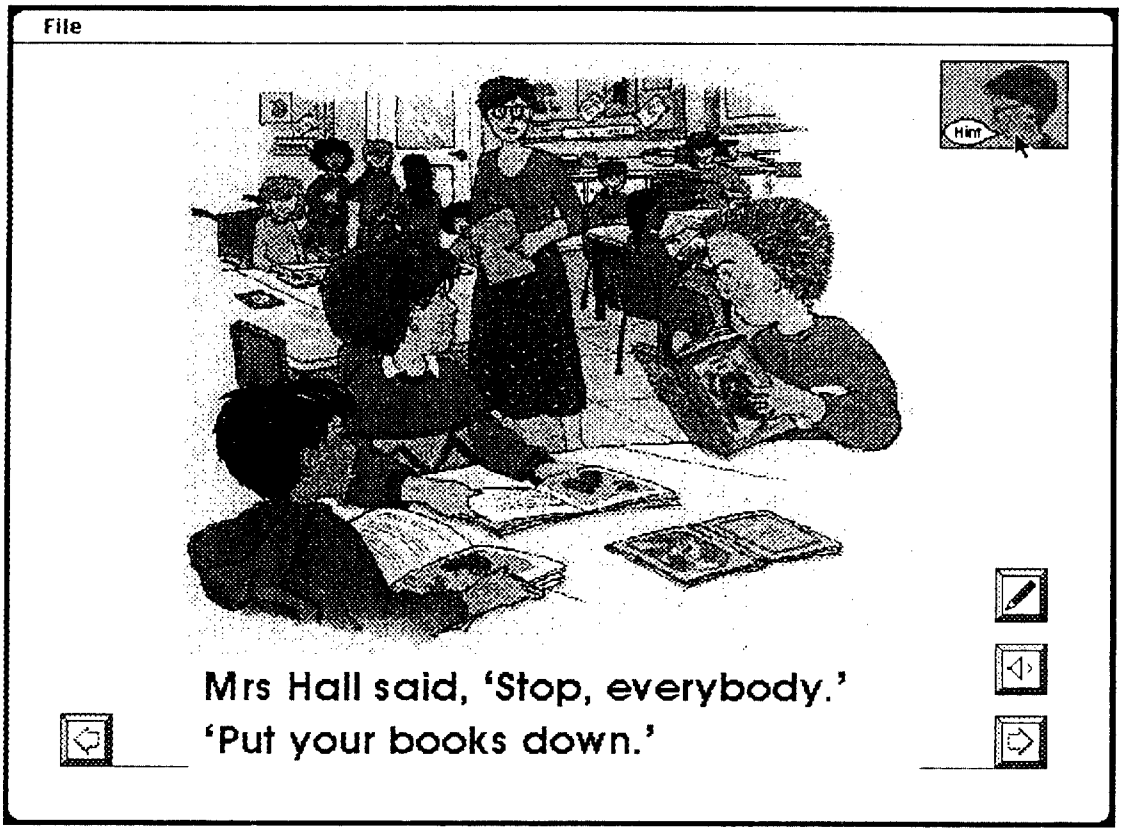


Figure 6.6: Example of 'hint button' toggled on

In the strand of the reading scheme that supports the development of phonological awareness, the Pattern and Rhyme books, the last page contains the complete rhyme. So every sentence from each page is repeated, which enables a child to appreciate the rhyme fully. To complement the hard copy books, this was replicated in the computerised texts (Figure 6.7). Clicking on the sentence button causes the complete rhyme to be read aloud with expression and each word is highlighted as it is spoken. As the illustration is general, and there are many sentences on the page, the activity button is not displayed. Clicking on an individual word is still supported in this type of screen.

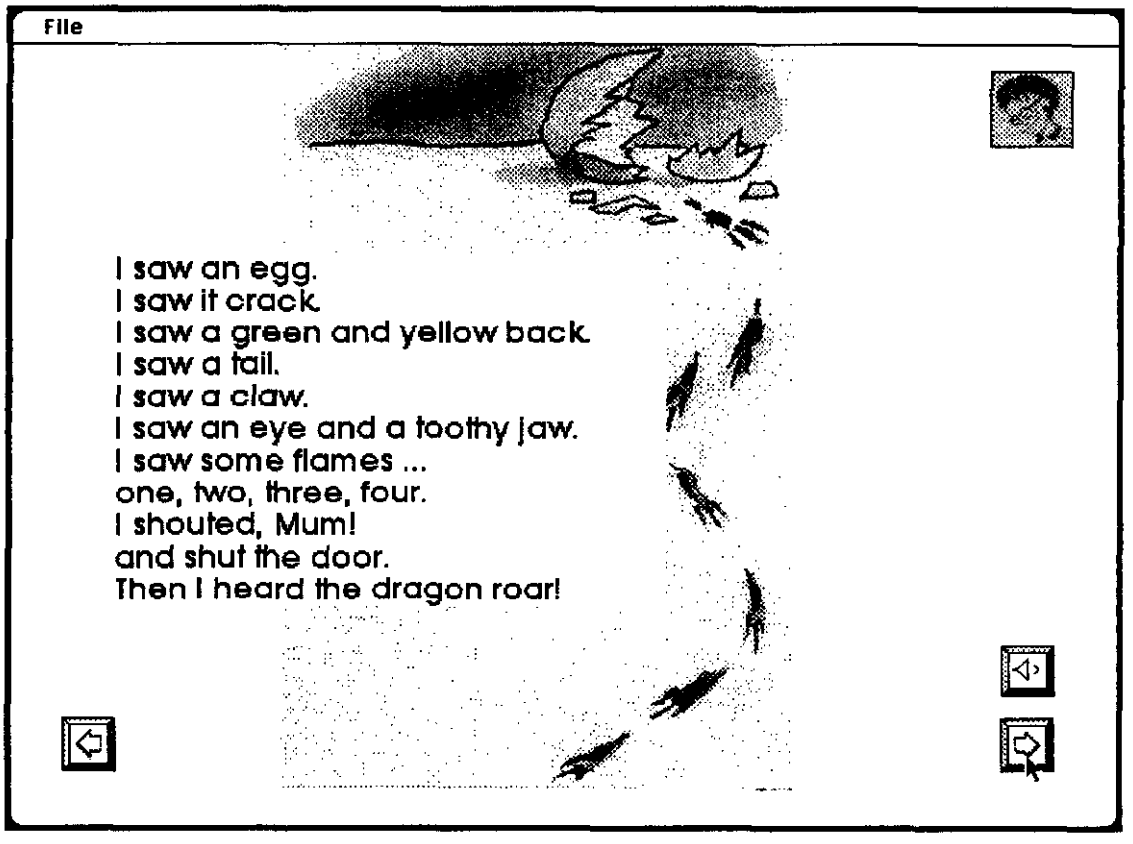


Figure 6.7: Example of last page of pattern and rhyme book, containing entire rhyme

6.5.3 Activities

All the activity screens contain a graphic in the top left hand corner such that it is clearly different from the standard page of the book. This is most important for the key word activity (Figure 6.8). In this task, the sentence and illustration from the page on which the activity button was invoked are replicated in the same positions on the screen. Even so, it could still be confusing for a less observant child.

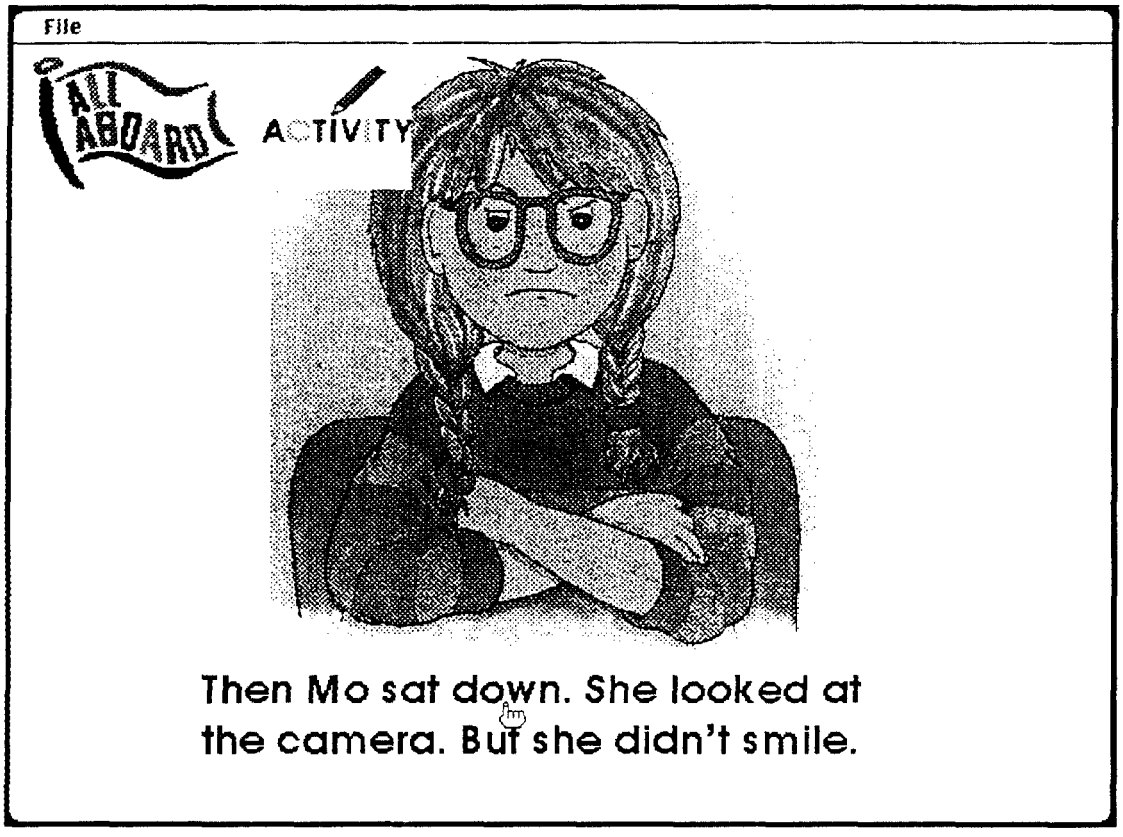


Figure 6.8: Example of key word activity, clicking on the word 'down'

In the key word activity, the child is given an instruction such as: "Read the sentence again and click on the word 'down'." Clicking on any word in the sentence causes it to be highlighted on the screen, and the pronunciation to be given. If the correct word is clicked upon, praise is given and the user is returned to the screen from which the activity was invoked. If an incorrect word is clicked upon, after the word is spoken, the child would hear: "Ahhh! Never mind. Try again!" Then the instruction would be repeated to minimise the possibility of failing again. It is assumed that the child will click on the correct word

eventually because there are a limited number of words in a sentence. Thus, the number of attempts is not otherwise limited.

Figure 6.9 shows an example of the illustration activity. In this case, the child is not correct and clicking on the 'OK' button would result in the child hearing "Ahhh! Never mind. Try again!" Following the third attempt, the child would hear: "Not quite. What you can see in the picture is ... Barney ... Babs ... sandwich ... bread." As each word is spoken, it would be simultaneously highlighted on the screen, drawing the child's attention to the correct answer. At any point during the activity, double clicking on a 'word card' would give the pronunciation of the word written on the card.

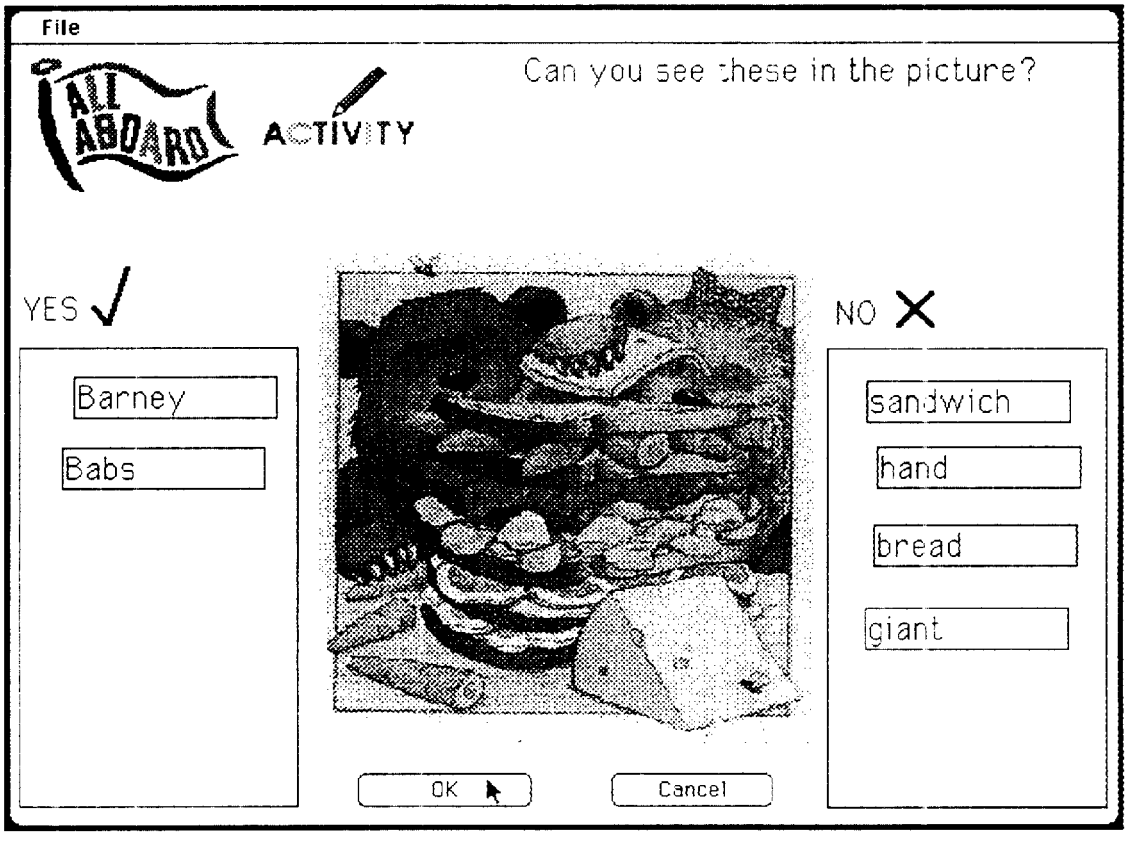


Figure 6.9: Example of illustration activity

Figure 6.10 is an example of the meaning and syntax activity, and depicts the word card 'had' being dragged across the screen to the sentence grid. When the activity has been completed and the child clicks upon the 'OK' button, the correctness of the sentence is verified. If it is not correct then a simple hint would be given as described in section 6.3. On the third attempt, the correct sentence would be displayed and read aloud. Each word card would be highlighted as its contents were pronounced, and then the child would be returned to the page they were on before they entered the activity. If the child correctly reconstructs the sentence, praise is given and the user returns to the page from which the activity was invoked.

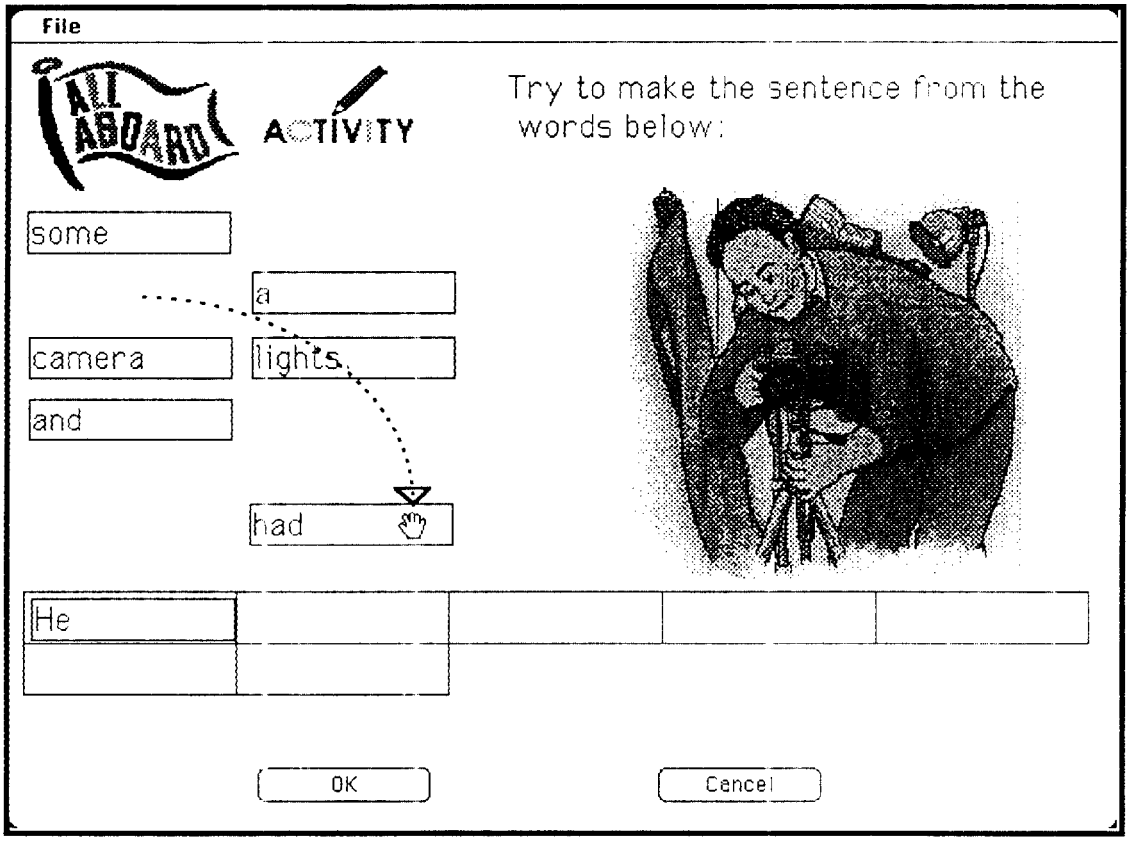


Figure 6.10: Example of meaning and syntax activity

The following three figures are examples of the onset activity for promoting awareness of blends. A child would be asked to guess what might be behind the blue door: "What is behind the door? Choose a sound that makes a word ending with 'aw'" (Figure 6.11). The user is required to click on one of the three blends, each of which is represented as a button at the bottom of the screen. If the child makes an incorrect choice, they would hear: "Try again". As the choices are limited, it is assumed that the child will eventually make the correct selection.

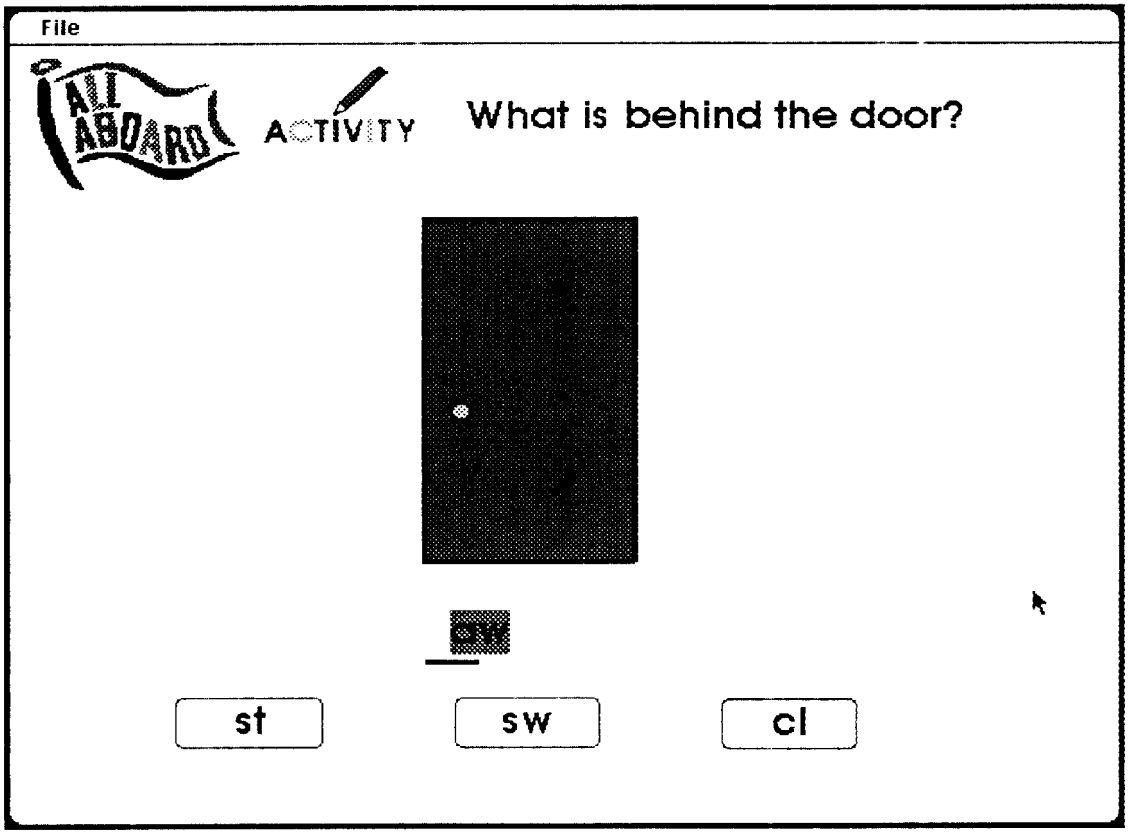


Figure 6.11: Example of onset activity, stage 1

In this example, the child has correctly clicked on the blend 'cl' (Figure 6.12). In this case, the whole word is displayed below the door, then the onset, rime and whole word are spoken. A simple animation has been implemented to reward success. The blue door opens to reveal a picture of the object representing the word that has been made (Figure 6.13). Thus, a restriction of this activity is that the only words that can be created are nouns and contained within at least one of the computerised texts. The images revealed behind the door are taken from illustrations contained in the talking stories, and therefore the activity is closely related to the books.

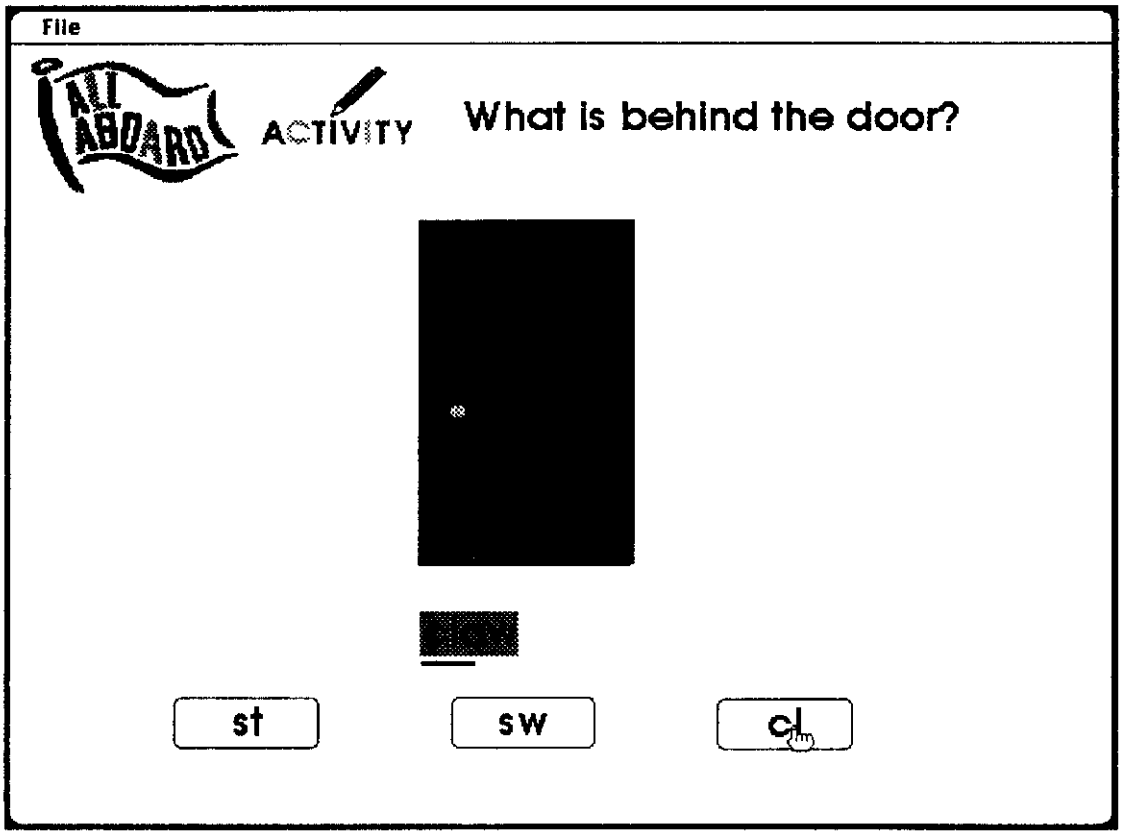


Figure 6.12: Example of onset activity, stage 2

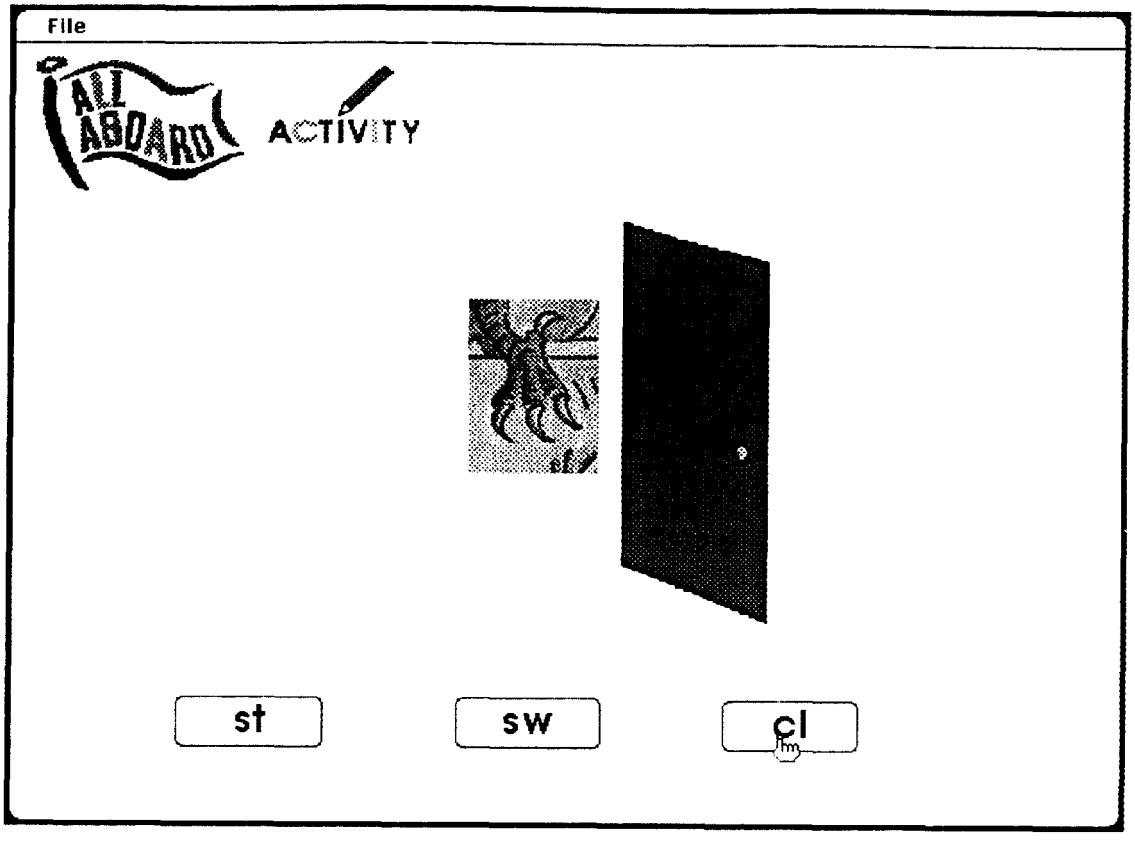


Figure 6.13: Example of onset activity, stage 3

6.5.4 Teachers maintenance module

The user interface example below (Figure 6.14) is the main screen seen on entry to the teacher's maintenance module. This allows various options to be set for an individual child, including which strategies are used regularly, their current confidence level, which books should appear on their personal book list, and feedback requirements.

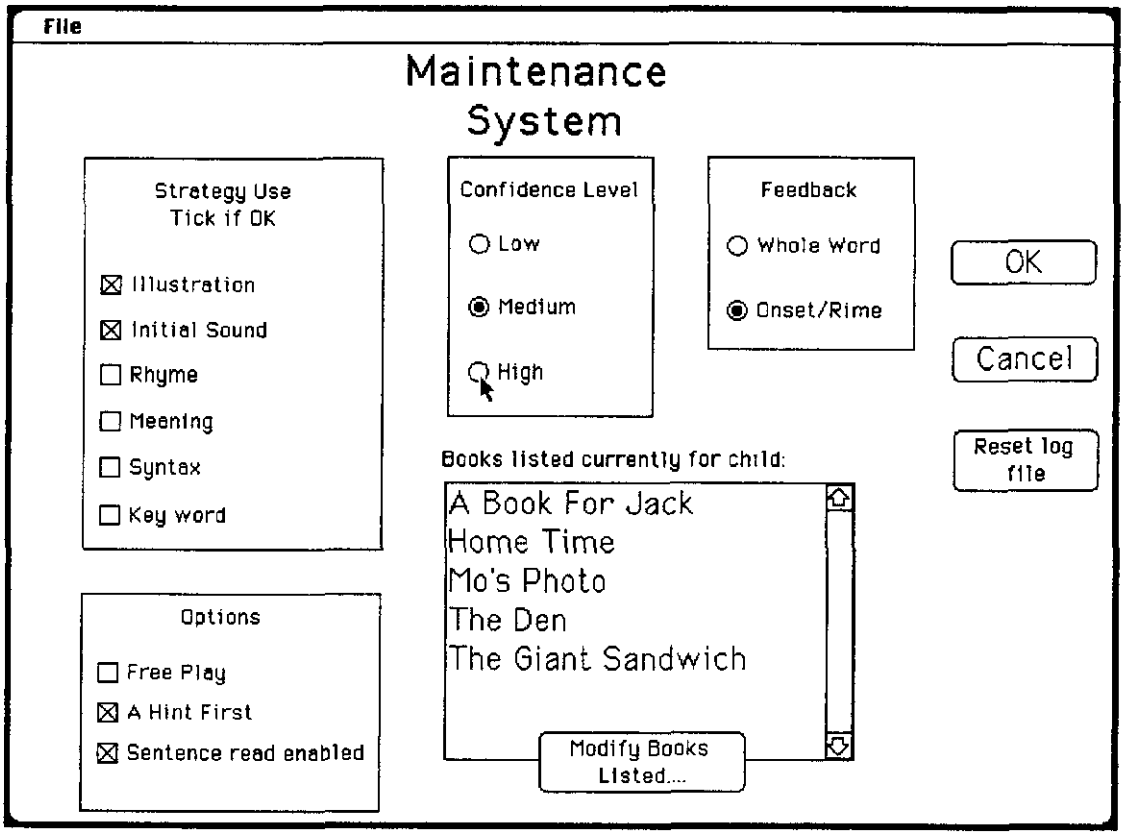


Figure 6.14: Example of main screen in teacher's maintenance module

6.6 Conclusion

This chapter has described how the design presented in chapter 5 has been implemented, including modifications made both due to time restrictions and as a result of formative evaluations. A number of examples of possible interactions with the system have been illustrated. Chapter 7 describes an empirical study, in which the computerised texts that have been developed were used, and presents a formal evaluation of this implementation.

7. Evaluating the effectiveness of the adaptive talking books

7.1 Introduction

This chapter describes an evaluation of the computerised texts created as a result of implementing the design (chapter 5 and chapter 6). This study was designed to contribute to existing knowledge on the use of electronic texts for providing supplementary reading practice. The study was conducted in the naturalistic environment of the primary classroom. Additional software features were introduced such as a greater variety of feedback modes, and a greater emphasis has been placed on the affective issues than in previous studies in this field (see Chapter 2, section 2.4).

The evaluation took place with children (referred to by pseudonyms) who were identified by their teachers as being ready to read the texts that had been implemented as electronic books. The study had the following aims:

- to explore the relationship between the software and the classroom environment, and ascertain how it can be integrated successfully
- to ascertain the effectiveness of a number of additional features designed to address the weaknesses in existing software and in particular to determine whether:
    - providing coaching in the use of specific reading cues would positively affect independent decoding of unknown words
    - providing an activity to reinforce sight recognition of the key words in the reading scheme would positively affect the key word vocabulary
    - giving a hint, rather than the whole word immediately, would encourage independent word identification and prevent a reliance on instant word pronunciations
    - providing feedback at the level of onset and rime would positively affect the development of phonological skills.
- to explore the effects of such a resource on levels of self-confidence and motivation
- to verify the technical adequacy of the software

The framework developed and piloted to ensure a holistic approach to evaluation (chapter 3 and chapter 4) was used, focusing on the pedagogical, cognitive, affective and technical aspects of the software. The context in which the software is used was considered together with the content.

Two versions of the computerised texts were created. The first version was an implementation of the design presented in chapter 5 in which the strengths of existing commercial software were incorporated and the weaknesses addressed. It was designed to be adaptive in order to meet users' individual needs, although this was not fully implemented due to a bug that remained undetected (chapter 6, section 6.4.4). A variety of feedback modes, such as hints to encourage independent decoding and feedback at the level of onset and rime, were incorporated. Another design feature was the provision of coaching in underutilised reading cues: initial letter, meaning, syntax, word analogies, and illustrations. This implementation is referred to in this research as the enhanced software. Secondly, basic versions of all texts were created. In these versions, feedback was available for whole words and sentences only and no additional features such as activities and hints were incorporated. The effectiveness of the enhanced software was investigated by comparing its use with that of the basic version of the talking books.

The remainder of this chapter is divided into five main sections in addition to the introduction: methodology used (section 7.2), results (section 7.3), case study data (section 7.3.4), discussion (section 7.5) and conclusion (section 7.6).

7.2 Methodology

Section 7.2 describes the method undertaken. The research methodologies employed in this project were described more fully in chapter 3 (section 3.4) and so the general design is only briefly summarised here. The participants are described and the selection procedure that was used is explained. One of the aims of the study is to explore the way in which the software can be integrated within a classroom and how it might complement current teaching approaches. Therefore, the classroom practice of the three schools is also summarised. The section concludes with a description of the procedure that was followed, including the pretests and posttests conducted, and how each group was instructed to use the software.

The use of enhanced and basic Talking Books was evaluated in the naturalistic environment of the classroom. The participants were instructed to read the Talking Books in a way that complemented teaching practices in the participating schools. They used the software on a

daily basis for a period of four weeks because regular practice is a primary factor affecting learning to read (e.g. Harrison, 1992).

A matched pairs design was employed based on the pretest results. In addition, an attempt to control for the teacher effect was made. As well as being matched on test scores, the pairs were matched as closely as possible for teacher. The pairs were then randomly assigned to two groups. One group used the enhanced software and the other used the basic software. The evaluation was conducted using both quantitative and qualitative research methodologies, combining word recognition tests with questionnaire, interview and observational data (see chapter 3 section 3.4).

7.2.1 Participants

Three schools using an integrated approach to reading instruction were identified. School A was a small combined primary school (age range 5 - 12 years) serving a wide rural area of mixed socio-economic-status (SES) in South Yorkshire. It had no nursery provision and 60 children in three mixed age range classes. School B was a combined Church of England primary school in a former coal-mining town in South Yorkshire with high unemployment. It had nursery provision and 200 children in total. On entrance to school, a large number of the pupils had minimal exposure to print and literacy skills, as judged by the early years teaching staff. School C was a multi-racial first school (age range 4 - 8 years) serving a mixed SES catchment area in a large market town in a central county. It had nursery provision and 320 children, 40% of who were from the local Asian community. It had extra staff support for children with English as a second language.

The teachers involved selected 32 participants from five Year 1 (aged 5 to 6 years) classes from these schools (Table 7.1). Their ages ranged from 5 years 1 month to 6 years 1 month at the beginning of the study. The mean age of all participants was 5 years 9 months.

Table 7.1: Breakdown of pupils from schools

School	Class	Pupils	Girls	Boys
A	1	4	2	2
B	2	5	4	1
B	3	7	4	3
C	4	8	4	4
C	5	8	5	3

It was requested that all children selected were considered to be ready to read, or would benefit from reading, books from Stage 3 of the Ginn All Aboard reading scheme. The children identified to participate in the study had a variety of reading abilities and teachers exercised a degree of flexibility in the selection process. Some of the children selected were still completing Stage 2 of the reading scheme. There is evidence to suggest that pupils who have not progressed as far as their peer group in reading achievement can make the greatest gains in word recognition measures (e.g. Reitsma, 1988). Therefore, it was assumed that these participants would still benefit, but that they would require additional support and may be more dependent than the other children on speech feedback. At the other end of the reading ability range, one child had already completed all the texts in Stage 3 but was considered by his teacher to be lacking in confidence. The teacher felt that this child would benefit from reading the texts again in an alternative format.

Pairs of children were then matched as described (section 7.2). Sixteen children were assigned to the enhanced software group and 16 children were assigned to the basic software group.

7.2.2 Classroom practice

The schools involved in the study were already using the reading scheme upon which the Talking Books were based. This was supplemented further by additional books of equivalent difficulty levels, both from other reading schemes and real books. The teachers

involved in the study used similar teaching strategies. They shared stories with the children daily and related activities included those designed to heighten awareness of the initial sounds of words. They requested that texts were read again if accuracy and comprehension were judged to be unacceptable. They also encouraged the children to use a variety of reading cues for decoding rather than telling them unknown words instantly. It should be noted that it was the policy of the year 1 teachers in school B to listen to the children read on an individual basis every day. In comparison, the teachers in schools A and C listened to the children two or three times a week. Nevertheless, it was considered that the style and approach of software complemented the classroom pedagogies in all schools.

7.2.3 Procedure

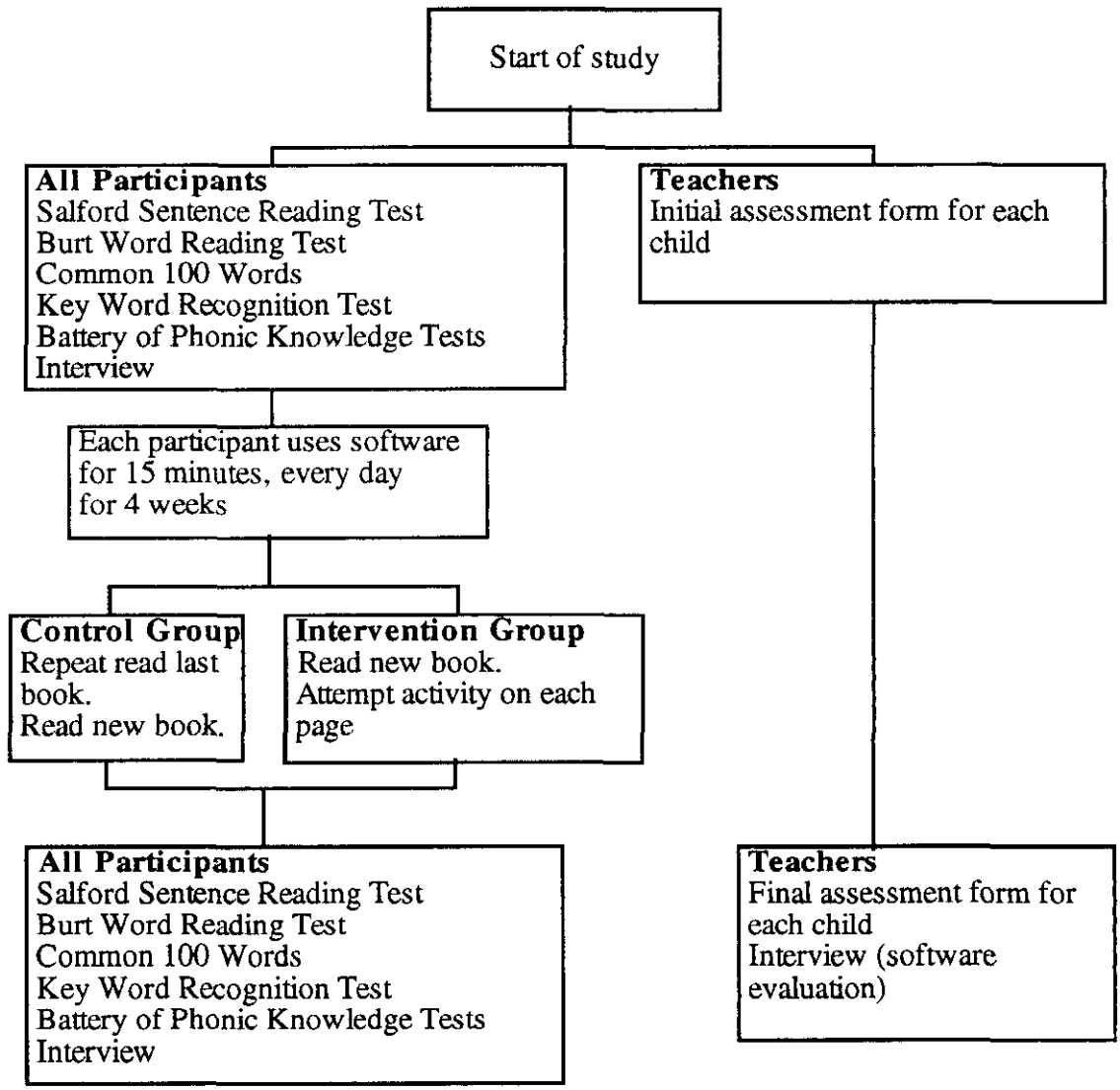


Figure 7.1: Schema for empirical study

At pretest and posttest the participants each undertook a battery of assessments and were also interviewed (Figure 7.1). The assessments included two standardised reading tests: the Salford Sentence Reading Test (Bookbinder, 1976) and the Burt Word Reading Test (1974 revision) (Scottish Council for Research in Education, 1976). The sentence based test was considered to be the primary measure as it concerned word recognition in the context of meaningful text rather than in isolation. The children were also assessed on their sight recognition of the key words specified in the reading scheme for stages 1, 2, and 3 (appendix I). In addition, they were given criterion referenced tests based upon the type generally contained in a Local Education Authority reading assessment pack (appendix C). These tests were used to ascertain the extent of their sight vocabulary of 100 commonly occurring words and their current level of phonic knowledge. The interview at pretest consisted of a number of pre-set questions covering aspects such as their attitudes to books and reading, and computer use at home and school (appendix B.1). At posttest the interview focused on their attitudes to reading and an evaluation of the software they had used (appendix B.2).

Teachers were given informal reading assessment questionnaires for each child that included aspects such as strategy use, motivation and self-confidence, general reading ability and specific problems encountered in reading (appendix J). They were also asked to indicate whether or not they considered that a child would benefit both from receiving speech feedback at the level of onset and rime, and receiving a hint first. This enabled the software to be configured to meet an individual learner's needs. In addition, the posttest questionnaire included a brief evaluation of the Talking Books used.

Each class either already had a Macintosh computer powerful enough to run the software, or were lent a Macintosh Performa 5200. Headphones were used to minimise distraction to other members of the class by screening out the sounds generated in the computer software.

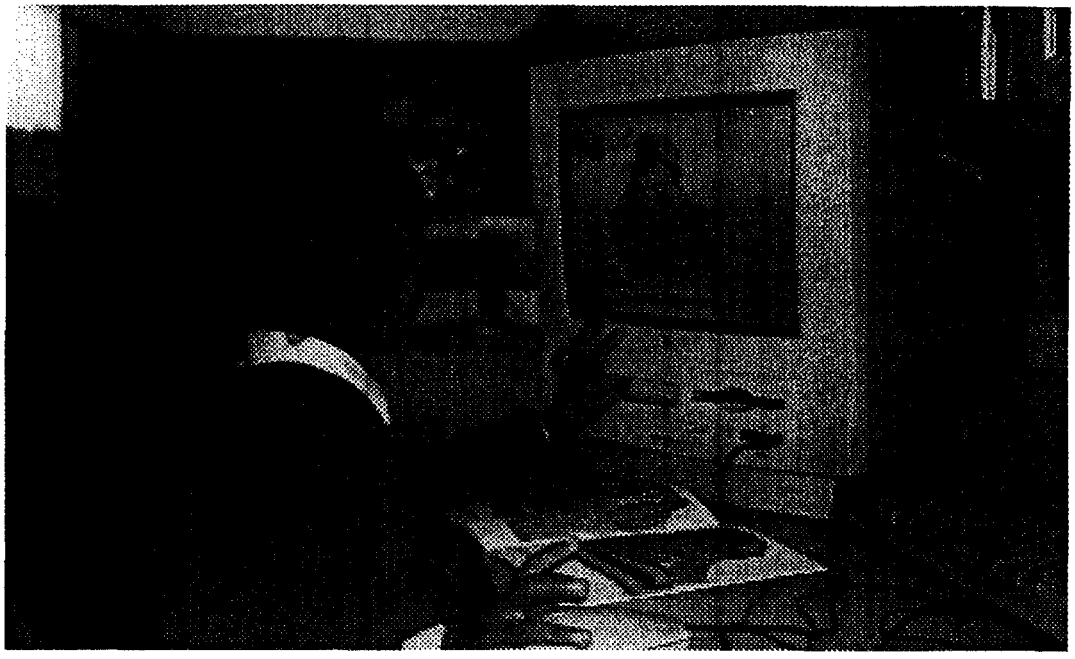


Figure 7.2: One participant interacting with a talking book

The children used the software individually (Figure 7.2), on a daily basis for 15 minutes, over a period of four weeks, during which time they read a total of 20 computerised books (appendix F). The classroom teacher at school A was not present for the first three weeks of the study due to ill-health. During this time, a variety of staff members from the head teacher to the school secretary supervised the children on different days. The head teacher did not feel that it was appropriate to ask the supply teachers to become involved in the study. In school B the children were supervised by their teachers, each of whom were given clear instructions on how to use the software and which books each participant should read every day. In school C, where a separate classroom was used for logistical reasons, a researcher supervised the children's use of the software. As, in this school, it was common practice for children to work individually and in small groups in rooms other than their own classroom, it was still considered to be a naturalistic setting.

Both those children using the enhanced version of the software and those using the basic version were asked to attempt to read the text independently, selecting words for speech feedback only as necessary. They were also instructed to select sentence level feedback after attempting to read each page themselves in order to self-check their reading accuracy. The group using the enhanced version of the software were asked to read the complete book

before attempting any activities to ensure that flow and comprehension were not disrupted. The length of time for which the software could be used was limited to 15 minutes by setting a kitchen timer before starting the session. The group using the basic version were instructed to repeat read the last text encountered, and then read a previously unseen book. This was to ensure that both groups spent about the same amount of time using the software in order to control for the practice effect. Reading a book more than once is a common teaching practice and has been shown to be beneficial for early readers (Carver, 1990). If a child was absent then alternative sessions were arranged within the intervention period such that each child used the software on a total of 20 occasions. All children were also instructed to read aloud to facilitate further data collection. Hard copy versions of all computerised texts were available for the children if required.

The first session was used to demonstrate how the software was operated, together with instruction on how the resource should be used. Each child was shown how to start up the system and the general layout of the computerised texts. Then they were asked to read the text and offered technical assistance as required. Some children were less familiar with technology and additional support was given if requested in subsequent sessions.

Each child was observed using the software for an entire session on a weekly basis. Miscue analysis was conducted by taking a Running Record (Clay, 1993b) (chapter 3, section 3.4.3). Miscue analysis records the behaviour of a child reading aloud. Unexpected responses that differ from the text being read and the use of strategies, such as repeating words in order to bolster comprehension, are noted. It was also noted when the speech feedback was requested instantly and whether or not it was used for confirming a guess. The software also logged details of all interactions such as word selections, which activities were undertaken and activity results.

7.3 Results

The results are now presented in relation to the framework for evaluation (chapter 3, section 3.3), which classifies the evaluation domains as pedagogical, cognitive, affective and technical. The findings for each category are described and then summarised.

7.3.1 Pedagogical issues

Data to support this section of the framework for evaluation was collected through teacher interviews and questionnaires, and also from classroom observations. The teacher in school A felt unable to comment in depth as she was only present for the last week of the study (7.2.3). The two class teachers from school B had more exposure to the software because the machines were located in their classrooms and they were involved in supervising the participants. The two class teachers in school C were aware of the functionality of the software but were not involved in supervising the children as the machines were located in a separate classroom.

Teaching approach

The framework for evaluation specifies that the software should be judged on whether it complemented classroom practice. This was investigated through observations and interviews with the teachers. The results of the interviews suggest that the electronic books did indeed complement their teaching approaches, which were similar, as all five teachers commented that the general format was successful in their classrooms. These teachers were already using the reading scheme upon which the software was based (see section 7.2.2). Thus, the computer-based resource complemented the other classroom materials. Furthermore, the software had been designed specifically to support an integrated approach to reading instruction and also to complement the aims of the Ginn All Aboard reading scheme (Fawcett, 1994).

Teacher 2 commented that some activities were more difficult than others, referring specifically to the onset and blend activities (chapter 6, section 6.5.3). She felt that it would be more beneficial to introduce this later in the software or when the children had reached a certain level of competence.

Teacher 2 strongly felt that giving hints to encourage independent decoding (chapter 6, section 6.5.2) made the software more compatible with her classroom practice. She said:

> The hints though are a good idea as I would not want them to be on a programme where they were not encouraged to 'have a go' and use cues and strategies. If they were given the word immediately it would go against what is being done in the classroom ... encouraging independence.

However, she did point out that sometimes the hints were not appropriate and that they should be more specific to the word and its context:

> If child is stuck on 'the' it is not helpful to be told 'Look at the first letter.' This can be confusing or unhelpful.

All five of the teachers indicated that they would integrate the computer-based resource into their normal classroom practice if they had software available. For example, teacher 4 said she would use the software throughout the day in a variety of ways, including individual use, mixed reading ability pairs and group reading sessions.

Teacher 2 chose to use the software with lower reading ability children who were not participating in the study. These children were instructed to listen to the complete story being read aloud first, then to try and read it for themselves. She also used the software as a focus for a group activity. She found that it generated "very good interaction between children". She further commented that the activities had been especially useful in this context. For example, she had anticipated that the sentence shuffling activity (chapter 6, section 6.5.3) would be too challenging for the poorer readers. However, it stimulated much discussion amongst all classroom members about initial sounds, which she considered to be constructive and beneficial.

Teacher 2 felt that the Talking Books had stimulated a greater interest in reading overall. She described how the hard copy versions of the books kept 'disappearing' only to be found in the hands of the last child who had used the software. She recounted a day when Andrew had picked up the book after his session was over, leaned against a cupboard and read the story again from cover to cover, greatly enjoying it.

Classroom management

One aspect of classroom integration relates to the management of the computer-based resource. During the posttest interviews, several comments were made relating to such issues. Teacher 1 commented that it had been problematic to time-table the children's computer sessions to fit in with other activities such as drama lessons. Furthermore, she felt

that the support that her children required was time-consuming, although it had given the participants a level of independence in their reading practice. Teacher 3 commented that, with only one computer in the classroom, access to other software for those children not participating in the study had been severely restricted. Teacher 2 also agreed with this:

> It was no problem to integrate into the classroom, only that it monopolised the computer for one session of the day.

Neither of them felt that it had been a distraction for the other children in the class, although there had been an initial interest in the software from those children not participating in the study.

All teachers, except for teacher 1, were asked what information they would like the software to record in order to monitor progress. The responses included:

- which books had been read and how many times
- time taken to complete a book
- the words which children needed most support with
- how often the sentence level feedback was selected
- percentage of activities correctly completed after one, two and three attempts

However, several concerns were raised. Teacher 3 recognised that it was difficult for the computer to record the mistakes that the children might make when reading the text. When it was pointed out that the current implementation could be modified to record the children reading aloud, she commented: "Then you've got to go through it". This teacher recognised that this kind of feature may require an additional investment of time to assess a child's accuracy in reading.

Teacher 2 felt that it would be too time-consuming to look at information generated by the software for all children. She stated, though, that it would be useful to refer to if a child did not appear to be making progress or if she had concerns that the software was not being used correctly.

The timetabling problems highlighted by teacher 3 were due to a lack of resources and not the software itself. Staff commented that it would be helpful to see records of children's progress, although they accepted that some aspects of use (mistakes made) could not be

recorded in a simple fashion. In order to complement classroom practice closely, talking books should provide detailed records of a child's progress and facilitate easy access to this assessment information.

Other pedagogical issues

The pedagogical area of the framework for evaluation also specifies that desirable attributes of software include supporting the holistic development of literacy skills, being suitable for the intended age range, and supporting student autonomy. Observations were used to evaluate how the software met these aims.

An integrated approach to reading instruction supports the holistic development of literacy skills, in that writing and reading skills are taught together. Thus, it is desirable to provide a similar level of support within software designed to support literacy teaching. An activity was designed to develop children's prediction skills. The children were asked to write (type) what they thought the story would be about before reading the book. At an appropriate point in the text, they were also asked to write what they thought would happen next. However, the prediction activity was disabled early on in the study since young children's typing skills are not efficient. The first users in the enhanced group spent a long time attempting to write their predictions. The length of time for which the software was used was restricted to 15 minutes. Completing the prediction activity restricted the effectiveness of the software use and limited the number of other activities that were attempted.

It can be stated that the software was appropriate for the children who used it because it was related to a stage of a reading scheme designed for a specific age range. It was observed that activities were challenging for the more able readers.

The software did support student autonomy. Although the children did not choose the book title each day, they were in full control of the software. Each child was able to read each page at his/her own pace and request as much or as little support from the computer as they needed. The software facilitated a level of independence from the teaching staff. Staff members commented that little additional adult support was required:

> It gives extra constructive reading help without needing a teacher there, and the children can use it independently. Very important in a busy classroom.

Some children found it useful to refer to the original text and illustrations in the hard copy versions that were provided. This was commonly observed when they were attempting the activities, particularly if they found the tasks to be challenging. Furthermore, the activity designed to develop the child's use of the illustration for decoding unknown words (chapter 6, section 6.5.3) demanded the use of the original texts. The picture that was replicated in the software in these cases was compacted and was not always clear. In the example given in

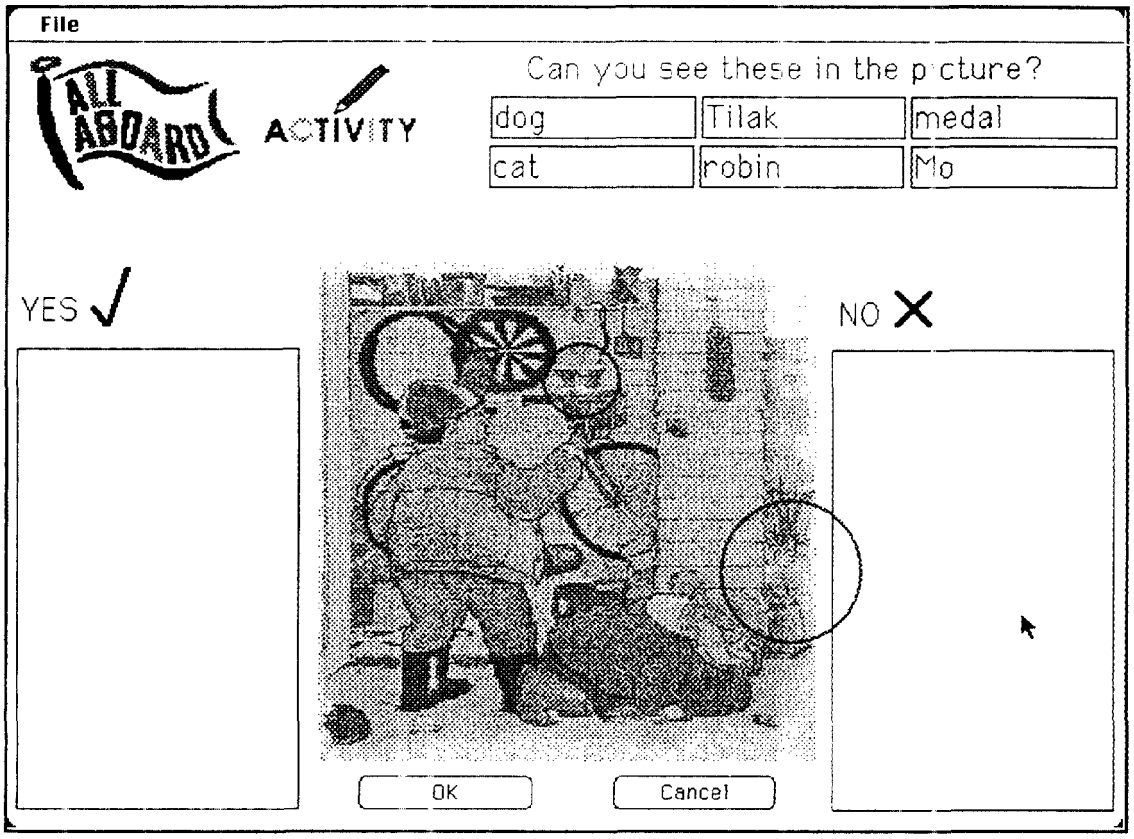


Figure 7.3, it is quite difficult to identify the presence of both the cat and the robin in the illustration (both circled). Yet this is required in order to complete the activity successfully.

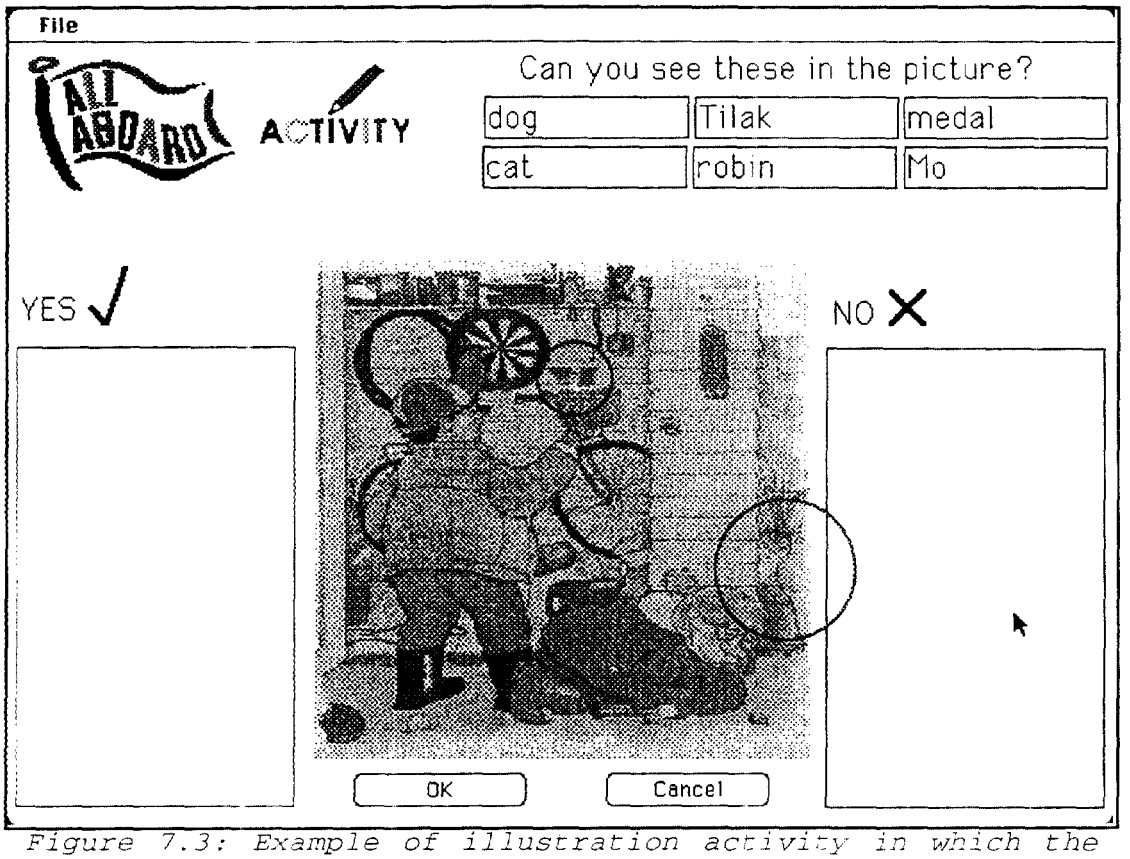


Figure 7.3: Example of illustration activity in which the picture is compacted and some objects are difficult to identify

Summary of results relating to pedagogical issues

- The teaching staff were extremely positive about the software and indicated that they would use it out of choice if it were available.
- Issues were raised concerning classroom management. In this study, because it was a requirement that the children have access to the computer daily, it had prevented access for other uses. This is a resourcing problem rather than one related to the software itself.
- The features in the enhanced talking books appear to complement classroom practice:
  - The provision of hints to encourage independent decoding and activities to develop the use of a variety of reading cues were seen to be beneficial. The hints should take into account the specific word and its context.
  - The activities were considered to be appropriate, although the prediction activity had to be disabled. It was perceived that the onset activity was more challenging than the other activities.
- Both forms of software facilitated independence and required little supervision.

7.3.2 Cognitive issues

The cognitive issues relate to learning outcomes measured by standardised reading tests and also to the use of reading strategies for independent decoding. The tests were administered

prior to use of the software and immediately after the intervention period. In addition to the standardised reading tests, sight recognition of the key words in the reading scheme and 100 common words were assessed. Further tests were used to assess phonic knowledge. Observation data from miscue analysis was used to assess each child's use of reading cues. In addition, the teachers were asked to identify whether each child regularly used each of five cues: initial letter, meaning, syntax, word analogy, and illustration. Due to the absence of teacher 1, four participants were excluded from the analysis of teachers' perceptions of reading cue use. Due to pupil absences, the posttests for three participants were conducted by their teachers a week later than the other children. The teachers were given clear instructions on the procedures to follow. However, the key word recognition test was not conducted in full for one of the participants in the enhanced group (Justin). His score on the posttest was partially estimated by using the pretest score, assuming that the words recognised at pretest would be successfully read at posttest. It may be that his score was a conservative estimate.

The overall results for the sight vocabulary and phonic tests are presented below. This is followed by the results by reading ability group. Then data relating specifically to the acquisition of key words is considered. Finally, data concerning the development of the use of specific reading cues (initial letter, word analogy, meaning, syntax, illustration) is described. The section concludes with a summary of the findings.

Overall results for sight vocabulary and phonic tests

Descriptive statistics for the outcome measures are given in Table 7.2. The score for the phonic tests (appendix C, section C.2) was calculated by totalling the scores for the 10 sub-tests. However, each sub-test was only administered if the child had scored at least 50% on the previous sub-test. Only the first two sub-tests were completed by all participants and have been analysed individually, in addition to the overall score. The first sub-test was designed to assess phonic decoding of consonant-vowel-consonant (CVC) words (for example, 'cat', 'sun', dig') (appendix C, section C.2, Test C.2.1). Twelve words were presented, of which four were pseudowords. The second sub-test was designed to assess blending skills. Each child was given three sounds and asked to blend them together to make

a word (for example, 'r - a - t', 'b - i - n', 't - o - p'). The sounds for nine words were given, of which three were pseudowords (appendix C, section C.2, test C.2.2).

Table 7.2: Descriptive statistics for all outcome measures

Outcome Measure	Software	Pretest Mean (SD)	Posttest Mean (SD)	Mean Gain (SD)
Salford Sentence (months)	Enhanced (n = 16)	77.75 (6.24)	80.81 (7.19)	3.06 (2.16)
	Basic (n = 16)	77.94 (6.69)	81.06 (8.52)	3.12 (2.28)
Burt Word (months)	Enhanced (n = 16)	75.36 (11.48)	81.48 (13.87)	6.12 (5.52)
	Basic (n = 16)	78.48 (13.89)	82.68 (13.28)	4.20 (4.20)
Common 100 Words	Enhanced (n = 16)	41.06 (33.95)	50.69 (31.16)	9.63 (9.11)
	Basic (n = 16)	45.56 (33.71)	50.88 (30.50)	5.31 (7.70)
Key Words (44)	Enhanced (n = 16)	23.13 (13.55)	30.19 (11.60)	7.06 (4.02)
	Basic (n = 16)	26.06 (12.00)	32.38 (9.15)	6.31 (4.59)
Phonic - CVC only (12)	Enhanced (n = 16)	5.00 (4.05)	5.81 (3.56)	0.81 (2.01)
	Basic (n = 16)	4.63 (3.58)	5.56 (4.02)	0.94 (3.43)
Phonic - Blend only (9)	Enhanced (n = 16)	4.31 (3.30)	6.06 (3.40)	1.75 (2.21)
	Basic (n = 16)	3.50 (3.25)	4.69 (4.01)	1.19 (2.01)
Phonic - all (75)	Enhanced (n = 16)	18.31 (18.28)	21.88 (18.49)	3.56 (5.37)
	Basic (n = 16)	13.56 (15.86)	18.88 (18.82)	5.31 (6.71)

The pretest and posttest mean gains are represented graphically for each outcome measure (Figure 7.5 to Figure 7.11). They are represented as E for the enhanced software group and B for the basic software group in all diagrams.

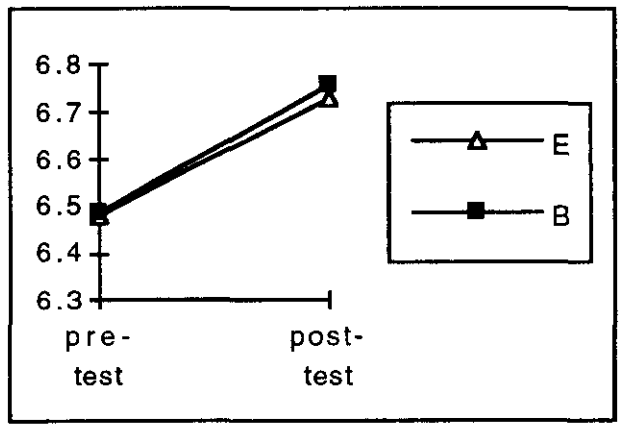


Figure 7.5: Pretest and posttest means for Salford Sentence Reading Test (years)

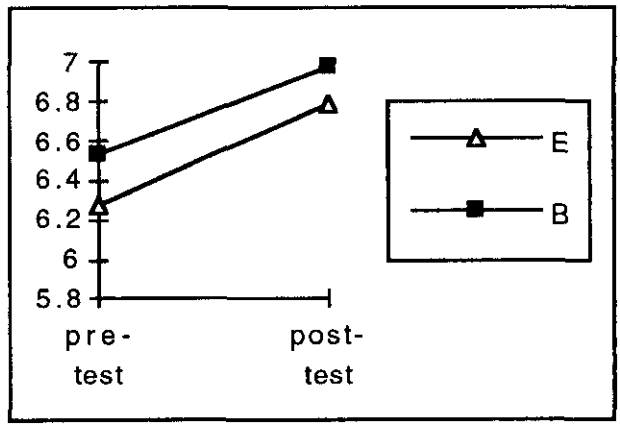


Figure 7.5: Pretest and posttest means for Burt Word Reading Test (years)

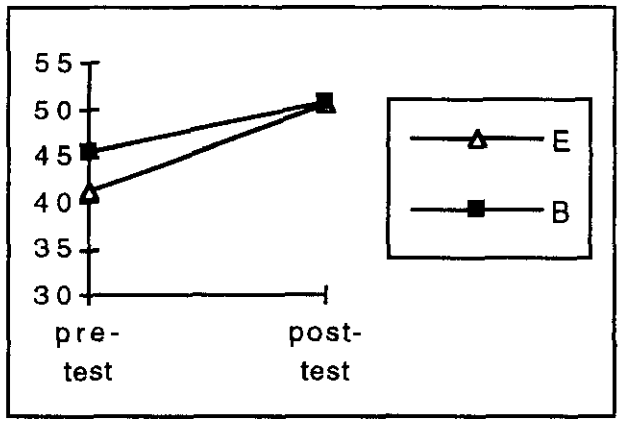


Figure 7.6: Pretest and posttest means for common 100 words test

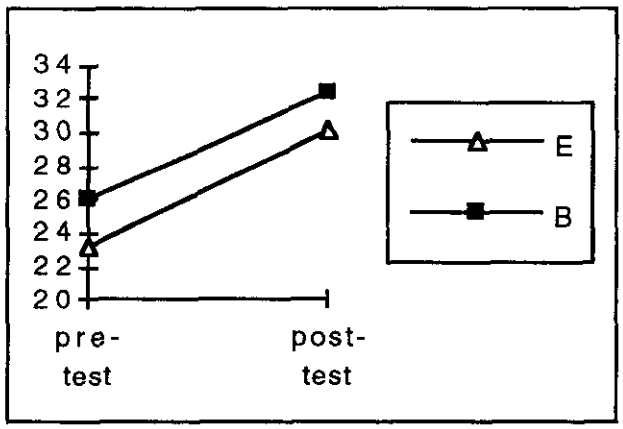


Figure 7.7: Pretest and posttest means for key word recognition test (maximum score = 44)

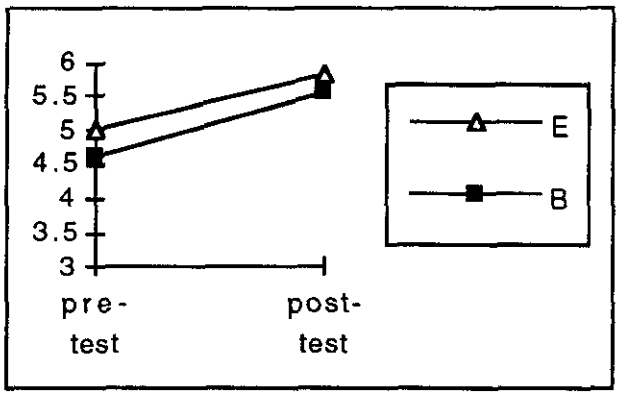


Figure 7.8: Pretest and posttest means for Phonic CVC Sub-test (maximum score = 12)

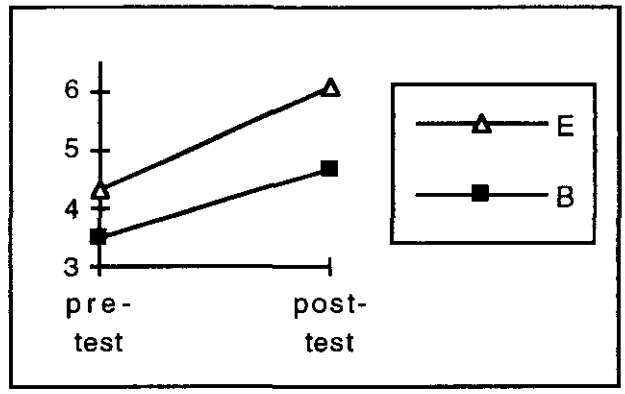


Figure 7.9: Pretest and posttest means for Phonic Blend Sub-test (maximum score = 9)

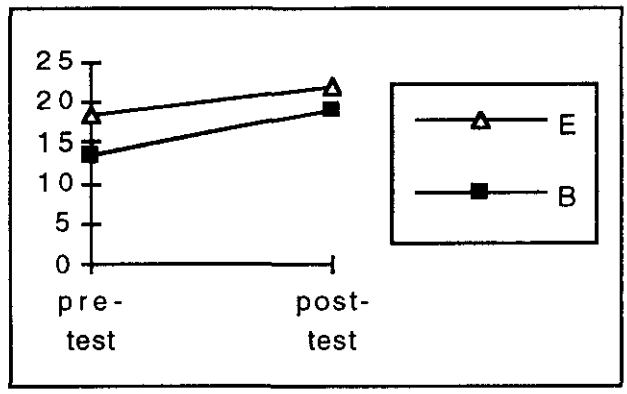


Figure 7.11: Pretest and posttest means for phonic tests (maximum score = 75)

Data was analysed using a repeated measures related ANOVA (2 x 2 x 3 x 5) with within subjects factors of program type (enhanced, basic) and time (pretest, posttest). In addition, between subjects factors for reading ability (low, medium, high) and class teacher (1-5) were considered. It was necessary to take reading ability into account as the range of pretest scores for all outcome measures had large standard deviations. The four pairs who failed to register a score on the Salford Sentence Reading Test before the intervention took place were categorised as low reading ability children. They each had a reading age of less than 6 years. It was considered that these children may not have been ready to read Stage 3 books independently, although they had been selected by their teachers as being suitable for the study. Four pairs, each averaging a score of 6 years and 11 months or greater on the Salford Sentence Reading Test, were identified to be high reading ability children. It was considered that the texts may not have been challenging to these children. The remaining 8 pairs were

categorised as being medium reading ability children. The pairs were matched as closely as possible for class to minimise the possibility of an effect from different teaching approaches. It was still necessary to take the teacher effect into account as a source of variance.

It was considered that the results of the statistical tests employed in this study, including ANOVAs and correlations, were significant if the alpha level was 0.05 or less. There was no overall significant difference in mean gains of all test measurements between the basic software users and the enhanced software users, as measured with repeated measures ANOVAs. Therefore, the results were analysed by reading ability group and are presented next. In the graphs in the next section, the low reading ability group is denoted by 1, the medium reading ability group is denoted by 2 and the high reading ability group is denoted by 3.

Results by reading ability group for sight recognition and phonic tests

It was assumed that the four pairs who had been categorised as low reading ability readers and the four pairs who had been categorised as high reading ability children may not benefit greatly from using the software. It was anticipated that the lower reading ability children would be more reliant upon computer support and that this would not facilitate improvements in word recognition. With the higher reading ability children, it was thought that using the software would not lead to improvements because they would not be exposed to many new words. Therefore, it was considered necessary to investigate the learning outcomes by reading ability to see if the assumptions were correct.

Figure 7.11 shows the mean gains by reading ability group for the standardised sentence reading test and indicates that both versions of the software were more beneficial for higher reading ability children. This was investigated by conducting a repeated measures ANOVA (2 x 3) with a within subjects factor of time (pretest and posttest) and a between subjects factor of reading ability (low, medium, high). Disregarding which version of talking books had been used, the difference in gains made on the Salford Sentence Reading Test between the three reading ability groups was statistically significant ($F(2, 29) = 4.68$, $p = 0.017$).

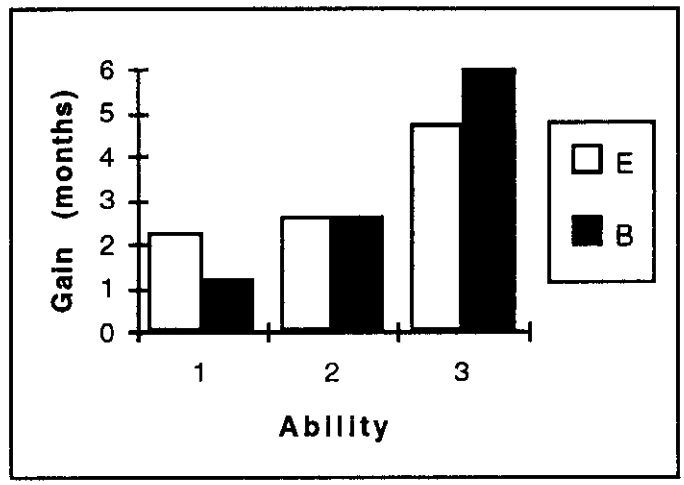


Figure 7.11: Mean gains by reading ability for Salford Sentence Reading Test

With the Burt Word Reading Test (Figure 7.12), in which words are presented in isolation, the magnitude of the gains made in the basic group was similar across the reading ability levels. It appeared as though the enhanced software was more effective for the better and medium reading ability children than it was for the low reading ability children. However, the differences were not statistically significant ($F(2,13) = 1.22$, $p = .328$). This may have been due to the variance of the gains made in the Burt Word Reading Test ($M = 6.12$, $SD = 5.52$).

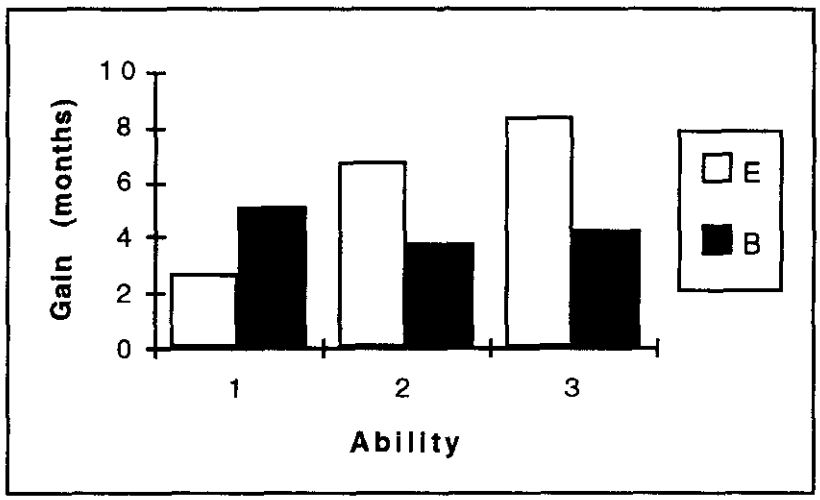


Figure 7.12: Mean gains by reading ability for Burt Word Reading Test

Figure 7.13 shows the mean gains by reading ability group for the common 100 words test. The results of the higher reading ability groups (3) are low due to a ceiling effect of the tests irrespective of the software version used. Neither of the low reading ability groups made large mean gains despite having low scores at pretest and the potential to make a greater improvement than the higher reading ability levels.

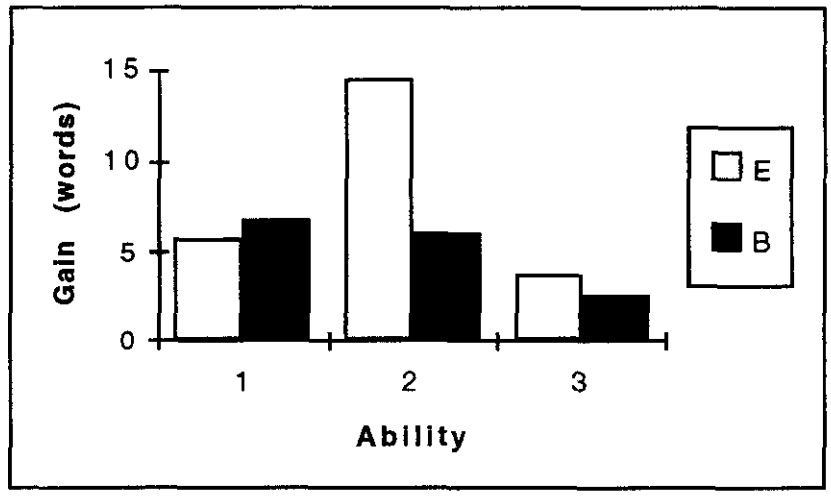


Figure 7.13: Mean gains by reading ability for common 100 words test

The difference between the medium reading ability children who used the enhanced and basic software versions was explored further in the common 100 words test. It was not statistically significant for this group: $F(1,3) = 5.45$, $p = 0.102$. The difference between the mean gains made by the two software groups appeared to substantial. However, the standard deviations of the scores were high (enhanced user group, SD = 9.75; basic user group, SD = 8.93). It may be that a significant difference between software versions could be demonstrated with a larger group of participants because this would reduce the effect of the large variances in scores.

Figure 7.14 shows the mean gains by reading ability group for the key words test. The results of the higher reading ability groups (3) are again low for both software groups because of a test ceiling effect. The low reading ability children had low scores at pretest in key word recognition irrespective of the software version used. This was probably due to previous lack of exposure to these words because they had not yet reached stage 3 of the

reading scheme. Therefore, they had greater potential to make gains because the key vocabulary is repeated regularly throughout all the texts at stage 3 of the reading scheme.

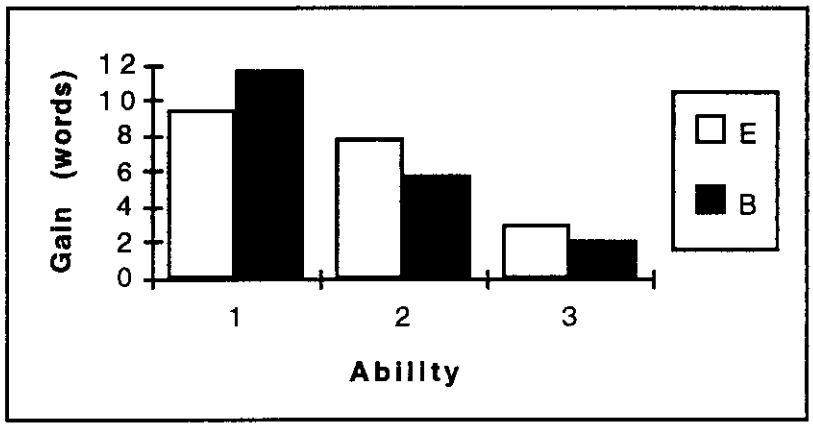


Figure 7.14: Mean gains by reading ability for key word recognition test (maximum score = 44)

The difference between the gains made by the medium reading ability groups was explored further. The analysis was restricted to the medium reading ability pairs only (8 pairs). In this case, the enhanced software users made significantly greater gains in key word recognition in comparison to the basic software users ($F(1,3) = 11.38$, $p = 0.043$). This statistical result provides support for the positive influence of the enhanced software on gains made in the sight acquisition of key vocabulary. The difference between the gains made was less than that of the common 100 words test, but the standard deviations were lower (enhanced user group, SD = 3.87; basic user group, SD = 3.01).

It was hypothesised that the onset and blend activities, together with feedback at the level of onset and rime would positively affect phonological decoding skills and in particular blending skills. The results of the CVC test and overall phonic test (Figure 7.15 and Figure 7.16) show that all reading ability ranges made little progress.

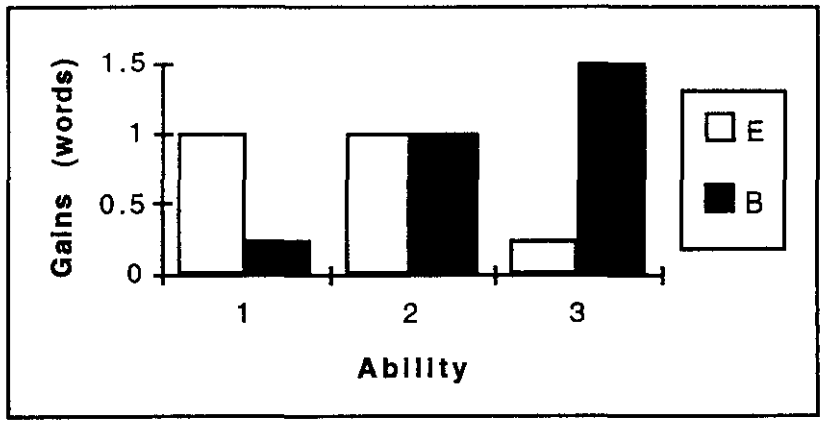


Figure 7.15: Mean gains by reading ability for CVC test (maximum score = 12)

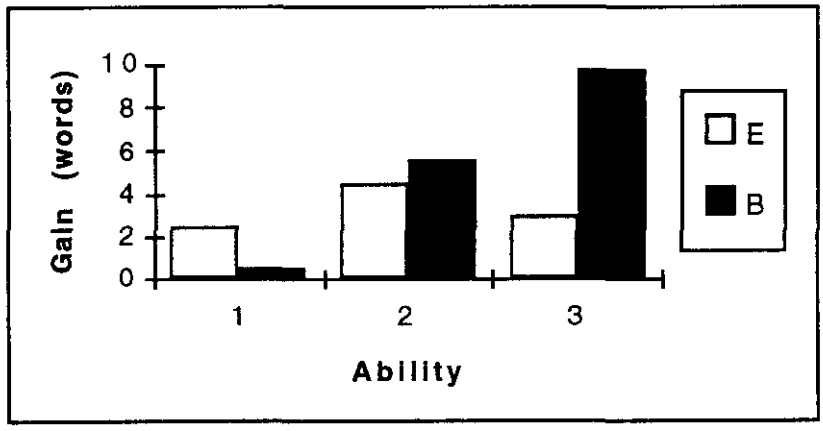


Figure 7.16: Mean gains by reading ability for phonic test (maximum score = 75)

The high reading ability levels of children who used the basic version of the software made greater mean gains in these tests than those children who used the enhanced software. This was due to the matching process, which did not take all the test scores into account. In the high reading ability group, the enhanced software users had stronger phonic knowledge prior to using the software ($M = 41.5$, $SD = 11.6$). In comparison, the basic software group ($M = 31.3$, $SD = 17.3$) had a lower mean score and were able to make greater gains.

The blending skills test results (Figure 7.17) indicate that similar gains were achieved by both groups across the reading ability levels. As the enhanced software was designed to develop phonological awareness, two Pearson product-momentum correlation coefficients were calculated for the 16 enhanced software users only. Firstly, it was predicted that there

would be a positive correlation between the gains made in the blend phonic test and the number of onset and blend activities undertaken. Secondly, it was predicted that there would be a positive correlation between the gains made in the blend phonic test and the number of words for which onset and rime feedback was given. These coefficients were .1765 (p = .513) and -.3592 (p = .172) respectively and were not statistically significant.

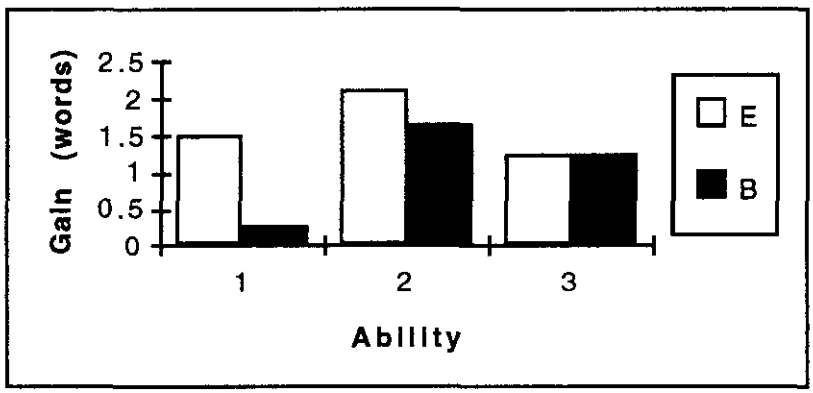


Figure 7.17: Mean gains by reading ability for blend test (maximum score = 9)

One issue arising was the validity of the blending skills test in measuring the benefits of coaching at the level of onset and rime. The test concerned the blending of phonemes or the smallest units of sound, rather than onsets and rimes. A test in which the children were asked to reconstruct words from their onsets and rimes may have been more appropriate. The onset and rime feedback is discussed more fully under technical issues (section 7.3.4, Feedback modes).

Exploring the relationship between the key word activity and sight recognition of key vocabulary

It was hypothesised that the provision of an activity designed to develop sight recognition of key words would positively affect the gains made in relation to the basic software group. Therefore, this was investigated by looking at the relationship between the key word activity and the measured gains in the key word recognition test for the 16 enhanced software users only. It was anticipated that a positive correlation would occur between the number of activities attempted and the key word recognition gains that were made. However, the

Pearson product-momentum co-efficient was -0.6627 (p = 0.005), providing evidence of a significant negative correlation (Figure 7.18).

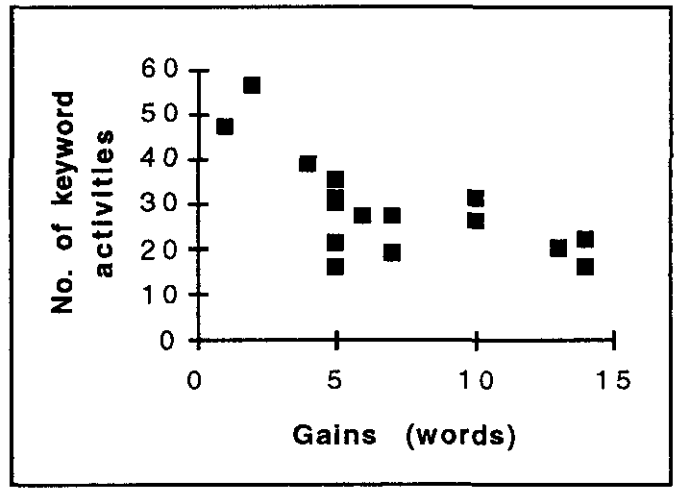


Figure 7.18: Relationship between key word gains and the number of key word activities attempted for the enhanced group

Further analysis revealed that there was a significant positive correlation between the number of activities participated in and the initial Salford Sentence Reading Test score, Pearson product-momentum co-efficient = 0.7590, p = 0.001. The higher reading ability children were able to read the complete texts in less time than the lower reading ability children and were more competent at the activities. Thus, they completed a greater number of activities during the session than the lower reading ability children. The lower reading ability children had low key word recognition scores and benefited from the additional practice and exposure to key vocabulary. In contrast, the higher reading ability children reached the test ceiling and despite completing more activities they had smaller gains to make.

Use of cues to aid independent decoding

It was considered that providing activities and hints in the enhanced software designed to provide coaching in the use of reading cues would have positive effects. This was investigated by considering teacher assessments of reading cue usage before and after using both versions of the software. In addition, the children were observed on a weekly basis. Their reading behaviour was noted in an effort to identify whether or not there was a change in their use of reading cues over time.

The teachers were asked to specify whether or not each of five reading cues (initial letter, word analogy, meaning, syntax and illustration) was used regularly by the children. They indicated their judgements for each child both before and after the software was used. Table 7.3 provides a summary of how many children were using each cue both prior to using the software and immediately afterwards.

Table 7.3: How many participants were using each cue before and after using the software (teacher assessed)

Cue	Enhanced Software		Basic Software	
	pretest	posttest	pretest	posttest
Initial letter	11	14	10	13
Word analogy	1	4	2	6
Meaning	7	10	6	6
Syntax	3	8	2	3
Illustration	11	13	10	13

Similar gains in strategy use were made by both groups for the regular use of initial letter, word analogy and illustration. The gains in strategy use were considered further by reading ability group. Lower reading ability children made gains in the use of initial letter. Higher reading ability children made gains in the use of word analogies. The teachers perceived that the children in the enhanced group's use of meaning and syntax cues improved more than their peers in the basic software group.

Each child was observed once a week and a Running Record (Clay, 1993b) was taken to enable miscue analysis to be conducted (chapter 3, section 3.4.3). This was problematic due to frequent interruptions by the children involved in the study and indeed, teaching staff who were unaware that systematic observation was in progress. Another issue was that a child must read texts at an appropriate level of difficulty for miscues to be generated. Seven children did not generate more than 10 miscues in total from the four observations and were discarded from the analysis. Three of the lowest reading ability children read with the computer (selecting most of the words for feedback) making miscue analysis difficult. The observation data for three children was incomplete due to absences. A total of 7 children

using the basic software and 6 children using the enhanced software were discarded from the analysis. There were no significant differences in strategy use over time between the two groups.

Figure 7.19 shows that the enhanced software users made a greater number of miscues initially that they were unaware of, and they were less reliant than the basic software group on instant feedback. That is they selected fewer word pronunciations. The basic software users made fewer reading for meaning errors and the number of these made did not change greatly over the four observed periods.

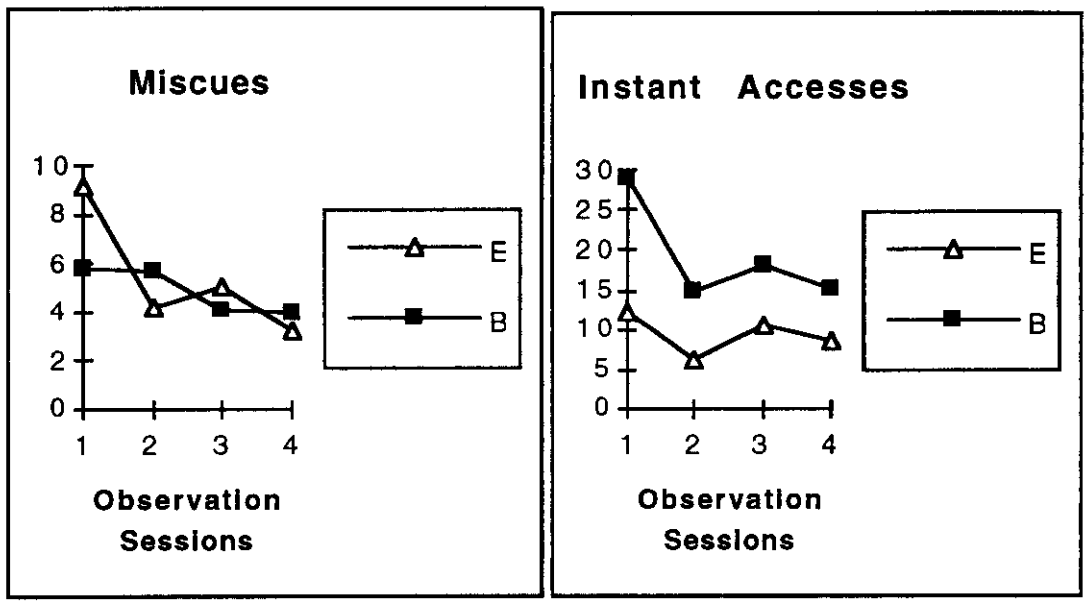


Figure 7.19: Number of miscues and words that were instantly accessed over four observations

The summary statistics presented in Table 7.4 from the software interaction logs for each child seem to confirm that the basic software users were more reliant upon instant feedback than the enhanced software users. It could be argued that this was due to the fact that the basic group were reading each text twice. However, the number of feedback requests during the second reading of the text decreased substantially. The enhanced group were also reading each text twice, because after completing the story they were instructed to repeat each sentence prior to invoking an activity.

Table 7.4: Summary of feedback requests and time spent for enhanced and basic software users

	Enhanced Group	Basic Group
Mean time overall	4 hours 30 minutes	3 hours 24 minutes
Mean no. requests for word feedback	463.1 (394.9 words, 67 hints)	641.9
Mean no. of key words given	172	249.8

It should be noted that the enhanced group spent more time on average using the software than the basic group. However, part of this can be attributed to the difference in the two versions of the software. For example, those children who received hints had to wait and listen to the suggestion that was given and then request the pronunciation again. The activities were time-consuming as well. The children had to listen to instructions and had to manipulate objects on the screen. The children who used the basic software actually read two books during each session and were exposed to as much text as those children who used the enhanced version.

Summary of results relating to cognitive issues

- The mean gain in the Salford Sentence Reading Test was just over 3 months for both the enhanced and the basic software users. The more able children made significantly greater gains than the less able children, irrespective of the software version.
- In the key word recognition test, the medium reading ability children in the enhanced software group made significantly greater gains than the medium reading ability children who used the basic software. The lowest reading ability groups made the greatest gains irrespective of the software version used and this is likely to be due to increased exposure to the vocabulary.
- There was no correlation between the gains made in the blend test by the enhanced software users and the number of onset and rime activities undertaken. There was also no correlation between the gains made in the blend test by the enhanced software group and the number of words for which onset and rime feedback was given.
- There was a significant negative correlation between the number of key word activities undertaken and the gains made in the key word recognition test. The higher reading ability children had more time and were able to complete more activities. However, they were approaching the test ceiling for key word recognition prior to using the software and had smaller gains to make.
- The teachers perceived that a greater number of children who used the enhanced software improved their use of meaning and syntax cues than those children who used the basic software.
- Analysis of the observational data indicated that there was no difference in strategy use between the two groups.

- The basic software users made more requests (63%) for word pronunciations than those using the enhanced software.

There was no overall difference in cognitive outcomes for the two groups, although the basic software users were reliant on word pronunciations and the enhanced software had a significant effect on key word recognition. However, the data relating to the affective domain of the framework for evaluation gives a different picture and is reported in the next section.

7.3.3 Affective issues

The affective issues of the framework for evaluation concern motivation and self-confidence. It is desirable that such software should have a positive affect on a child's level of motivation and self-confidence. This was assessed through interviewing the children before and after using the software, and through observations. Their teachers were also asked to rate each child on motivation and confidence levels on a scale of 1 (negative) to 5 (positive). In addition, the children's attitudes towards using the computer to practise reading are reported.

Motivation towards reading as assessed by teachers

From the teacher's assessments (Figure 7.20), it appears that the motivational levels of the enhanced software users became more positive when compared to the basic software users group, following the intervention period.

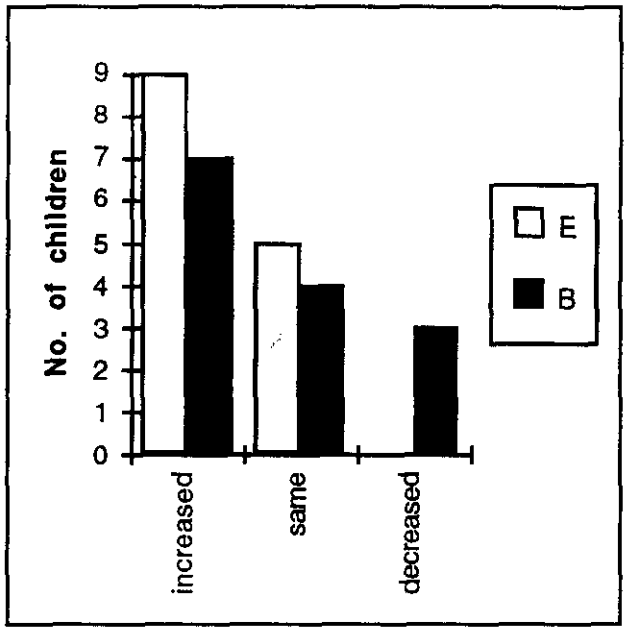


Figure 7.20: Overall changes to motivation levels as assessed by teachers

Confidence in reading as assessed by teachers

The teachers considered that the affect of the talking books on each individual's confidence level was positive, regardless of software group (Figure 7.21).

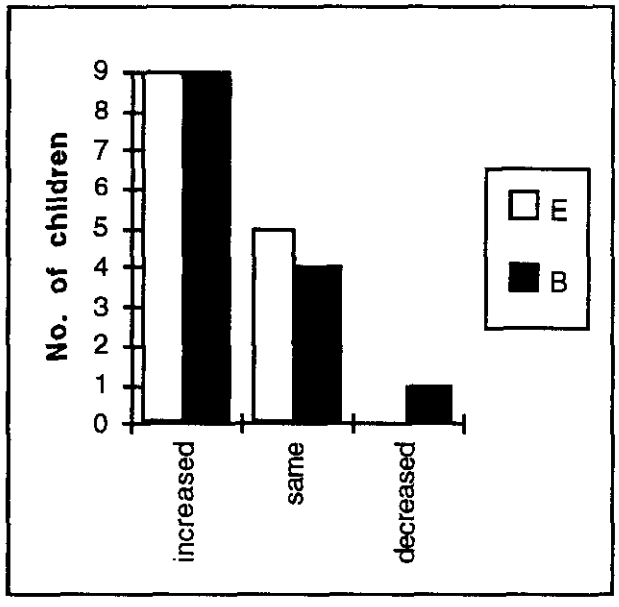


Figure 7.21: Changes to confidence as assessed by teachers

Participants' perceptions of motivation

The children were asked a number of questions before and after using the software to identify their attitudes to reading (appendix B).

Prior to using the electronic books, both the enhanced and basic software groups expressed positive attitudes to reading. At posttest, some of the enhanced software group expressed a more negative attitude towards reading than at pretest whereas the children using the basic software maintained a positive attitude. Six of the enhanced software users enjoyed reading less after the intervention compared to one child in the basic software group.

All children were asked what their favourite and least favourite school subjects were from a choice of reading, writing, science and mathematics as an alternative means of investigating their motivation towards reading. Figure 7.22 shows that the enhanced software users had a more negative attitude to reading after using the software whereas the basic software group appear to have developed a more positive attitude.

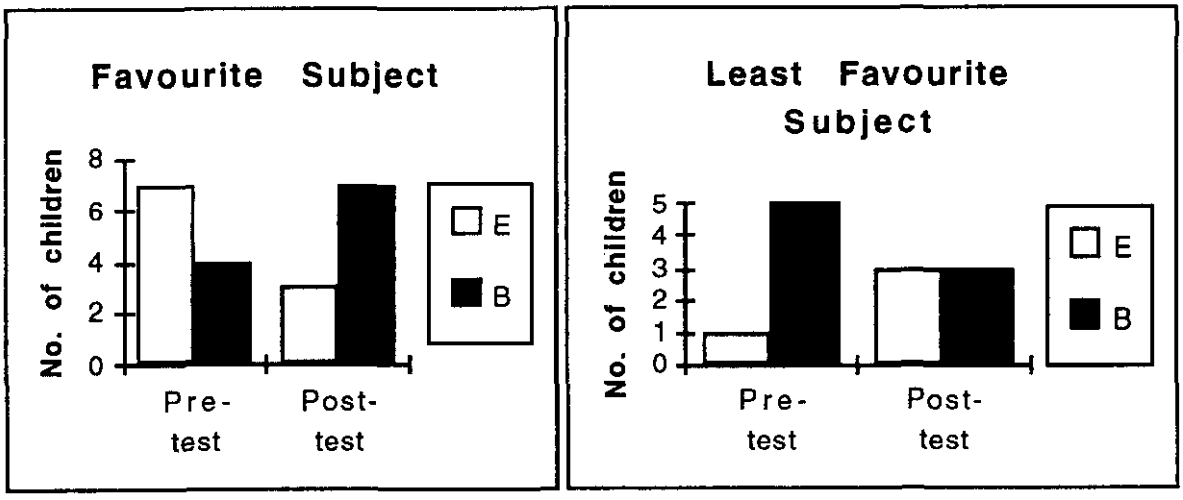


Figure 7.22: The number of children in each software group who identified reading as their favourite and least favourite subjects

The children were asked at posttest if, having used the computer to practise reading, they enjoyed reading more, the same or less (Figure 7.23). Again, it seems that the basic group developed a more positive attitude to reading than the enhanced group.

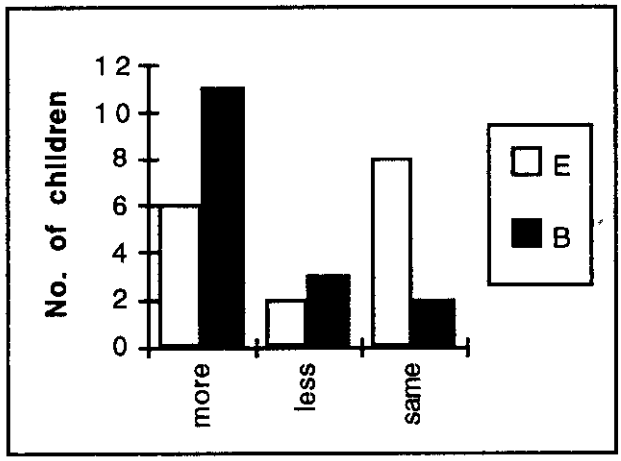


Figure 7.23: Perceived change in enjoyment of reading

Participants' perceptions of confidence

The children were asked a number of questions before and after using the software to identify their self-confidence levels in their reading ability (appendix B).

In both groups, the number of children who perceived that reading was the most difficult subject fell from pretest to posttest (Figure 7.24). In the enhanced group the number of children who identified reading as being the subject that they found the easiest increased. In comparison, fewer children identified reading as the easiest subject in the basic group after using the software.

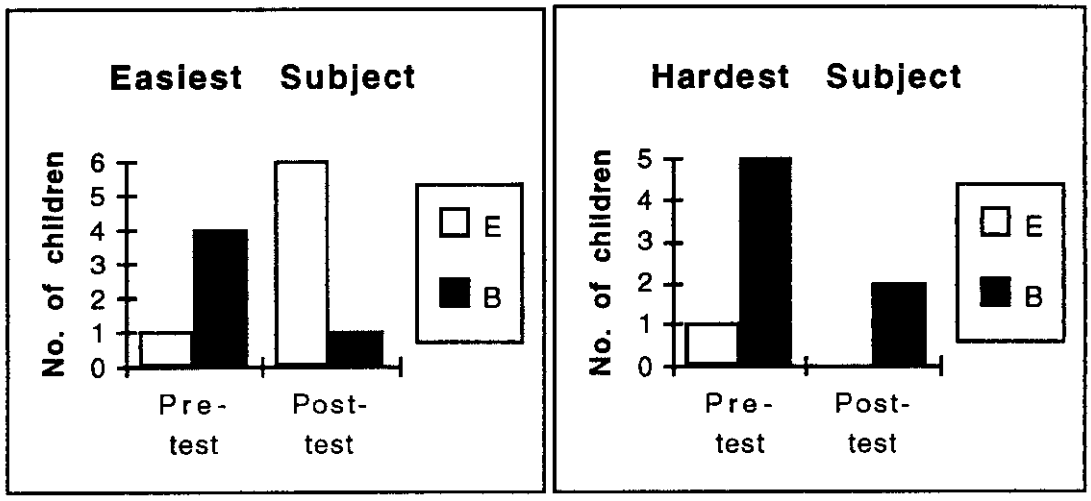


Figure 7.24: The number of children in each software group who identified reading as the easiest or hardest subject

Figure 7.25 summarises the children's perceptions about their own reading ability. They were asked whether they thought that their reading ability had got better, stayed the same, or got worse.

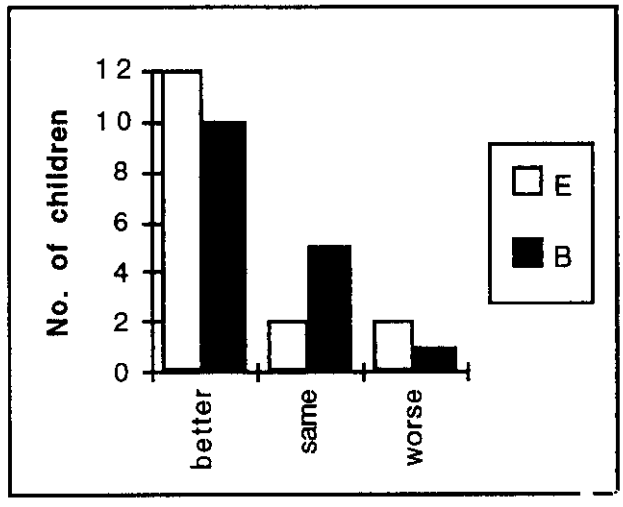


Figure 7.25: Self-confidence in reading ability

The majority of both the enhanced group and basic group perceived that their reading ability had improved. Yvonne, who used the enhanced software, commented that she felt her reading had improved because: "I used to always press the words and now I don't". For her, a perceived decrease in reliance on computer support had clearly boosted her confidence. She was able to judge for herself that her reading was improving.

Participants' attitudes to computers and reading

All but one of the enhanced software group and all but three of the basic software group indicated a positive attitude to using a computer. All the children were asked at posttest whether they would like to spend more time using the computer, the same amount of time or less time than they had been during the study (Figure 7.26). Both groups made similar responses, with at least half the group members indicating that they would like to spend more time on the computer than they had been.

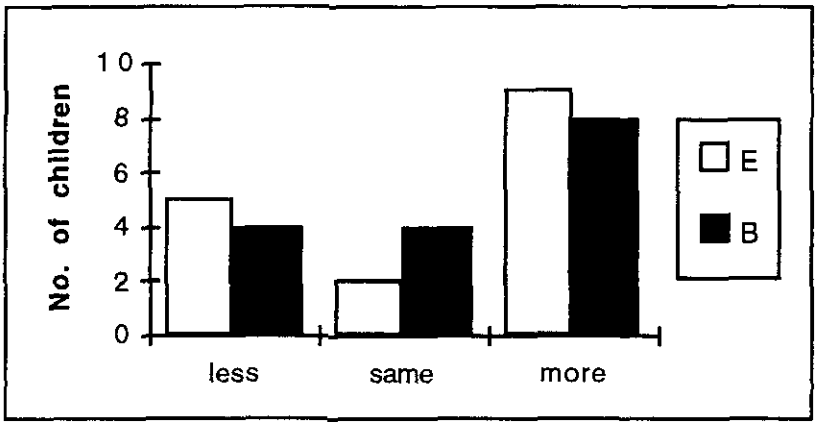


Figure 7.26: Desired changes to frequency of computer use

Two of the children who indicated that they would like to spend less time using the computer, one from each group, expressed concerns about missing other classroom activities. They were from school C, at which the children used the computer in a separate room from their own classroom.

The children were asked how they felt about different situations where they might be reading on a scale from 1 (do not like it) to 4 (enjoy it a lot). The situations were reading:

- on their own (alone)
- to their parents

- to their teacher
- with the computer

The children's attitude towards reading alone (without the presence of anyone else) (Figure 7.27) changed more dramatically than their attitudes towards the alternative ways of practising reading, which remained similar from pretest to posttest. The enhanced software users developed a more positive attitude towards reading alone whereas the basic users were more positive about using the computer to practise reading (Figure 7.28).

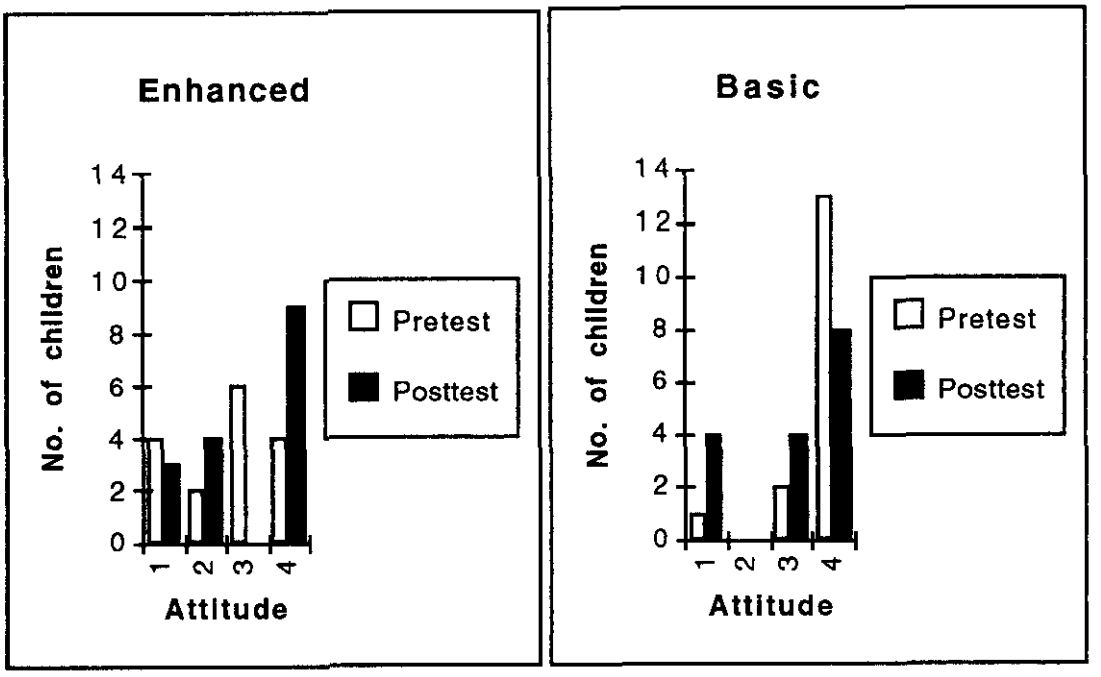


Figure 7.27: Children's attitudes towards reading alone

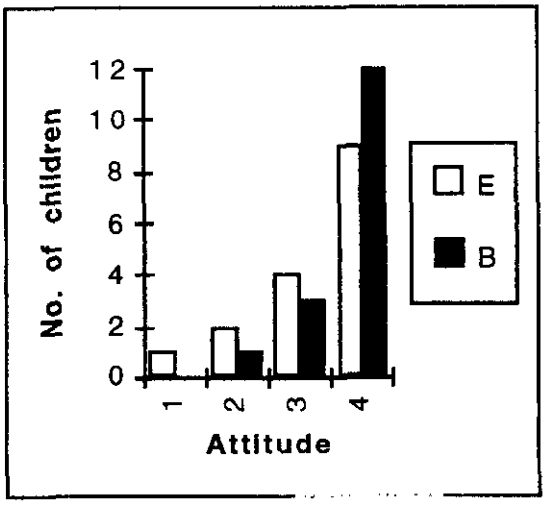


Figure 7.28: Children's attitude toward reading with the computer after using the software

All but 3 of the enhanced software group indicated that they would like to continue to use the software to practise reading. Of these three, Edward stated that he would rather read 'real' books and that he had not found it very exciting. The other two children could offer no reason as to why they did not wish to continue. In the basic group all but one child wished to continue [Miles], who expressed a desire for a short break but commented that he would like to carry on shortly.

The children in the enhanced software group who expressed an interest in continuing gave the following reasons. Seven children said that they thought it was fun and they enjoyed it. Five children perceived that it helped them learn to read. The children in the basic software group who expressed an interest in continuing gave the following reasons. Nine children said that they thought it was fun and they enjoyed it. Four children felt that it helped them to become better readers.

Summary of results relating to affective issues

- The teachers perceived that the enhanced software users were more positively affected by the experience than the basic software group with regard to motivational levels.
- From pretest to posttest, the enhanced software group developed a more negative attitude to reading whilst the basic group maintained a positive attitude.
- The teachers perceived that the confidence levels of both groups were improved.
- The confidence levels of both groups were improved, but more so for the enhanced software users.
- Both groups were positive about using computers.
- The enhanced group developed a more positive attitude to reading alone from pre to posttest whereas the basic group were more positive than the enhanced group about reading with the computer at posttest.
- Overall, both groups expressed a desire to continue using the electronic books with:
    - more of the enhanced group than basic users perceiving that it would improve their reading ability
    - more of the basic group than enhanced users commenting that they felt it was fun

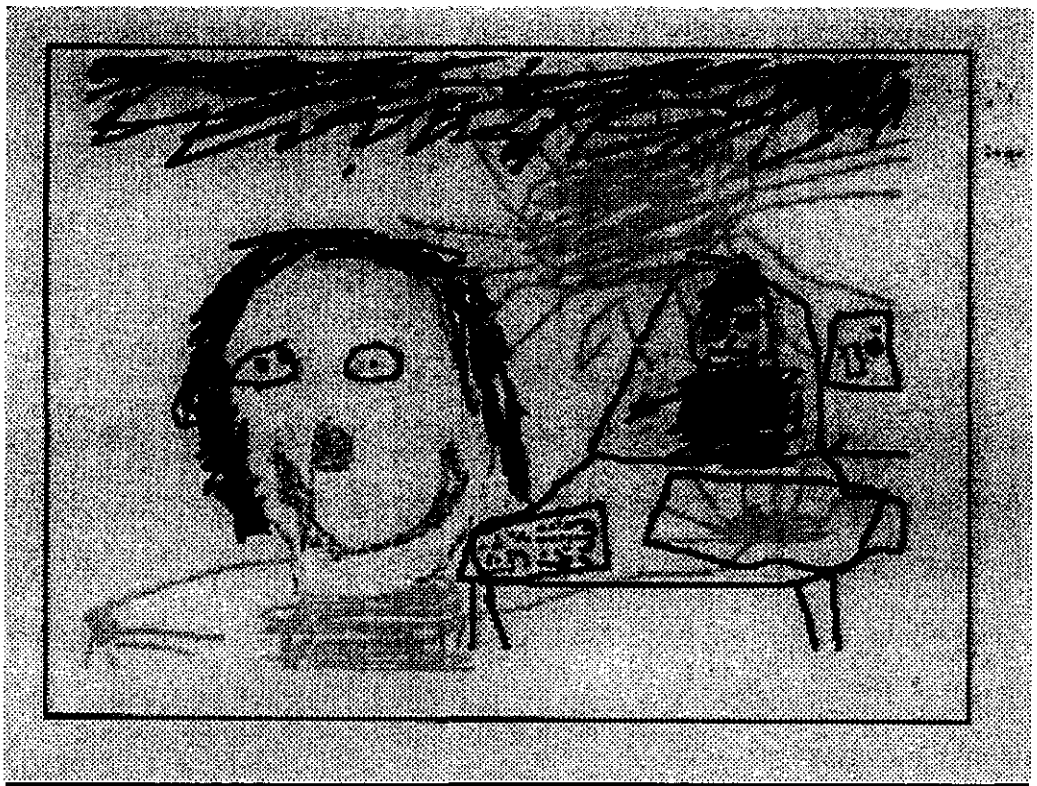
7.3.4 Technical issues

Technical issues relate to exploiting technology to its maximum advantage, operational transparency, instructional adequacy and cost-effectiveness. In this study, the usability of the two software versions has been monitored through observing the children using the

computer. The effectiveness of the modelling mechanism has been assessed through analysing the software interaction logs In addition, the children were asked to identify at the posttest interview what they thought the benefits and the disadvantages of using the software were.

Feedback modes

During the posttest interviews the children were asked what had been beneficial about reading with the computer (appendix B). Several children from both groups made comments relating to the types of computer feedback available and are described below. The software design included a number of forms of feedback (chapter 5, section 5.3.4). Sentence level feedback could be invoked by clicking on a button and then the complete text on the page was read aloud with expression and intonation. Whole word and segmented feedback were available for individual words.



I lac the activities and wen I rid the whds at the hend I lak the kopyut to rid ol of the book to me. I wud wont to gow on it ol diy if I cud.

[I like the activities and when I read the words at the end I like the computer to read all of the book to me. I would want to go on it all day if I could.]

Figure 7.29: One participant's opinion of the enhanced talking book software

Of the enhanced software group, nine of the 16 children referred to obtaining whole word pronunciations for unknown words as being one of the best features of the software. Four children mentioned how useful the sentence level feedback was for checking whether they had read the text correctly (Figure 7.29). Of the basic software group, five of the 16 children referred to obtaining whole word pronunciations for unknown words as being beneficial. Again, five children mentioned the sentence level feedback and how it facilitated self-checking:

> Because I like to hear if I'm right.
>
> I get all the words right. I press 'speak' and I get it all right.

No comments were received on the feedback at the level of onset and rime. For eight of the talking books, feedback was given at the level of onset and rime by pronouncing the onset, then the rime and then the complete word. Simultaneously, the onset, rime and then the

whole word were highlighted. For example, feedback for the word 'shut' would be given as 'sh', 'ut', 'shut'. Some of the children were seen to blend the onset and rime, following the presentation and highlighting of these segments, and then vocalise the complete word before its pronunciation was given. They had not been instructed to do this. They seemed to have taken the opportunity to create an element of competition by synthesising the word from the parts given and then trying to beat the computer and say the whole word first.

The modelling mechanism

In order to emulate a human tutor and provide tailored instruction, a modelling mechanism was implemented. It was designed to represent a child's use of five cues (onset, rime, meaning, syntax and illustration) and their competence at key word recognition. Each model was set initially using information provided by each teacher on which strategies were used on a regular basis. The models were then adjusted according to words selected for feedback and the activities undertaken, each of which was designed to provide coaching in the use of a cue or in key word recognition. If a word was selected for feedback it was assumed that one or more of the cues had been under-utilised and thus their associated models were decreased. The model representing a cue for which coaching was being given through an activity was incremented if it was completed and decremented upon each unsuccessful attempt.

Analysis of the software logs revealed that the modelling mechanism was not sensitive enough for those children who required more support with decoding words. By selecting as little as 10% of the words for immediate speech feedback, the models that were based on linear scales of 1 - 10 were soon all reduced to one. Each time a word was selected, a number of cue models were decremented.

For those users who requested less than 10% of the words for immediate feedback, the modelling mechanism appeared to function successfully. An example is presented in appendix K.

In the current implementation, the use of this feature in normal classroom situations would probably be impractical. It was originally intended to make this fully automatic, but without the incorporation of speech recognition in this implementation, teacher input was required in

the initial stages at least. Furthermore, the teachers completed paper forms only and did not use CARIS to input this information. This was performed by the researcher. However, this feature could be useful for children that required additional support and monitoring. One teacher participating in the research commented that she would only need to monitor closely those children who were experiencing problems in specific areas.

The hint mechanism

Of the 16 children who used the enhanced software, only ten were identified by their teachers as potentially benefiting from receiving a hint first to help them to decode unknown words. During observation sessions it was noted that many of these children soon learned that clicking twice on the target word in rapid succession enabled them to obtain the pronunciation. Initially, the hint offered seemed to be considered, possibly due to the novelty of the interaction, but this observed behaviour soon diminished over time. The participants involved later showed signs that this style of interaction was frustrating and/or irritating to them. For some of these children, it was clear that this style of interaction did not suit their specific needs. It did not prevent users from becoming reliant upon instant feedback, if they chose to do so. One child [Justin], when answering a question concerning what was good about reading with a computer, referred indirectly to clicking twice on a word:

> Because it says 'Try Again' and if you press again it will tell you the word.

Another child [Jessica] answering a question concerning problems related to reading with the computer commented:

P: When I don't know a word I click on the thing and it says 'Try starting again at the beginning of the sentence'.
I: Right and what don't you like about that? Is it annoying for you?
P: Yeah and I have to press it again.
I: And why do you press it again?
P: 'Cos then it doesn't tell me the first time so I have to press it again.

The six children who did not have mandatory hints were shown how to access hints if required. It was anticipated that having the option of accessing a hint would assist those children who might be fearful of making mistakes. One child [Edward] referred to it in the posttest interview:

P: An' also I like it because ... because it's got this kind of clue thing.
I: Oh right. Did you ever use that?
P: Just once.
I: Oh right and what did you think of it?
P: 'Guinea-pigs' ... I thought it was 'guinea-pigs'.
I: But what did you think about the clue. Did it help you?
P: Uhuh. Just a little bit. And then I clicked on the word and it said it was 'guinea-pigs'!

Other children, for whom hints were not mandatory, were observed to toggle the hint button on and off. Sometimes they obtained hints for words. The behaviour that they exhibited indicated that they were experimenting with an interaction feature rather than deliberately choosing to obtain a hint.

Activities

A number of activities were implemented to provide coaching in the use of reading cues and improve sight recognition of the key vocabulary (chapter 6, section 6.5.3). The activities were well received by all the children using the enhanced software. Some of the children using the basic software asked why they did not have activities in the Talking Books that they had used. The children using the enhanced books were restricted to 15 minutes use. However, in many cases, they would have happily continued for longer periods.

Six children who used the enhanced software commented that they thought that the activities were a useful feature. The following extract from an interview is given to illustrate that some children perceived them as being beneficial. In this case, the child not only indicated that the activities were a good feature of the software, but implied that the sentence shuffling activity had helped him to read the sentences in the story.

I: What is good about reading with the computer?
P: 'Cos we were doing activities. 'Cos of that door thing where you had to press them letters ... sound things.
I: Uhuh. So that was your favourite activity then was it?
P: Yeah!
I: Why do you think that was good for you and your reading?
P: What do you mean?
I: Do you think that it helped you with your reading?
P: What ... erm computer?
I: The activities.
P: Yeah!
I: Why did they help you with your reading?
P: 'Cos ... erm ... we did those sentenc' and ... [illegible] they were same sentences as words that were on picture when you were reading it.

Teacher's comments relating to the appropriateness of the activities have already been reported (0). Three children referred to a bug in the illustration activity when asked if they had experienced any problems with the software. One child [Mark] commented that the sentence shuffling activity and the illustration activity were too time consuming.

Usability issues

During the observed sessions a number of usability issues came to light. Some concerns had been raised regarding the 'drag and drop' interface style and whether young children would find this concept difficult. This was not found to be the case. Even the less confident users experienced little difficulty with this operation. The only problem that became apparent was that it was possible to drop a word card in the sentence shuffle activity or the illustration activity under the menu bar. In these cases, the word card became no longer accessible (see chapter 6, section 6.5.3 for examples of activity interactions). This technical problem lay with the development environment.

Some concerns had been raised about the button sizes and mouse manipulation skills of young children. No problems were observed. However, sometimes the children experienced difficulties when clicking on individual words, particularly when faced with more than one line of text. This seemed to be more problematic on the rhyme pages of the eight pattern and rhyme books, where the text was smaller and closer together. Figure 7.30 is an example of the type of problem that occurred. The child attempts to click on the word 'three'. The cursor is presented as a hand with a pointing finger, and the point is defined as the cursor hot spot. In this case, it appears to be pointing to the correct word. However, the cursor hot spot is actually over the space occupied by the word 'mittens' in the line below. In this case, the word 'mittens' is highlighted and its pronunciation given. This was confusing for many children. It could be addressed by modifying the software such that each word is highlighted as the cursor passes over it, before the button is clicked. This would provide a visual check for the user that they were selecting the word that they actually wanted.

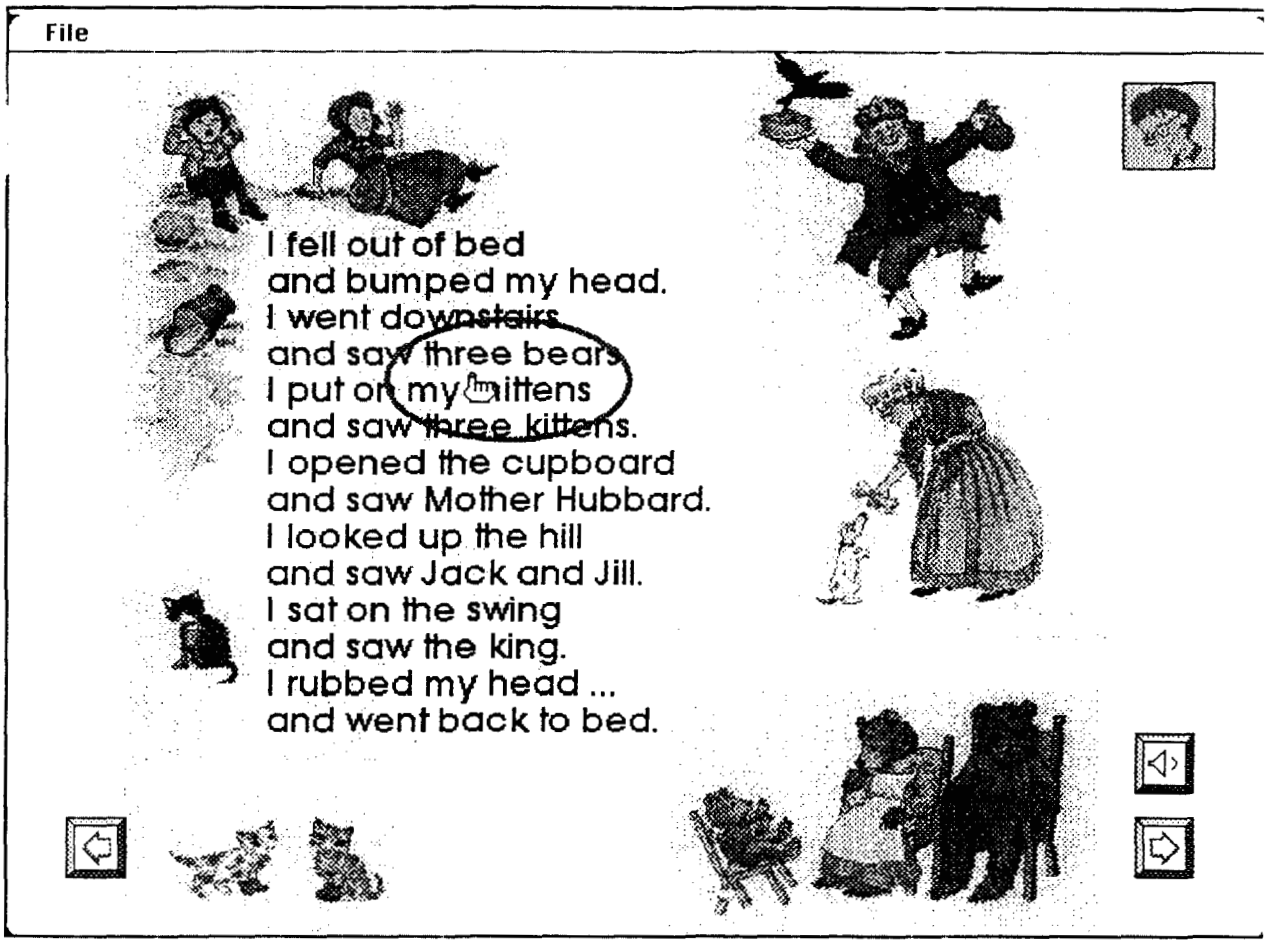


Figure 7.30: Example of difficulties in word selection

In the sentence shuffle and illustration activities it was possible to obtain the pronunciations of individual words by double clicking on the appropriate card. This appeared to be a difficult operation for many of the young children. Clearly, fewer children in this age range had developed their manipulation skills sufficiently enough for this type of interaction to be feasible.

Instructional adequacy

The children using the basic software had no problems learning how to use it. They found it straight forward. The original enhanced software design included spoken instructions relating specifically to the use of activities. During early formative evaluation, it seemed to be intrusive because the users had to wait for a number of seconds for the instructions to be spoken. The users appeared to learn how to complete the activities quickly without this time consuming intervention. Instead, all users were given a verbal introduction on how the

software should be used. Some of the children using the enhanced software found it a little difficult to remember the instructions that had been given. The procedure that they had to follow was more complicated than that followed by the basic software users. They had to remember to read the story first, then return to the beginning, read each sentence again and then do an activity on each page. It may have been less confusing for these children if the activity button had been disabled until they had read the complete book. Furthermore, a reminder of what to do after completing the story may have alleviated some of the problems that one or two of the children had.

Summary of results relating to technical issues

- The facility to obtain feedback for individual words was perceived by the users as being useful.
- The hints initially seemed to prompt the users to consider alternative strategies for decoding words independently. However, it rapidly became obvious that this was a novelty effect and all the users who received hints soon started to click twice on a word in order to obtain its pronunciation.
- The enhanced users selected fewer words than the basic users.
- By using the sentence level feedback not only were the learners able to identify any mistakes made, but they were able to verify that they had read the text correctly.
- There is little evidence to suggest that feedback at the level of onset and rime is beneficial.
- The activities were enjoyed by most of the children and had the potential to extend the period of concentration applied to reading tasks, although the time spent on task was controlled in this study.
- The modelling mechanism was not sensitive enough for those children who were more dependent on computer support but appeared to operate correctly for the rest of the users.
- Some usability issues were raised:
  - 'Drag and drop' style interfaces are appropriate for this age group.
  - The children did not experience any difficulties as a result of small button sizes, but found individual word selection to be problematic when the font size was small and the text was close together.
  - 'Double clicking' can be extremely difficult for children of this age group.

7.4 Case Studies

Five case studies, which serve two purposes, are now presented: four boys and one girl (their test results are summarised in appendix L). Firstly, they were used to consider the effects of both talking book versions on participants who were identified by their teachers as

lacking in motivation or self-confidence at pretest. Case studies are used to provide a deeper insight into these affective issues. Secondly, this approach can provide a much richer and more detailed view of the way in which an individual interacts with electronic texts that support reading instruction. This level of detail and richness of data cannot be achieved with purely quantitative approaches. The opinions of both the teachers and the children were elicited through questionnaires and structured interviews. However, it is imperative to consider alternative sources of evidence, and in particular observational data, rather than rely solely on the "static nature of questionnaire measures" (Oldfather & Wigfield, 1996, p106). As there was much variance both in the styles of interaction and outcomes, it is necessary to provide illustrative cases rather than summative results alone.

7.4.1 Framework for behavioural analysis

One of the primary sources of data is analysis of video recordings of the children using the software. A number of reading specialists have focused upon motivation and reading. Turner (1995) identified a number of literacy behaviours indicating that a child is motivated towards reading:

- effective use of strategies
- repeating words or sounds
- explaining the relevance of the picture to the text
- sounding out words
- pointing to words
- re-reading text to bolster comprehension
- predicting events
- interpreting text to peers
- persistent effort
- signs of enjoyment of stories (smiling or laughing)

In comparison, through analysing the video recordings of the case studies, examples of demotivated behaviour were identified as:

- physical signs of loss of concentration
- comments unrelated to text or task
- ignoring hints

- immediately accessing a word pronunciation

In the research reported here, it could be argued that ignoring hints and accessing word pronunciations are not signs of demotivated behaviour. In this research however, having a hint meant that the user was off task for a short period because they could not continue until the hint had been given. Selecting words for immediate speech feedback could be associated with either motivated or demotivated behaviour and depends on each individual, but in these cases below clearly indicates a lack of motivation. Physical signs of loss of concentration were defined as being obvious actions that were unrelated to the task such as playing with the microphone. It could be argued that a child who briefly looks away from the screen is not necessarily off task and so this is not considered to be a sign of concentration loss in this research.

Given these modifications, the framework suggested by Turner (1995) has been used to analyse the reading behaviours of the five children studied in depth here. Each child was video recorded after approximately one week and in 3 cases, again towards the end of the study. A full breakdown of observed behaviour for each child is included in appendix M.

7.4.2 Case study 1

Pretest. Jason was identified as lacking in confidence by his teacher. He had used the enhanced version of the software with both onset and rime feedback and hints enabled. In his pretests there was some evidence of use of phonics and initial letter when presented with words in isolation. He did better on the Salford Sentence Reading Test than on the Burt Word Reading Test (Appendix L). This implied successful use of meaning and syntax. He had experience of computers at home, as well as in school, and seemed quite at ease with their use. He indicated that he enjoyed reading a lot. He did not like reading alone and greatly enjoyed reading to both his parents and his teacher.

Interaction with software. Early observations recorded in field notes revealed that he did not require hints and would have made effective use of instant word pronunciations:

> [He] actually tries to decode unknown words himself before clicking on them. Therefore, when he clicks on a word what he really wants is instant feedback not a hint. Frequently, the hint directs him to a strategy that he has already attempted to use.

Later on it became apparent that he was clicking twice on the word to access its pronunciation and ignoring the hint. Apart from four selections, all words for which hints were given were selected again in order to obtain the pronunciation. Three of the four selections in which just a hint was obtained occurred within the first week. It is likely that the hints were used effectively at this time due to the novelty effect of this new feature. The child heeded the advice given, was able to decode the unknown word independently and did not need to access the word pronunciation. In sum, the hint mechanism did not help Jason but it interfered with his use of the software.

During the observed sessions, there was a decrease in the number of words that were instantly requested without any attempt to decode them independently. It seemed that he was making use of a variety of cues.

Feedback was given at the level of onset and rime for 65 words. Whole word feedback was given for 164 words, of which 33 were from the key vocabulary. Hints were given for 155 words. He spent a total of five hours and 57 minutes using the software, spending an average of 17.85 minutes per session. He was confident in its use and experienced no difficulties. He did not seem to become more reliant on immediate whole word feedback (Figure 7.31).

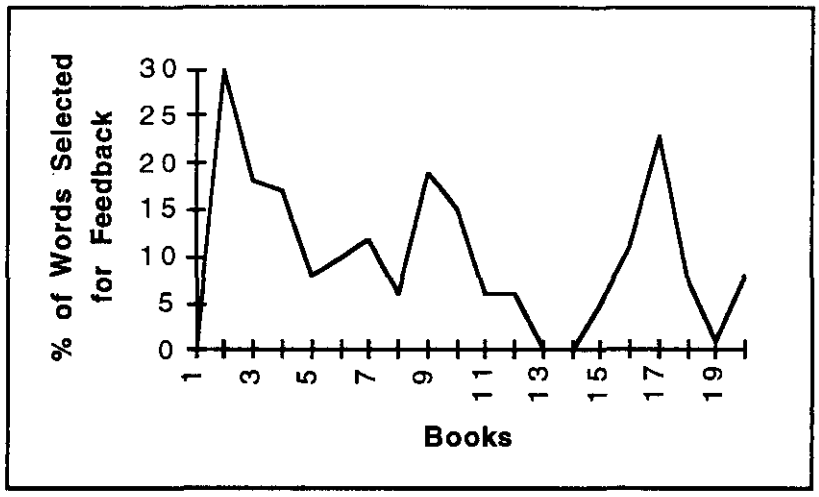


Figure 7.31: Use of instant feedback over 20 sessions

The majority of attempts at activities were successful the first time (Table 7.5). He therefore experienced a higher level of success than some children. Experiencing success at challenging tasks should improve confidence and motivation.

Table 7.5: Results of activities

Activity	First Attempt	Second Attempt	Third Attempt	Failure	Total
Sentence Shuffle	45	0	2	0	47
Key Word	16	0	0	n/a	16
Illustration	12	4	1	2	19
Onset	8	1	0	n/a	9
Blend	4	4	2	n/a	10

Observations. Jason was video recorded during two sessions and a summary of his observed behaviour is summarised in appendix M, section M.1. The first recording was made during the second week, to ensure that the child had become comfortable with using the software. The second recording was made during the last session.

In the first recording, he ignored all hints and clicked twice on each word to obtain its pronunciation. Physical signs of loss of concentration mainly occurred whilst the hints were being given. Whilst a hint is being given the software does not recognise any other actions such as mouse clicks. The child was aware that he had to sit back and wait before continuing, and took the opportunity to take a break. There were a number of signs of

motivated behaviour: attending to words as they were highlighted, repeating words, rereading text, pointing at words and enjoyment of the story. He paid attention to each picture and followed the highlighting of text as sentence level feedback was given. He acknowledged the achievement of completing the text: "I've done it!" He did make one comment indicating low self-esteem. It was related to use of the technology (not being able to double click with the mouse) rather than to reading itself: "I can't click on this word what it says. I'm not very quick at this." He read slowly and with little expression.

In the second recording, he made fewer miscues in reading the text and so the number of occurrences of motivating behaviours, such as reading text again and repeating words, was naturally lower. He paid attention to the illustrations and followed the highlighting of text when sentence level feedback was given. As in the first recording, he ignored all hints and clicked immediately on the words to obtain the pronunciations. Similarly, he showed physical signs of loss of concentration whilst the hints were being given. However, he also appeared to lose concentration during other periods. He was playing with the mouse, rapidly moving the pointer across the screen, throughout the session. Although he quickly and confidently completed two of the sentence shuffling activities, during a third he played with the word cards and eventually 'lost' one under the menu bar. His reading was slow, stilted and without expression. He made comments that were unrelated to the tasks. It appeared that Jason was becoming more demotivated towards reading with continued use of the software.

Posttest. His teacher commented at posttest that:

> [He] still lacks confidence in his reading and had to be prompted to use strategies.

She perceived that he was no longer making use of meaning and syntax when attempting to decode unknown words. His score on the Salford Sentence Reading Test remained unchanged and his score on the Burt Word Reading Test improved by eight months (Appendix L). However, he seemed less inclined to attempt unknown words and said "don't know" more readily. He improved his score through acquiring a larger vocabulary of words recognised by sight. This was also reflected in the common 100 words and key word recognition tests, in which he showed improvement (Appendix L). In the phonic tests, his ability to decode CVC words remained the same, but his blending skills improved.

Jason indicated that he did not like reading. He still did not like reading alone, but was happy reading to his parents and teacher, and also reading with the computer. He stated that he enjoyed using the computer, referring to the activities, but that its use worried him sometimes. He also perceived that he had been missing too much work in the classroom and would like to use the computer less frequently. He felt that his reading ability had got worse and stated that he enjoyed reading less. He identified instant word pronunciations and the activities as being the most helpful features of the software. At posttest, Jason's motivation towards reading had decreased and his self-confidence in his abilities remained low.

Summary of case 1

- The hints were not helpful because Jason only requested feedback after he had tried alternative strategies for decoding.
- Towards the end of the study he became less motivated and appeared to make less effort.
- His low self-confidence was not improved, despite experiencing a high level of success.

7.4.3 Case study 2

Pretest. Justin had also used the enhanced version of the software with both onset and rime feedback and hints enabled. He was described by his teacher as being "reluctant to tackle new words. Will sound out but then finds it difficult to blend" and she felt that he was lacking in motivation. Justin stated that he liked reading a little. He had exposure to computers at home as well as at school, and greatly enjoyed using technology. He stated that he did not like reading alone, but he enjoyed greatly reading both to his parents and teacher. In his pretests he made little use of initial letters in the Salford Sentence Reading Test, but their use was evident in the Burt Word Reading Test. He had progressed further through the reading scheme than his peers who were identified to participate in the study, yet his test scores fell in the average of the ranges obtained (Appendix L). He readily said "don't know" when he reached a previously unseen word, and as his teacher had commented, he didn't attempt to decode independently. He was not able to decode CVC words, nor able to blend sounds together.

Interaction with software. During the intervention period, his teacher commented informally that she felt that Justin was not using the software correctly. She had to keep a close eye on him. She gave the impression that she had to monitor him more closely than the other children in many classroom activities. He used the software for a total of five hours and 48 minutes, spending an average of 17.4 minutes per session. He did not appear to become more dependent on accessing immediate word pronunciations as the study progressed (Figure 7.32). He requested a high proportion of words for computer feedback whilst reading the first book, but it is likely that this is due to the novelty effect.

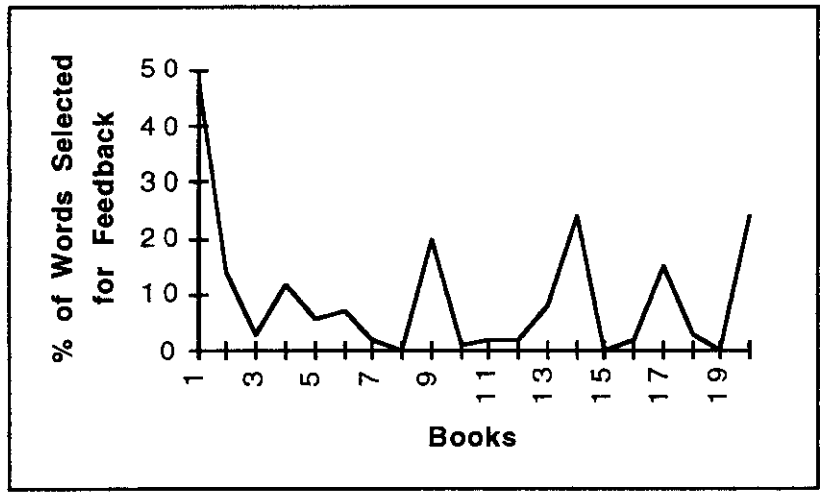


Figure 7.32: Use of instant feedback

Table 7.6 summarises the activity results. On 13 occasions Justin did not complete the sentence shuffling activity at the first attempt. As he was considered by his teacher to be one of the more proficient readers in the group, he should not have found this activity to be too challenging. Furthermore, he rarely completed the onset and rime activities correctly on the first attempt. He even made a comment that revealed that he had discovered that activities did not necessarily require much effort:

I sometimes click on OK afore I've put all the things in!

Justin, therefore, was not always motivated by the use of software and clearly avoided making an effort whenever he could.

Table 7.6: Results of activities

Activity	First Attempt	Second Attempt	Third Attempt	Failure	Total
Sentence Shuffle	37	6	4	3	50
Key Word	25	1	1	n/a	27
Illustration	11	1	3	13	28
Onset	2	5	7	n/a	14
Blend	4	5	4	n/a	13

Observations. During observed sessions (appendix M, section M.2) Justin showed signs of being easily distracted. He sometimes appeared not to follow the highlighting of words as they were spoken by the computer. He also readily selected unknown words rather than attempting to decode them independently and ignored any hints given. There appeared to be a problem with the software which on several occasions gave more than one hint before the word pronunciation. He found this to be frustrating. When being observed, he frequently turned and made comments unrelated to the reading tasks. Although the activity results indicate a higher than expected failure rate, he did seem to concentrate on the activities during the observed sessions, particularly the sentence shuffle. He read the words out as he reconstructed the sentence correctly. In this activity, he appeared to be concentrating and making an effort.

In the first observation during week three, it seemed that the activities generated a greater interest in the task and a higher level of motivation for Justin. He asked if he could start the activities at the beginning of the session. At the end of the session he expressed an interest in reading another talking book.

In the second observation he obtained hints for two words and successfully decoded them without the whole word feedback, which was considered to be effective use. On this occasion he seemed very pleased with his success. Justin exhibited physical signs of loss of concentration on 25 occasions. He frequently made comments that were unrelated to the task such as:

Shall I go down to assembly?

Where are the window cleaners now?

He made several comments relating to the use of the software. Although he did not often exhibit much behaviour that might be associated with high motivation, such as repeating words, he did comment on aspects of the story. His use of meaning and syntax improved during the observed sessions, although his use of initial letter was not consistent. He appeared confident, but unwilling to make an effort particularly in situations where the computer would do the work for him.

I just press twice on the word and it tells me.

Justin appeared to be motivated at times by the software during observed sessions, particularly when involved in the activities, but this was not consistent and he was easily distracted.

Posttest. At posttest, his reading age had improved in both standardised reading tests, as had his scores on the common 100 words and key word recognition tests (Appendix L). He appeared to be taking more notice of initial letters. His teacher perceived that his use of strategies remained unchanged and she commented that he "quickly realised that he could get the computer to do the work for him! He will try to get anyone to do his work if he can." She felt his confidence had improved a little but his motivation towards reading remained unchanged. His blending skills improved slightly, but he was still unable to decode CVC words (Appendix L). In his interview, he was extremely positive about his experiences with the talking books. He stated that he enjoyed reading a lot and loved reading on the computer. He indicated that he would like to spend more time using the computer than he had done (daily use). He thought that being able to access word pronunciations was useful. He felt that he enjoyed reading more after participating in the study, and felt that his reading had improved "because it [the computer] makes me read quicker".

Summary of case 2

- Justin made effective use of the software sometimes but occasionally experienced lapses in concentration.
- He was sometimes unwilling to make an effort, but he did not seem to become over reliant on immediate word feedback although he readily selected unknown words.

- He did seem to be motivated by the activities, but his performance on these was adversely affected by lapses in concentration and willingness to apply himself to the tasks.
- The use of the software had a positive effect on the learning outcomes. He avoided work at every opportunity, but using the software as a means of providing supplementary practice was still beneficial for him. However, he seemed to prefer adult attention.

7.4.4 Case study 3

Pretest. Mark had already completed stage 3 of the reading scheme and was on stage 4. Selected as lacking in confidence, his teacher felt that he would benefit from re-reading stage 3 texts in an alternative format. Again, he used the enhanced version of the software with both onset and rime feedback, and hints enabled. At pretest, he stated that he enjoyed reading a lot and that it was his favourite subject in school. He had access to a computer at home and was quite comfortable with their use. He said that he did not like reading alone, he enjoyed reading to his parents a little and that reading to his teacher was not bad but that he did not enjoy it. There was little evidence of use of initial letter, meaning, syntax and phonics in the Salford Sentence Reading Test. There was evidence of use of all strategies in the Burt Word Reading Test and the phonic tests. He was approaching the ceiling in the key word recognition test (Appendix L). That is, although he was lacking in confidence, he already recognised most of the key words up to stage 3 of the reading scheme by sight. He had little difficulty with the blending task and decoding CVC words (Appendix L). He was also beginning to develop proficiency in decoding consonant-consonant-vowel-consonant words (e.g. 'trap', 'slap', 'clap') and those beginning with consonant digraphs (ch, th, wh, sh).

Interaction with software. Mark spent a total of five hours and 15 minutes using the software, an average of 15.75 minutes per book. As he made little use of immediate word pronunciations, it was not necessary to consider whether he was reliant on such support. He was given 35 hints and whole word feedback for 14 words. From this it might be inferred that he made use of some of the hints. There was some evidence from the observed sessions of ignoring hints and clicking immediately on the word again in order to obtain its pronunciation. During the observed sessions he only made three miscues and he requested

computer feedback for two of them. He received feedback at the level of onset and rime for a total of five words. He had no problems with the activities and appeared to be motivated and engaged with the texts.

Table 7.7 summarises the activity results and shows that Mark successfully completed most of them, and the total number of activities attempted was well above the average attempted by the enhanced group. He did not have much difficulty with the activities and was able to complete them quickly, enabling him to try more than the average enhanced group member. He achieved a higher degree of success than most of his counterparts, which should have boosted his self-confidence. The evidence here suggests that Mark appeared to make use of the hints and use the software effectively.

Table 7.7: Results of activities

Activity	First Attempt	Second Attempt	Third Attempt	Failure	Total
Sentence Shuffle	61	3	1	4	69
Key Word	36	3	0	n/a	39
Illustration	26	8	2	8	44
Onset	11	2	3	n/a	16
Blend	9	4	0	n/a	13

Observations. During the first video-recorded session (appendix M, section M.3), Mark was pointing to words as he read. He read confidently and fluently throughout the entire book. He only made three miscues, and solved them by alternative strategies, making no use of computer feedback other than at sentence level. His approach to the activities was confident and he was generally successful on the first attempt. He did, however, appear to be a little distracted, making comments that were unrelated to the task and fiddling with the hardware. This may have been partly due to the fact his 15 minute session overlapped with the school playtime and he was aware of this. He was left unattended for a short while and had great fun experimenting with the video recorder.

During the second video-recorded session Mark seemed to be much more focused on the task. He made several comments relating the story itself and enjoyed it greatly. Again he

recognised the miscues he made and decoded them via alternative strategies. He did request one word pronunciation and ignored the hint. He read confidently and fluently, and seemed to enjoy the activities. He did comment that the sentence shuffle and illustration activities were time consuming. He indicated a preference for the onset and rime activities, exclaiming "yes!" when he got one.

Posttest. At posttest, his teacher commented that:

> Mark reads with much more confidence and fluency. He has benefited from reading every day.

His reading age as measured by the Salford Sentence Reading Test increased by three months. His reading age as measured by the Burt Word Reading Test increased by one year and he appeared to make better use of initial letter and phonic information. His scores on the common 100 words and key word recognition tests improved, despite approaching ceiling levels. The results of the phonics tests were similar to those conducted prior to using the software. (See Appendix L for full test results.)

However, Mark's attitude to reading seemed to be less favourable. He did not identify it as being his favourite subject and he indicated that he liked reading a little bit, rather than enjoying it a lot. His confidence in using computers appeared to improve and he still seemed to enjoy using technology. He indicated a preference for reading with the computer, when compared to reading alone, with parents or his teacher. He stated that he would like to continue to use the computer and more frequently, and he perceived that his reading had improved. He commented that he had enjoyed the activities, although he felt they were time consuming. He identified that feedback both for individual words and for complete sentences had been extremely helpful. Mark had already read the texts at Stage 3 before. However, it seems that presenting the texts in an alternative format ensured that his motivation was not adversely affected through repeated exposure to the texts and that his self-confidence improved.

Summary of case 3

- Mark's confidence level was greatly improved through additional practice and the opportunity to experience success.
- He was able to experience success through completing the activities.

- Through receiving sentence level feedback he was aware that he was reading the text accurately, contributing further to improving his self-confidence.
- Although he had already read the texts before, presenting them in an alternative format ensured that he was not demotivated by the repeated exposure.

7.4.5 Case study 4

Pretest. David was identified by his teacher as having a good level of confidence and an average attitude to reading, and was assigned to the basic software group. He was selected as a case study nevertheless because it became apparent, through informal observations, that he was not motivated by the use of computers. He was still on Stage 2 of the reading scheme and was making steady progress. His teacher perceived that he was regularly using initial letters, meaning and syntax, and illustrations to decode unknown words. In the reading tests, he made little use of decoding strategies, readily saying "don't know" rather than making any sort of attempt. He could not decode CVC words and displayed no blending skills. His scores in the common 100 words and key word recognition tests were relatively low, as were his standardised reading test scores (Appendix L).

David stated that he enjoyed reading a lot, but found it hard. This was related to his lack of experience and lower reading ability. He said that he greatly enjoyed reading alone, to his parents and to his teacher. He had exposure to computers in his home environment. Although he indicated that he was not worried about using computers and did not find them hard to operate, he did not like using technology. He commented:

> Sometimes it's [the computer] a bit scary, sometimes it's a bit nice, like a funfair.

Interaction with software. David spent a total of five hours and one minute using the software, an average of 15.05 minutes per session. He requested word pronunciations for 1389 words in total, of which 530 were key vocabulary. Clearly, he was more dependent on computer support than some of the other participants. For 11 of the 20 books read, the requests for feedback were at least 50% lower on the second reading than on the first (Figure 7.33). For example, the first time that David read 'The Big Turnip' he requested 82 words for feedback and the second time he requested only 30 words.

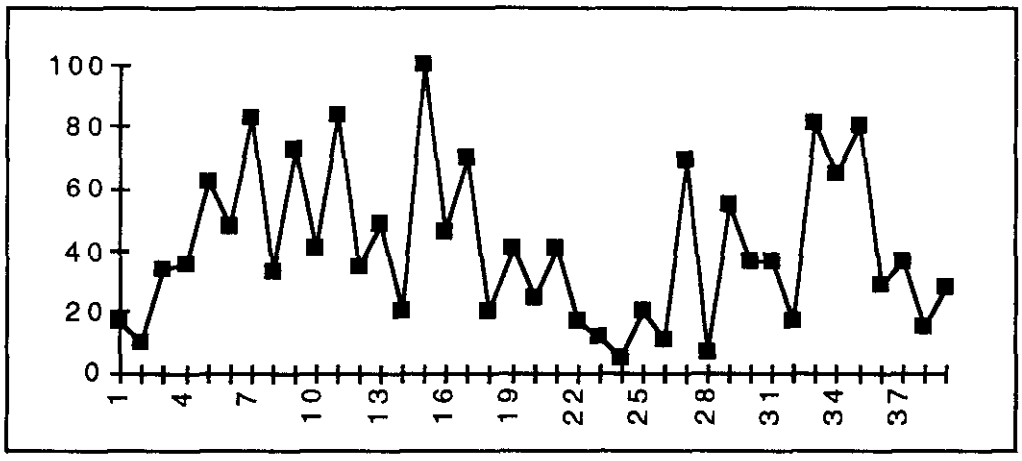


Figure 7.33: Use of instant feedback over 20 sessions

Observations. Due to David's absence, he was only video-recorded at the beginning of the study. During the observed sessions he appeared to make better use of initial letters, meaning and syntax as he progressed through the study. However, he requested a large number of word pronunciations, frequently clicking on every word in a sentence and then repeating them. He rarely attempted to decode an unknown word independently. It was evident that some of the words selected were already known because he was selecting words such as 'the' which he had correctly identified at both pretest and posttest. He sometimes used the computer feedback to check that he had guessed correctly. He occasionally made errors that he was unaware of, such as reading 'tripped' for 'fell'.

David made a lot of use of instant feedback during the observed session, although he tended to repeat the word after the computer pronunciation. This may be considered to be motivated behaviour. As he required a greater level of support he may have opted for instant feedback in order to maintain comprehension levels. Without computer assistance he would have been reading a text at a frustration level. He was seen to make use of the illustration and the initial letter for one or two words. The first book was a repeat read, during which he appeared to be concentrating and quite focused. However, during the second book, which was previously unseen, he seemed to lose concentration and made several comments that were unrelated to the task. He commented: "its a bit boring!" It may have been the case that the books were too difficult for him. He did seem to be reliant on computer generated word pronunciations and for him, using the computer was not motivating.

Posttest. At posttest his teacher commented:

> David does not seem to have benefited greatly from this. He still doesn't seem interested in reading and now won't try any words he doesn't know.

His reading age as measured by the Salford Sentence Reading Test improved by two months. His reading age as measured by the Burt Word Reading Test improved by three months. He made some improvement on the number of key words which he recognised by sight, but not in the common 100 words test. His phonics scores remained unchanged. (See Appendix L for full test results.)

In his interview David stated that he enjoyed reading a lot. However, he indicated that he did not enjoy using the computer. Not only did he find it boring, but on occasions he was worried about using the technology. When asked if it was the stories themselves that he found uninteresting or being on the computer, he readily replied "'cos I'm on the 'puter". He indicated a preference for reading alone. He said, though, that he would like to continue to use the computer to practise reading. He perceived that his reading had improved because the computer had helped him to learn new words and helped him to concentrate.

Summary of case 4

- David did not enjoy using the computer, although he commented in his posttest interview that he enjoyed reading.
- David soon developed a high level of reliance on immediate feedback. For him, this style of behaviour resulted in the use of computerised texts affording minimal, if any, challenge. This led in turn to decreased motivation and a more negative attitude to reading. As he was identified by his teacher as not quite ready to read stage 3 texts independently, it is likely that the enhanced software would have had similar effects.

7.4.6 Case study 5

Pretest. Rifat was a girl with English as a second language. Her teacher perceived her as being highly motivated, but lacking in confidence:

> [She] is very quiet and she is making good progress for an E2L [English as second language] child. She is very keen to read and motivated to make good progress.

In her interview Rifat said that she enjoyed reading a lot and that she perceived it as being easy. She also identified reading as her favourite subject. She had exposure to a computer at home as well as at school, and stated that she enjoyed using such technology. She indicated that she felt the same about reading alone, to her parents and her teacher: enjoyed it a lot. She

appeared to make more use of initial letters, word endings and their overall shape, than of meaning and syntax. She already recognised half the common 100 words and most of the key words up to Stage 3 (Appendix L). She experienced some difficulty in decoding CVC words and had limited blending skills.

Interaction with software. Using the basic software, Rifat requested a total of 257 word pronunciations of which 36 were key vocabulary. She did not appear to be reliant on immediate feedback (Figure 7.34) and made effective use of this facility. She spent a total of four hours and 50 minutes using the software, an average of 14.5 minutes per session.

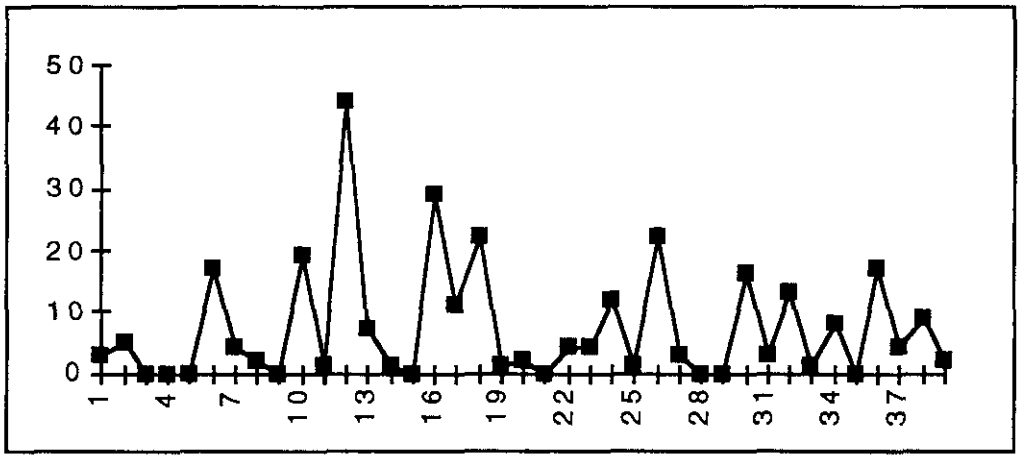


Figure 7.34: Use of instant feedback over 20 sessions

Observations. Observation data indicated that she frequently made use of initial letter, meaning and syntax cues, as well as illustration cues where appropriate. However, their use was not consistent. In the video-recorded sessions (Appendix M, section M.5) Rifat was focused and read with expression. She rarely hesitated or showed any physical signs of having low self-confidence. She only required computer feedback for four words, which she accessed instantly. She miscued on ten occasions and seemed to be unaware of them. As English was her second language, it is likely that she had less experience of syntactical rules and would therefore be more likely to make such mistakes in reading texts. Overall, she was motivated and made effective use of the software.

Posttest. At posttest her teacher perceived that her confidence level had improved:

> Rifat has been very excited about this study and this has really motivated her reading and positively affected her writing. She has raced through her reading stage confidently.

Her reading age as measured by the Salford Sentence Reading Test improved by four months. Her reading age as measured by the Burt Word Reading Test improved by three months. Her results in the common 100 words and key word recognition tests improved slightly. Decoding CVC words improved, as did her blending skills. (See Appendix L for full test results.)

In Rifat's interview, she indicated that she still enjoyed reading a lot, although she no longer identified reading as her favourite subject. She still enjoyed using the computer. She stated that she enjoyed reading alone a little bit, and greatly enjoyed reading to her parents, her teacher and reading with the computer. She expressed a preference for reading both to her teacher and with the computer. She did not perceive that there were any problems relating to reading with the computer. She thought it was useful because it was easy to use and it talked. She expressed a desire to continue using the computer. She felt that her reading had not improved, but had remained the same, yet her teacher had perceived that her self-confidence had improved.

Summary of case 5

- Rifat was progressing well. She was extremely motivated and thus enjoyed her experiences with the computerised texts.
- The opportunity for regular practice greatly improved her self-confidence.
- She made gains in learning outcomes, although they were related to an increased sight vocabulary rather than an improved use of reading cues.
- The software was beneficial in its basic format. However, providing the enhanced version may have provided more challenge and more opportunities for experiencing success.

Overall summary of case studies

The evidence arising from the case studies indicates that the use of computers to support reading instruction is highly complex. Furthermore, it suggests that it is not always the case that computer-based learning is motivating, even for the boys who are commonly believed to be more interested than girls in technology. Only one of the two boys who had been identified as lacking in confidence appeared to use the software effectively and improved in

both the cognitive and affective domains. The other boy's level of self-confidence remained unchanged. Moreover, his motivation decreased with continued use of the software despite being in the enhanced group and having additional features such as the activities that were designed to enhance motivation.

Two of the boys, one from the enhanced software group and one from the basic software group, who were identified as lacking in motivation, did not benefit greatly from their experiences. They both appeared to avoid making an effort, although one was clearly not motivated by computer use rather than the reading itself. The girl who was selected as lacking in self-confidence greatly enjoyed her experience with the basic software version. Therefore, it seems that a self-motivated child will benefit from the use of talking book software, whatever its format.

7.5 Discussion

The main findings of the study are now discussed in relation to the aims of the study presented in the introduction to this chapter.

Overall, it is considered that both versions of the software were beneficial and could make a valuable contribution to primary classrooms. In addition, the enhanced software seemed to offer the most effective environment for children who have already acquired a limited sight vocabulary. A number of issues arising from the study are highlighted. In particular, it appears that matching the reading ability level of an individual to the type of features provided is of paramount importance. There are also issues that concern the appropriateness of some aspects of the methodology used to conduct the evaluation.

The mean gains in the standardised reading tests, which were achieved by both groups, are greater than might be expected during the elapsed time of one month. The children using the enhanced software achieved mean gains of three months in the Salford Sentence Reading Test and six months in the Burt Word Reading Test. The children who used the basic version of the software also achieved mean gains of three months for the Salford test and four months for the Burt test. As the study concerned the comparison of two software versions it is not possible to contrast the results with a non-intervention control group.

Nevertheless, as standardised reading tests were used it could be argued that either format of the resource is effective in providing additional support for practising reading. It seems that for some children, practising reading texts without additional features, such as reinforcement activities, is sufficient to lead to gains in reading ability. These children may be well-motivated and already be making progress with little support in the classroom.

7.5.1 The software and the classroom environment

The software met closely with the pedagogical criteria specified in the framework for evaluation. It complemented classroom practice and was well received by the teaching staff. In common with the findings of Matthew (1997), the use of electronic texts appeared to stimulate an interest in reading for some of the children. Favourable comments were received from the teaching staff regarding the provision of a hint prior to obtaining the whole word pronunciation and activities, features that were available in the enhanced version. However, one teacher commented that the onset and blend activities (chapter 6, section 6.5.3) were more challenging for some of the children than the other activities. She considered that they may not have been appropriate for the lower reading ability children.

Table 7.8: Percentage of success at each activity for the children using the enhanced talking book software

Activity	First attempt	Second attempt	Third attempt	Failures
Sentence shuffle	82%	6%	3%	9%
Key word	90%	7%	3%	N/A
Illustration	59%	12%	4%	25%
Onset	53%	25%	22%	N/A
Blend	49%	21%	29%	N/A

The software interaction data was examined to determine whether the onset and blend activities were more challenging. Table 7.8 indicates the percentages of success at the first, second and third attempt for each activity for the group of children using the enhanced software. The illustration activity proved to be difficult for all of the children. This was due to the poor quality of the picture generated by the software, rather than the cognitive demands of the task. Furthermore, in this activity, some of the books contained bugs. This

meant that the children were told that they had not completed the task successfully, even though they had. This partially accounts for the high failure rate. It was anticipated that the sentence shuffle activity (chapter 6, section 6.5.3) would be the most demanding task and subject to the highest failure rate. This was not found to be the case. If children found this activity too difficult (e.g. not remembering the sentence from the page they invoked the activity from), they used the hard copy book in order to complete it successfully. The interaction data seems to confirm the assertion from teacher 2 that the onset and blend activities were more difficult. However, the interaction data from the software logs highlights a flaw in the design of the activity. Sometimes it was possible to create words, other than the target word, that were either real or sounded correct (e.g. 'jour' sounds like 'jaw'). It should be noted that this could be influenced by local dialects.

Table 7.9: Recorded attempts at onset and rime activities for one case

Target Word	Rime given	Onsets available	Attempts
sausages	ausages	d s c	(c) (s)
dad	ad	d s c	(s) (d)
four	our	p f j	(p) (f)
jaw	aw	p f j	(p) (f) (j)
jack	ack	p f j	(p) (f) (j)
crack	ack	br cr ch	(ch) (br) (cr)
tail	ail	b c t	(b) (t)

For example, in one case [Clare], the total number of onset and blend activities attempted was 19, of which 12 were successful on the first attempt. The remaining attempts are recorded in Table 7.9. Of the 10 incorrect onset selections, five represented real words that were not the target words (sad, pour, paw, pack, bail) but the child was instructed to try again.

Signs of frustration were also observed in another case [Sally]. This child clearly recognised that the onset she selected made a real word ('bed'), yet she was instructed to try again as the target word in this example was 'Ted'. Furthermore, one child [Katie] who was a less experienced reader was observed to sound out the onset and rime, and make the correct

selection. It appeared that she understood what was required despite the fact that she was lacking in experience. The high failure rate could be attributed to the ambiguous design of the activity rather than the difficulty of the task.

The prediction activity (chapter 6, section 6.5.3) was disabled early on in the study. It appeared to be too difficult for children of this age range who had not yet acquired typing skills. It may be that this kind of activity is more appropriate for older children who may be more adept at typing. Alternatively, the software could be adapted to record an oral explanation rather than receiving a type written prediction.

Providing activities in Talking Books has the potential to extend the period of concentration that a child may have when reading. Thus, it may increase the amount of time that a child uses the software for practising reading. The issue of when and where to provide activities must be considered in future implementations. Activities should support the development of reading strategies and not form a distraction or interfere with comprehension.

One of the weaknesses of existing commercial software identified in the pilot study (chapter 4) was that children frequently made mistakes of which they were unaware. Without speech recognition, this problem is difficult to address. For this study, the children were instructed to request sentence level feedback after attempting to read the text themselves. This was really effective and the children themselves made many comments relating to the benefits of such a feature. They were able to verify that they had read the text correctly or were made aware of any errors that had been made. By confirming their attempts at decoding they were able to experience success which in turn could contribute to motivational and confidence levels. As Parham (1993) commented, this form of feedback provides a good model against which a child can evaluate their mistakes. Moreover, by instructing all children to use the software in this way, it was possible to ensure that children were made aware of any errors they made. This was particularly important for those who might not request help otherwise.

7.5.2 The effectiveness of additional features

It is considered that the complete system should be assessed rather than the individual features. The original design was intended to emulate a human tutor, adapting to provide

tailored instruction to meet individual needs and requiring a number of additional interrelated features that were introduced simultaneously. Evaluating a specific feature in isolation is problematic and the findings cannot be attributed to one feature alone, but have to be considered in relation to the whole system. However, as the features were each designed to address a particular skill or an undesirable style of interaction, the merit of each is now considered.

Providing coaching in the use of specific reading cues may positively affect word attack skills

There was no evidence to suggest that providing coaching in the use of specific reading cues through activities and hints positively affected word attack skills. The gains made in the standardised reading tests and common 100 words test were not significantly different between software groups. The variances in the gains achieved were high which could partially account for this.

The teachers perceived that the use of meaning and syntax cues improved more for the enhanced software users than the basic software users. However, miscue analysis detected no differences between the two groups in reading strategy and cue use, for which there are two possible explanations. Firstly, the study took place over four weeks. The development of the use of reading cues may take place over a number of months. Thus, it may not be possible to detect such subtle changes in reading behaviour over a short space of time. Secondly, the results of the study may have been confounded by the classroom instruction that took place simultaneously. Both the enhanced and the basic software users were receiving additional instruction throughout the intervention period and being coached by their teachers on the appropriate use of reading cues and strategies.

Providing an activity to reinforce sight recognition of the key words could positively affect the development of key word vocabulary

With the medium reading ability children, the gains made in sight acquisition of key vocabulary were significantly greater for the enhanced software users than the basic software users. It is not possible to attribute this to one feature alone as they were introduced as a whole. It is possible that the key word activity and the sentence shuffling activity contributed

to this effect (see chapter 6, section 6.5.3 for activity interaction examples). Both the high reading ability groups were subject to a test ceiling effect.

The low reading ability children made the highest gains in key word recognition irrespective of the software version used. It seems that for the lowest reading ability children, exposure to the key vocabulary within the texts and the computer support of word pronunciations is the most beneficial aspect. In further support, there was a significant positive correlation between the gains made in the key word recognition test and the number of word pronunciations requested. The activities may have had little affect on the group of children.

Giving a hint could encourage independent word identification and prevent a reliance upon instant word pronunciations

The hint mechanism did not operate as anticipated and after the novelty effect had worn off, most children ignored them completely. It may be that certain human tutoring techniques, such as this, demand more sophisticated techniques than those implemented in this version of CARIS. A human tutor would not necessarily give a word pronunciation instantly if they knew that a child was capable of decoding it themselves. Rather, they would draw upon a wide range of information, such as non-verbal communication, in order to decide whether or not such action was appropriate. Another possible explanation is that the hints available were all general, although it was intended that they be generated according to the type of word selected. For example, if the word selected was a noun then the child was instructed to look at the picture. It may be that the implementation of the hint mechanism was too simplistic. There is clearly room to improve the sensitivity and range of hints that could be given. Furthermore, giving a hint at every request for a pronunciation may not be beneficial. It is possible, that varying the feedback given between a hint and the whole word may improve the outcomes of the interaction. In addition, giving a hint for some words may never be beneficial. As a common teaching strategy is to give a hint, this type of approach warrants further investigation. However, in order to be effective it may demand the implementation of more sophisticated techniques commonly associated with intelligent tutoring systems such as student modelling and tutoring protocols (Hume, Michael, Rovick & Evens, 1996).

It was assumed that the lower reading ability children would be more reliant on word pronunciations than their peers and so the data was analysed by reading ability group (Figure 7.35). The lower reading ability children accessed significantly more word pronunciations than the higher reading ability children ($F(2,30) = 5.595$, $p = 0.009$) whether they used the enhanced or basic version of the software.

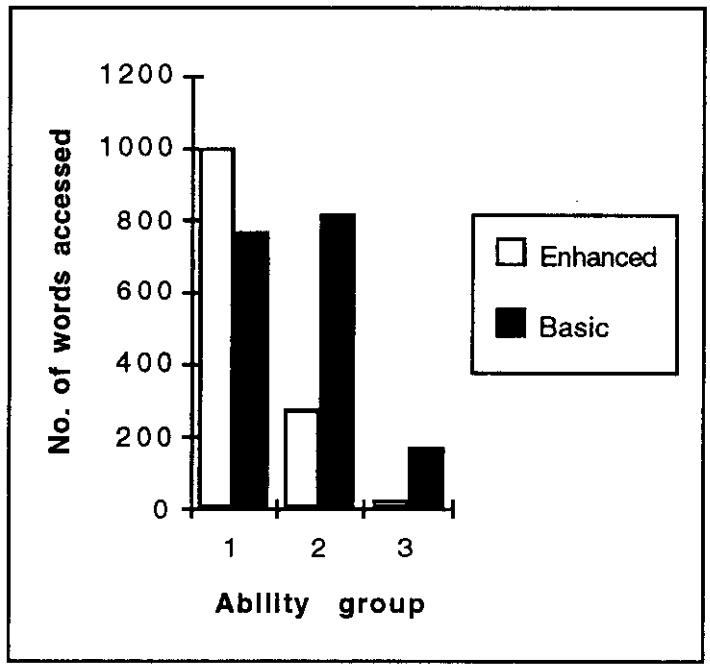


Figure 7.35: Mean number of words selected for enhanced and basic software groups by reading ability

There was no significant difference between the mean word pronunciations accessed by low reading ability basic and low reading ability enhanced software users. The medium and higher reading ability children who used the basic software accessed significantly more word pronunciations than those children of similar abilities who used the enhanced software ($F(1,23) = 4.341$, $p = 0.049$). One explanation for this is that the basic software users perceived the software to be fun and felt more inclined to make use of this feature. In contrast, the enhanced group may have perceived the software to be an educational task and made more effort to decode unknown words independently. An alternative explanation is that a side effect of providing hints may have been to act as a deterrent against unnecessary feedback selections. Some of the basic users may have selected more feedback in order to cross-check a guess. Alternatively, they may have accessed a word pronunciation without

attempting to decode it independently because it was easy to do so. For the enhanced users, receiving hints was irritating and time consuming, and affected fluency and comprehension. They are more likely to have opted for help as a last resort. This is in accordance with the findings of van Daal and Reitsma (1990) who found that poor readers made fewer requests for help when segmented feedback was provided than when whole word pronunciations were given. They too considered that for these children the segmented feedback had acted as a form of deterrent.

Three of the eight children in the medium reading ability range in the enhanced group did not have hints enabled. Discounting their data, the average number of hints given was 84 and the average number of word pronunciations obtained was 85. This implies that the children for whom hints were mandatory generally accessed the word again in order to obtain its pronunciation. All four of the high reading ability enhanced software users had hints enabled. They requested an average of 43 hints and 28 word pronunciations. It is likely that these children made effective use of some of the hints that were received.

Providing feedback at the level of onset and rime could positively affect the development of phonological skills

It was anticipated that the group using the enhanced version would benefit from the phonological training, and that the gains on tests measuring phonic ability would be greater than the basic group. However, no differences were found. It is possible that the sensitivity of these measures is low or inappropriate. They were designed to focus on phonic analysis and synthesis of words at the level of phonemes, rather than at the level of onset and rime. This was not detected in the pilot work because the talking books in that evaluation did not provide segmented feedback and so changes in the test results were not expected.

The provision of feedback at the level of onset and rime did not appear to be effective. It could, though, have contributed to the gains made in non-practised words in common with the findings of previous related research (van Daal & Reitsma, 1990; Wise et al., 1989). The implementation of this feature was simplified by giving the starting sound and remaining sound of a whole word rather than for each syllable. Thus, some of the multisyllabic words selected in these books may not have been suitable for developing an awareness of word

analogies or word families. Words such as 'downstairs', 'sausages', and 'vegetables' were given rime feedback of 'ownstairs', 'ausages' and 'egetables'. There is little doubt that these endings do not form common word families and therefore would not be likely to enhance a child's use of word analogies. Of the total of 1218 words for which onset and rime feedback was requested by the users, only 601 (49%) could be considered to be from common word families.

It should be noted that 12 of the 20 books contained feedback at the level of onset and rime, and so feedback was only available in 40% of the books read. Also, only 11 of the 16 users of the enhanced version were identified by their teachers as being suitable to receive this type of feedback. In the study, the children had limited exposure to this form of feedback.

7.5.3 The effects of talking books on motivation and self-confidence

The evidence relating to the affective aspects of the framework for evaluation and in particular that relating to motivation seems to be inconsistent. The teachers perceived that the enhanced group became more motivated towards reading than the basic group as a result of using the software. Yet the children who used the enhanced software, although expressing a more positive attitude to computers after the study, developed a more negative attitude to reading. The opinions expressed by the children themselves seem to contradict the improvements to motivation levels as perceived by the teaching staff.

One reason for these findings might be that the enhanced software group perceived the activities as being hard work and requiring additional effort. Yet, the basic users may have perceived the talking books to be more fun because they were able to read with the additional support of computer feedback, which required less effort. This is further supported by the findings, which indicate that the basic users made more use of the computer support than did the enhanced users.

With regard to changes in self-confidence, the teachers perceived that both groups improved. The children who used the enhanced software, however, appeared to find reading easier after the intervention period than those using the basic software. In addition, more of the enhanced group perceived that their reading ability had improved than in the basic group.

One explanation is that the enhanced group experienced a degree of success with the activities that they attempted, which improved their levels of self-confidence. In contrast, the basic users had fewer opportunities to experience success whilst using the software.

Of the three case studies examined who used the enhanced software, only one pupil was positively affected by his experience. In the second case, for a child who was lacking in confidence, there was no improvement in this affective aspect. His teacher commented that following the study he had to be prompted to use some of the reading cues again. In the final case, using the computer was not necessarily effective for the child who demanded a great deal of adult attention. It was anticipated that he would be motivated by the computer-based resource. He still managed to avoid making an effort however and it was still necessary for his teacher to keep a close eye on him.

Of the two case studies selected from the basic software group, one pupil benefited greatly and showed improvements in self-confidence but was highly motivated, whereas the second pupil seemed to have been negatively affected. In this case, his teacher commented that he was still uninterested in reading and at the end of the study he was no longer attempting to decode unknown words.

Prior to the study it was anticipated that the provision of activities and multiple feedback modes would enhance the affective aspects of software use, making it more fun and interactive. However, this did not hold in all cases. Evidence from case studies supports the statement made by Balajthy (1989) that computer-based resources for teaching reading do not provide a "magical motivational cure". Children who enjoyed using computers or were self-motivated prior to the study made effective use of the software, irrespective of which version they used. Those enhanced software users who were lacking in motivation were not necessarily inspired by the additional features of the talking books. Rather, in some cases, they perceived the additional features, such as the activities, to be academic tasks and requiring effort.

7.5.4 Technical issues

A number of technical issues have already been discussed that concern the provision of hints, sentence level feedback, activities and feedback at the level of onset and rime. Issues that relate to a number of additional features are now presented.

The modelling mechanism (chapter 5, section 5.3.5) was not sensitive enough for those children who were reliant upon word pronunciations and accessed more than 10% of the vocabulary in the talking books. This had been anticipated in the design. An attempt had been made to address it by only adjusting the models when words that had not been previously selected in the book were requested for feedback. For example, if the word 'there' was selected on five pages of a book then the models would only be adjusted on the first request. However, the children who were more dependent on computer support selected a large number of different words in each computerised text (more than 10%). The outcome of this was that the models were all reduced to their minimum value and the software recorded under-utilisation for all cues for these children and did not provide tailored instruction. This problem could be addressed by setting individual thresholds for the models in order to complement the variety of levels of support required by each user.

It is not possible to comment in depth on the adaptive features of the design due to a bug that remained undetected until more extensive use occurred in the full evaluation of the software. Some elements of adaptivity functioned correctly such as being able to configure the feedback to support individual requirements. The under-utilised reading cues and strategies, although correctly modelled, were not addressed directly. Rather, each child received coaching for all reading cues in equal proportions.

7.5.5 The limitations of the empirical design

A combination of research methodologies were employed to address issues of generalisability and to highlight the effects of such a resource on individuals. The quantitative analysis did not reveal overall statistical differences between the learning outcomes of the children who used the enhanced and basic versions of the software. Specific differences were detected in the medium reading ability range. There are a number of factors

that may have influenced the non-significant results overall and limit the inferences that can be drawn from this study.

The number of participants was small and there were substantial variances in their reading abilities. This had been anticipated and an attempt at addressing this issue had been made. It was hoped that the teaching staff would have little difficulty in identifying children who were ready to read the books because the software was based on a specific stage of the reading scheme. Davidson (1993) identified similar problems in her research design, but attributed it more to the differing judgements of individual teachers. As she used a quasi-experimental approach in her research, the groups for comparison were formed from separate classes and so there was a greater degree of variance between them. It was anticipated that, by selecting matched pairs from the same class, these variances would still occur but would be evenly distributed and be less problematic. However, it was evident that many of the children were not yet ready to read the computerised texts independently. Many of the teachers commented that they were unable to identify enough children of the right reading ability. This problem was more extensive than had been anticipated. They had to include some who were less able in order to make up the numbers and so the variance remained large. As the problem was anticipated, it seems appropriate that part of the analysis was conducted on reading ability ranges, although this did further reduce the number of participants within each grouping.

It seems that, while there are only a limited number of computerised texts available, the problem of matching texts to each reader will persist.

The empirical work was intended to be naturalistic, although in one school the research was conducted in a separate room. It could also be argued that the software was used more regularly than it might have been if the staff were integrating the software within their classes out of choice. This was because the frequency of use was specified and the teachers agreed to allow the children to have daily access to the computer. This was necessary though because it was imperative to control the amount of time spent using the computers in order to compare the two groups. The amount of time spent practising reading can have an affect on improvements that might be made. Recent research (Davidson, 1993; Medwell, 1996), in

which the teachers have been responsible for arranging access to the computer, could be criticised for not allowing equivalent amounts of practice between participants. A number of reading programmes, such as Reading Recovery (Clay, 1993b), specify that children must spend regular amounts of time on a daily basis in order to benefit from their participation. Therefore, it was felt that it was realistic to impose similar restrictions in this study. Due to the time taken to complete activities there was a difference in the overall time spent using the software between the two groups. This may have influenced the results obtained.

This study has compared the use of two versions of talking book software without considering a non-intervention control group. It was not relevant to compare the use of the software with a group who did not participate in equivalent non computer-based tasks and hence were not engaged in similar amounts of practice. Moreover, it is not necessarily valid to compare such diverse resources as computer assisted learning with pen and paper tasks or even working with an adult. These alternative instructional methods offer different advantages and features. It is more appropriate to consider the individual merits of a particular resource.

It is difficult to draw general conclusions from these results due to the reliance on gains analyses, which can be problematic, and to the limited sample size used in this study. There was a greater variance in the reading abilities of the children in the study than had been anticipated and this may have affected the results obtained. There was no assessment of the sustainability of the results obtained. It may be that the novelty effects of computer based learning could have contributed to the results that were obtained and that delayed post-testing would not provide evidence that the progress made by the participants was sustained.

7.6 Conclusions

- Talking books complement classroom practice, although teachers would not always provide a word pronunciation instantly.
- Both versions were beneficial overall, but the medium reading ability children who used the enhanced version made significantly greater gains in the sight recognition of key vocabulary than their counterparts who used the basic version.
- Both versions of the software may be effective in developing the sight acquisition of key vocabulary for those children who are less experienced readers.

- *Instructing users to select sentence level feedback after attempting to read the text* themselves may overcome the problem of children not being aware of making mistakes.
- The hint mechanism was not successful in its current implementation. As it is a common teaching strategy and the concept of embedding such a technique within educational software was well received by practitioners, it warrants further investigation.
- Medium and higher reading ability basic software users accessed significantly more word pronunciations than the children with similar levels of reading ability who used the enhanced software.
- Lower reading ability children were significantly more reliant on word pronunciations than higher reading ability children whether they used the enhanced or the basic software version.
- The provision of activities has the potential to extend the period of concentration of a user and hence the amount of time spent practising reading skills.
- The way in which feedback at the level of onset and rime was provided was too simplistic and may not have been beneficial. However, previous research suggests that such feedback can improve the decoding of non-practised words (Wise et al., 1989; Goswami & Bryant, 1990). Therefore, this form of feedback should not be discounted in future implementations.
- It appeared that the activities were perceived by some as tasks rather than an element of fun, and thus may not have enhanced motivation in the way that had been anticipated.
- It is imperative to consider the reading ability of an individual when selecting appropriate software. It may be that:
  - children who can read less than 90% of the vocabulary and are more dependent on word pronunciations could benefit more from a basic version
  - children who are able to read at least 90% of the vocabulary should be provided with simple activities and feedback at the level of onset and rime
  - children who can read more than 95% of the vocabulary could benefit from more challenging activities and a greater variety of feedback modes such as onset and rime and the provision of hints

This study has shown that both versions of the software were beneficial and the resource was positively received by teachers and children alike. The basic version was successful in improving sight recognition of key vocabulary for children who may not have been able to read the texts independently otherwise. The enhanced electronic books were effective for those children who had already acquired a limited sight vocabulary and these users achieved significantly greater gains in key word recognition than the basic software users of similar reading ages. These findings are considered in the next chapter in the context of the project overall and the contributions that this research has made to the field.

8. Conclusions

The conclusions of the thesis are now presented. Firstly, a summary of contributions of the research is described. This section of the chapter ends with a list of implications for teachers, designers and researchers. Secondly, the limitations of this research are summarised. These include limitations relating to the empirical study itself and, in addition, a discussion of the research questions that remain unanswered in this study. Thirdly, future research directions are discussed. These relate to software design modifications identified as a result of the evaluation process and alternative populations with whom the software could be evaluated. This chapter ends with a brief summary.

8.1 Contributions to the field

This research has provided an exploratory study of using talking book software in the classroom and investigated the effectiveness of specific design features. In order to become a competent reader a child needs to acquire a wide variety of skills and strategies, which in the past have been taught in isolation. Recently, there has been a move towards a more holistic approach in which all aspects of literacy are taught in the context of real and meaningful text (e.g. described in Harrison, 1992). Moreover, despite the importance given to the role of phonics in the National Literacy Project, it should not be taught in isolation from other reading tasks. Rather, although an essential part of teaching reading, phonics instruction is more meaningful if it is contextualised.

Talking book software replicates texts on the computer, but with the additional feature of sound to provide speech feedback for words and sentences, and has the potential to complement a holistic approach. Talking book software can also be designed to provide contextualised phonics instruction. It is imperative to establish the most effective format for computerised texts as commercial software developers continue to expand the range of titles available and practitioners begin to integrate such resources in their classrooms. Furthermore, an understanding of how such a resource might be integrated and which aspects of the resource are successful (or not) is best investigated in the context of real classrooms.

Previous research has suggested that the use of talking books can be beneficial for early readers (Reitsma, 1988; Davidson, 1993). The main focus of these studies, however, has been the effects of whole word feedback on sight vocabularies. Additionally some researchers (e.g. Wise et al., 1989) have focused on those children experiencing difficulties in learning to read and have argued that the use of software can have positive effects. Some researchers have considered segmented feedback and whether or not it improves the independent decoding of both previously unseen words and pseudowords. Little account has been taken of how closely such a resource complements current classroom practice, its effects on motivation and self-confidence and how future implementations could be improved. Moreover, little research on electronic books has been conducted in British classroom environments. These aspects are important and this research has addressed these areas.

The pilot study (chapter 4) identified strengths and weaknesses of commercial talking book designs. In addition, it provided evidence that regular use of talking book software by children can contribute to the improvement of sight vocabularies. Furthermore, the use of such a resource can positively affect both attitudes towards and self-confidence in reading. Each child in this study used the talking book software for at least five hours over a period of one month. They were observed on a daily basis and the study provided a substantial insight into the use of such a resource. The framework for evaluation (chapter 3), which was developed to ensure that the investigation of computer assisted software was comprehensive, proved to be effective (Lewin, 1997a).

The pilot study also highlighted four styles of interaction that should be considered in the design of talking book software. McKenna et al. (1996) and Miller et al. (1994) described similar styles of user interaction in their research on talking books.

Firstly, children were observed (by the researcher) to use the software effectively by attempting to decode unknown vocabulary independently and obtaining computer generated word pronunciations when alternative strategies, such as looking at the initial letter, had failed. The children were also seen to use word pronunciations in order to confirm that their

guess was correct. The second style of interaction concerned the underuse of the speech feedback facilities. In this case, children were observed to make little use of word pronunciations and make errors, although usually they substituted a different word with the same meaning.

The third and fourth styles of interaction were both related to the overuse of word pronunciations. Children with minimal sight vocabularies were clearly reliant on a greater level of computer support. They were able to experience success and enjoy the stories without additional adult support. Of the children who were able to read the talking books independently (90-94% of words recognised by sight), one child accessed word pronunciations immediately rather than using reading cues to identify unknown words. This style of interaction was also observed in the other children with this level of reading ability, but to a lesser extent.

A practitioner survey was conducted in order to ensure that the design for enhanced talking book software would complement classroom practice (Lewin, in press-a). This highlighted the extent to which such software was being used in classrooms and how talking books have been integrated within the curriculum. Furthermore, the respondents confirmed that incorporating both segmented pronunciations and hints to encourage independent decoding of unknown vocabulary were perceived to be potentially beneficial features. Educational researchers have argued that providing segmented feedback can be effective (e.g. Wise et al., 1989) and that encouraging independent word identification is pedagogically sound (e.g. Clay, 1993a). The results of the survey also highlighted the need to provide activities to reinforce the strategies and skills required to become a proficient reader.

When this research commenced, the availability of commercial British talking books was extremely limited and they had restricted functionality. An innovative design, which incorporated the strengths of commercial talking books together with a number of additional features to address the identified weaknesses, was presented (chapter 5). The implementation of this software design, CARIS, generated 20 talking books based on a popular UK reading scheme, Ginn All Aboard (Fawcett, 1994). This enabled the

effectiveness of the additional features to be investigated, whilst ensuring that the books on which the software was based were appropriate for the target age range and the schools involved.

The main empirical study comprised an evaluation of CARIS, comparing the use of two software versions in the natural environment of the classroom. Sixteen children from three different schools used the enhanced version of CARIS, which included additional features such as:

- segmented word pronunciations
- hints to encourage independent word identification
- a number of activities designed to develop the use of specific reading cues

An additional sixteen children, from the same three schools used a basic version of the software that was intended to replicate the functionality commonly available in commercial talking books. Both user groups used the software every day for a period of four weeks. On average each child used the software for four hours, again being considerably greater than many CAL evaluations that do not consider repeated exposure to the software.

The overall results suggest that both versions of the software were beneficial, complemented current teaching practices and were successfully integrated within the classroom environment (Lewin, 1997b). The activities were well received and moreover, had the potential to extend periods of concentration and the amount of time spent practising reading skills. Surprisingly, the motivational levels of the enhanced software users were not positively affected by the additional features as was anticipated. Rather, they perceived the features, such as the activities, to be educational tasks. Furthermore, although it had been considered that engaging with both versions of the software would be motivating for the users, this was not the case for all the participants, notably some of the boys. Use of both versions of the software had positive affects on the participant's levels of self-confidence, although again not in all cases.

One of the aims of the study was to investigate the effectiveness of the additional features in the enhanced implementation. It was assumed that use of the enhanced software version

would positively affect a child's ability to identify unknown vocabulary using a variety of reading cues including those relating to phonics. This was not found to be the case as the reading test scores of both groups of users were similarly improved. The effects of some of the additional features may have been too subtle to detect in a short period of time and using only one stage of a reading scheme. The effect of regular practice may have also contributed to the lack of difference between the two groups of users in this case. Some children develop their use of reading strategies and skills simply through exposure to books with little, if any, direct instruction (e.g. Rayner & Pollatsek, 1989).

The evidence from this research suggests that use of the enhanced talking book software significantly affected gains made in the sight acquisition of key vocabulary when compared to the use of the basic version. However, it should be noted that the sample size was limited and that the participants' reading abilities at the start of the study varied more than was expected. Furthermore, gains analyses can be problematic. No assessment of the sustainability of these effects was made.

One style of interaction that was observed during the pilot study was a dependency on immediate access of word pronunciations. It was anticipated that emulating a human tutor and providing hints to encourage the independent decoding of unknown words would address this, but it was not effective. Human tutoring techniques such as this may as yet be too complex to implement successfully, although they may be appropriate for higher reading ability children. This undesirable style of interaction was still observed during the main empirical study and alternative methods to encourage independence must be investigated in future designs.

One of the most important issues arising from this study is the relationship between the reading ability level of a child, in terms of sight vocabularies, and the appropriateness of specific features. The use of both versions of the software was beneficial for children of lower reading ability who may be dependent on word pronunciations, contributing to the development of key word sight vocabularies. For those children who have already acquired a limited sight vocabulary, the enhanced version of the software appeared to be more

beneficial. It could be argued that the activities contributed to this effect because they were designed to provide coaching in the use of specific reading cues and reinforce the sight acquisition of key vocabulary. Furthermore, they had the potential to extend periods of concentration and were challenging for these users.

This finding provided evidence to support the work of Vygotsky (1978) and Wood et al (1976) concerning tutoring techniques that provide scaffolding and enable a learner to progress when they may have been unable to proceed with no assistance. In this case, the computer software has provided the support. Word pronunciations enabled the lower reading ability children to successfully read the books. The hints facilitated independent decoding of words for the higher reading ability participants. Varying levels of support were provided within the activities, from word pronunciations to modelling the correct solution. Additionally, instructing the users to listen to the sentence level feedback after reading the text for themselves provided a demonstration of the correct pronunciation, one of the techniques proposed by Wood and his colleagues. The participants were able to select as much or as little support as they required and were thus able to read the text successfully and complete the reinforcement activities.

Implications for teachers:

1) Talking book software provides an appropriate format for computerised reading instruction.

2) The regular use of computerised texts can contribute to gains in sight vocabularies and, with additional features, may encourage the use of cues for independent decoding.

3) Supporting activities have the potential to increase the periods of concentration that might be sustained by a child and hence the amount of time spent practising reading. However, such features may not necessarily enhance motivation.

4) Instructing children to attempt to read the text independently and then self-check their correctness by selecting sentence level feedback can ensure that they are aware of any errors made.

5) The reading ability of an individual must be considered when selecting an appropriate level of electronic book. Children with small sight vocabularies may still benefit from using computerised texts particularly in the case of key vocabulary, which is repeated frequently throughout the texts. They may, however, have a greater reliance on computer generated word pronunciations than children who have already begun to extend their sight vocabularies. Children with wider sight vocabularies may benefit from hints and activities that are challenging.

Implications for designers:

1) Teachers would like to see, in future implementations, more supporting activities and multiple feedback modes including hints, rather than whole words, and sub-syllabic feedback at the level of onset and rime.

2) Different kinds of features may be more appropriate for specific reading ability levels. The software should become more challenging as a child develops proficiency in the required skills and strategies that need to be utilised in reading. Lower reading ability children may benefit more from simple and basic formats, whereas higher reading ability children may need to have a greater range of feedback modes and activities.

3) Human tutoring strategies, such as providing a hint to encourage independent decoding, may not be effective in computer assisted reading software. Designers must be aware that teachers do not always provide instant whole word pronunciations and that alternatives should be explored in future implementations.

4) Young children have little problem with 'drag and drop' style interfaces but some of them have difficulty executing a 'double click'.

Implications for researchers:

1) The framework for evaluation facilitated a holistic approach to the evaluation of computer assisted reading software.

2) In literacy research, it may not be possible to detect changes in the use of reading cues and strategies in such a limited period of time.

8.2 Limitations of this research

This research has been subject to a number of limitations, which are described below. Firstly, the overall limitations of the project are described. Secondly, the limitations of the empirical study design are summarised. Finally, some objectives of the empirical study were not met and these are discussed.

8.2.1 General limitations of this research

The pilot study concerned older children who had reading ages (as measured by standardised reading tests) that were at least two years below their chronological age. There are many reasons for the problems experienced by such older readers and an experimental design could be problematic because this population requires different approaches and levels of support in literacy instruction. As the approach undertaken in this research required quantitative techniques, the main empirical study focused on young readers who were progressing normally rather than those experiencing difficulties in learning to read.

Another limitation has resulted from the restricted range of computerised texts that were implemented, shortening the length of time over which the study could be conducted. It was problematic for the teachers to identify children who were at a suitable level to benefit from reading the books available because computerised texts from only one stage of the reading scheme were generated. A wider range of electronic books would both facilitate a longer term study to be conducted and enable children to be matched to reading material of an appropriate level.

8.2.2 Limitations in empirical design

The limitations of the empirical design have already been discussed in chapter 7 (section 7.5.5). There were no overall statistically significant differences in the word recognition scores and phonic knowledge measurements between the children who used the enhanced software and those who used the basic version. This may have been partially due to the small number of participants in the study, the large variances in reading ability and the

degree of appropriateness of stage 3 texts for each child. As was noted in Chapter 3 (section 3.4.2) identifying an appropriate control group is problematic and so the design involved two comparison groups who used different versions of the software. In a larger study, including both a non-intervention control group and an additional group participating in an equivalent pen and paper task could have enhanced the findings. However, this approach was beyond the scope of this research. It is not possible to attribute any of the findings to one individual feature as they were introduced simultaneously. It is only possible to state that any differences in the effects of use of the two software versions are due to the enhanced system as a whole.

It should be noted that the sample size was limited and that the results that were obtained should be interpreted with caution. In addition, the statistical analyses concerned gains analyses which can be problematic, particularly as there was a variance in the reading abilities of the participants at the start of the study. The sustainability of the results obtained was not evaluated.

8.2.3 Unanswered research questions

Wise et al. (1989) demonstrated that segmented feedback in computer assisted reading software is beneficial. Goswami and Bryant (1990) argue that phonological training at the level of onset and rime can be beneficial for beginning readers. Yet, surprisingly, this research has not provided any evidence to support the hypothesis that feedback at the level of onset and rime will positively affect the development of phonological skills. Neither has it been able to support the hypothesis that providing coaching in the use of specific reading cues can encourage independent decoding. Yet, the activities (one aspect of coaching provision) were based on commonly used classroom techniques. It may be that there were methodological limitations that prevented the detection of differences in outcomes between the enhanced software users and the basic software users.

There are four reasons that may have contributed to the lack of findings concerning the development of phonological skills. Firstly, tests administered before and after intervention did not address phonological skills at the level of onset and rime; rather they focused upon

phonic synthesis and analysis. Secondly, the frequency of exposure to such feedback was minimised as it was only available in 40% of the texts. Thirdly, the implementation was simplified, giving the initial sound and then the remaining sound in a word. A segmented pronunciation was unlikely to be beneficial for words such as 'downstairs' (given as 'd' + 'ownstairs'). These issues could be addressed by improving the sensitivity of the tests or developing more computerised texts with feedback at the level of onset and rime. Additionally, modifying the implementation of this feature such that the feedback given was true onset and rime (the beginning sound and remaining sound of each syllable) may be more effective. Fourthly, related research, which has demonstrated significant effects of segmented pronunciations, have focused on those children experiencing difficulties in learning to read. Perhaps the participants in this research were not at an appropriate developmental stage to benefit from this form of feedback.

One of the problems partially responsible for the lack of findings relating to the utilisation of reading cues concerns the use of miscue analysis (analysing errors made in oral readings). It was difficult to conduct accurate miscue analysis in naturalistic settings due to frequent interruptions by other children and also occasionally by staff members. Furthermore, it was extremely difficult to identify the oral readings of the lower reading ability children. They were more dependent on computer support and reading in unison with the computer feedback as they clicked on most words. Also, the more able readers made very few oral errors. It now seems unlikely that changes in reading behaviour would be observed over the period of a week. It may have been more effective to conduct miscue analysis at pretest and posttest rather than weekly, on previously unseen text, and in a location where risk of interruptions would be minimised.

8.3 Future research

There are many aspects of this research that could be explored further. These are discussed in terms of design improvements, which were identified as a result of conducting the main empirical study, and the use of the software with alternative populations.

8.3.1 Design modifications

There are a number of major software enhancements that should guide future implementations. The onset and rime feedback feature could be improved by ensuring that each word was broken first into syllables and then giving pronunciations for the beginning and remaining sound in each segment. Furthermore, this form of feedback could be disabled for words where it may not be appropriate, such as the names of characters in the story. The hint mechanism could also be improved by implementing a more complex model for selecting an appropriate clue. Again, this feature could be disabled for words that may not benefit from this form of feedback. A number of minor improvements, such as preventing access to the activities until the complete text has been read, could also be made.

The design could be adapted to support collaborative use. This option should be investigated because technological resources in classrooms are limited and so it is common practice to have pairs or small groups working together at a computer. As speech recognition technology becomes more readily available, this feature should also be incorporated and investigated. Mostow et al. (1994) have already demonstrated the potential of such technology for reading instruction and practice.

8.3.2 Alternative populations

The research could be extended to explore the effects of such software on alternative populations. The advantages that computerised texts can afford those children who are experiencing difficulties in learning to read have already been proposed by some researchers (e.g. Wise et al., 1989). The effects of the innovative design presented in chapter 5 should be ascertained with this group of children. Another population that could be investigated is those children in nursery education. Goswami and Bryant (1990) highlighted the importance of exposure to nursery rhymes and its relationship to reading development. Some of these children may not be able to recognise many words other than their own name and would therefore be heavily dependent on computer feedback. However, the opportunities both for exposure to 'the act of reading' and the development of phonological awareness should be explored. Another application of electronic books that warrants investigation is their role in addressing adult literacy problems.

The most important direction for future research should focus on the whether or not talking book software hinders or helps children as they learn to read. In order to do this, a larger study should be conducted with control groups such that it is possible to evaluate the overall effectiveness of the software. Future research should also focus on design enhancements in electronic books, particularly the most appropriate ways to develop phonological awareness and to provide contextualised instruction.

8.4 Summary

Electronic books are being used in the classroom at a vastly increasing rate. It is therefore vital to show that benefits can be obtained by grounding the development of such software in theory and investigating its use in the classroom. This study has evaluated two versions of electronic books in the naturalistic environment of the primary school classroom. It has demonstrated that both versions of the software were beneficial, contributing to improvements in both cognitive and affective outcomes. The electronic books complemented the pedagogical aims of the teachers who were extremely positive about the resource, as were the children themselves. The basic software version contributed to improving sight recognition of key vocabulary for lower reading ability children. The enhanced electronic books were effective for those children who had already acquired a limited sight vocabulary. The evidence suggests that talking book software can scaffold the process of learning to read. The framework for evaluation of computer assisted reading software was effective in ensuring a comprehensive approach to the investigation of talking books in the classroom.

References

Adams, M. J. (1990). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Afflerbach, P. (1996). Engaged assessment of engaged readers. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), Developing engaged readers in school and home communities (pp. 191-214). Mahwah, NJ: Lawrence Erlbaum.

Almasi, J. F., McKeown, M. G., & Beck, I. L. (1996). The nature of engaged reading in classroom discussions of literature. Journal of Literacy Research, 28(1), 107-146.

Balajthy, E. (1989). Computers and reading: Lessons from the past and the technologies of the future. Englewood Cliffs, NJ: Prentice Hall.

Balajthy, E. (1996). Software classifications: Trends in literacy software publication and marketing. Paper presented at the International Reading Association, New Orleans, USA, April 96.

Ball, S. J. (1993). Self-doubt and soft data: Social and technical trajectories in ethnographic field work. In M. Hammersley (Ed.), Educational Research: Current Issues (pp. 32-48). London: Paul Chapman Publishing Ltd.

Blease, D. (1986). Evaluating educational software. Beckenham, UK: Croom Helm Ltd.

Bolton, M. (1995). Talking Computer Project. (Available from the Author at The Mount, Taunton, Somerset, UK.)

Bonham, B., & Bonham, L. (1993). Naughty Stories [Computer software]. Malmesbury, UK: Sherston Software Ltd.

Bookbinder, G. E. (1976). Salford Sentence Reading Test. London: Hodder & Stoughton Ltd.

Bork, A. (1995). Guest editorial: Why has the computer failed in schools and universities? Journal of Science Education and Technology, 4(2), 97-102.

Breakwell, G. M. (1995). Interviewing. In G. M. Breakwell, S. Hammond, & C. Fife-Schaw (Eds.), Research methods in psychology (pp. 230-242). London: Sage Publications.

Bryant, P., & Bradley, L. (1985). Children's reading problems. Oxford, UK: Basil Blackwell Ltd.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Company.

Carver, R. P. (1990). Reading rate: A review of research and theory. San Diego, CA: Academic Press Inc.

Clay, M. M. (1993a). Reading recovery: A guidebook for teachers in training. Auckland, NZ: Heinemann.

Clay, M. M. (1993b). An observation survey: Of early literacy achievement. Auckland, NZ: Heinemann.

Cohen, L., & Manion, L. (1989). Research methods in education (Third ed.). London: Routledge.

Collins, J., Littleton, K., Longman, J., Mercer, N., Scrimshaw, P., & Wegerif, R. (1996). CD-ROMs in primary schools: An independent evaluation. Coventry, UK: National Council for Educational Technology.

Coolican, H. (1994). Research methods and statistics in psychology (Second ed.). London: Hodder & Stoughton.

Costanzo, W. V. (1989). The electronic text: Learning to read, write and reason with computers. Englewood Cliffs, NJ: Educational Technology Publications Inc.

Crook, C. (1991). Computers in the zone of proximal development: Implications for evaluation. Computers and Education, 17(1), 81-91.

Crook, C. (1994). Computers and the experience of collaborative learning. London: Routledge.

Davidson, J., Coles, D., Noyes, P., & Terrell, C. (1991). Using computer delivered natural speech to assist in the teaching of reading. British Journal of Educational Technology, 22(2), 110-118.

Davidson, J. (1993) The evaluation of a system which gives computer-delivered speech feedback for the teaching of reading. Unpublished PhD thesis, University of Bristol, UK.

Davidson, J. (1994). Using computer-delivered natural speech to assist in the teaching of reading. Computers and Education, 22(1/2), 181-185.

Davidson, J., Elcock, J., & Noyes, P. (1996). A preliminary study of the effect of computer-assisted practice on reading attainment. Journal of Research in Reading, 19(2), 102-110.

DeJean, J., Miller, L., & Olson, J. (1997). CD-ROM talking books: What do they promise? Education and Information Technologies, 2, 121-130.

Delamont, S. (1992). Fieldwork in educational settings: Methods, pitfalls and perspectives. London: The Falmer Press.

del Soldato, T. (1994) Motivation in tutoring systems. Unpublished PhD thesis, University of Sussex, UK.

Draper, S. W., Brown, M. I., Henderson, F. P., & McAteer, E. (1996). Integrative evaluation: An emerging role for classroom studies of CAL. Computers in Education, 26(1-3), 17-32.

Dupagne, M., & Krendl, K. A. (1992). Teachers' attitudes toward computers: A review of the literature. Journal of Research on Computers in Education, 24(3), 420-429.

Ehri, L. C. (1991). Development of the ability to read words. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), Handbook of reading research (pp. 383-417). New York: Longman.

Fawcett, V. (Ed.). (1994). All aboard teachers' resource book. Aylesbury, UK: Ginn and Company Ltd.

Feldmann, S. C., & Fish, M. C. (1991). Use of computer-mediated reading supports to enhance reading comprehension of high-school students. Journal of Educational Computing Research, 7(1), 25-36.

Fife-Schaw, C. (1995). Questionnaire design. In G. M. Breakwell, S. Hammond, & C. Fife-Schaw (Eds.), Research methods in psychology (pp. 174-193). London: Sage Publications.

Gardner, J., Morrison, H., Jarman, R., Reilly, C., & McNally, H. (1994). Learning with portable computers. Computers and Education, 22(1/2), 161-171.

Gardner, J. (1997). ILS and under-achievers. In J. D. M. Underwood & J. Brown (Eds.), Integrated Learning Systems: Potential into practice Oxford, UK: Heinemann for National Council of Educational Technology.

Goodman, K. S. (1973). Analysis of oral reading miscues: Applied psycholinguistics. In F. Smith (Ed.), Psycholinguistics and Reading (pp. 177-182). New York: Holt, Rinehart and Winston, Inc.

Goodman, K. S. (1976). Reading: A psycholinguistic guessing game. In H. Singer & R. B. Ruddell (Eds.), Theoretical models and processes of reading (pp. 497-508). Newark, DE: International Reading Association.

Goswami, U., & Bryant, P. (1990). Phonological skills and learning to read. Hove, UK: Laurence Erlbaum Associates.

Goswami, U. (1996). Rhyme analogy. Child Education(May 1996), 28-29.

Gough, P. B. (1972). One second of reading. In J. Kavanagh & I. Mattingly (Eds.), Language by ear and eye (pp. 331-358). Cambridge, MA: MIT Press.

Greenlee-Moore, M. E., & Smith, L. L. (1996). Interactive computer software: The effects on young children's reading achievement. Reading Psychology: An International Quarterly, 17, 43-64.

Guthrie, J., McGough, K., & Wigfield, A. (1992). The motivations for reading questionnaire. Athens, GA: National Reading Research Centre.

Hall, J., & Rhodes, V. (1986). Microcomputers in primary schools: Some observations and recommendations for good practice. London: Educational Computing Unit, Centre for Educational Studies, Kings College.

Hammond, M. (1995). Learning from experience: approaching the research of CD-ROM in schools. In J. D. Tinsley & T. J. van Weert (Eds.), World Conference on Computers in Education VI, (pp. 283-291). Birmingham, UK: Chapman & Hall.

Harrison, C. (1992). The reading process and learning to read: What a teacher using a 'real books' approach needs to know. In C. Harrison & M. Coles (Eds.), The reading for real handbook (pp. 3-28). London: Routledge.

Hawkridge, D., & Vincent, T. (1992). Learning difficulties and computers. London: Jessica Kingsley Publishers Ltd.

Heller, R. S. (1991). Evaluating software: A review of the options. Computers and Education, 17(4), 285-291.

Hudson, J. (1992). Reading delays. In C. Harrison & M. Coles (Eds.), The reading for real handbook (pp. 224-246). London: Routledge.

Hume, G., Michael, J., Rovick, A., & Evens, M. (1996). The use of hints by human and computer tutors: The consequences of the tutoring protocol. In D. C. Edelson & E. A. Domeshek (Eds.), International Conference on the Learning Sciences, (pp. 135-142). Evanston, IL: Association for the Advancement of Computing in Education.

Hunt, R. (1986). Oxford Reading Tree teacher's guide. Oxford, UK: Oxford University Press.

Issroff, K. (1991). Motivation and computer assisted learning. Technical Report No. CITE 174. The Open University, Milton Keynes, UK.

Issroff, K. (1992). Computer-assisted cooperative learning. Technical Report No. CITE 183. The Open University, Milton Keynes, UK.

Issroff, K., Jones, A., & Scanlon, E. (1994). Case studies of children cooperating with computers: a timebased analysis. In H. C. Foot, C. J. Howe, A. Anderson, A. K. Tolmie,

& D. A. Warden (Eds.), Group and Interactive Learning (pp. 73-80). Southampton, UK: Computational Mechanics Publications.

Johnson, P. (1992). Human computer interaction: Pschology, task analysis, and software engineering. Maidenhead, UK: McGraw-Hill Book Company Europe.

Jones, A., Scanlon, E., Tosunoglu, C., Ross, S., Butcher, P., Murphy, P., & Greenburg, J. (1996). Evaluating CAL at the Open University: 15 Years on. Computers in Education, 26(1-3), 5-15.

Kantrov, I. (1991). Talking to the computer: A prototype speech recognition system for early reading instruction. Technical Report No. 91-3. Newton, MA: Education Development Center.

Keller, J. M. (1987). Strategies for stimulating the motivation to learn. Performance and Instruction, 26(8), 1-7.

LaBerge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

Lange, H. R. (1993). Speech synthesis and speech recognition: Tomorrow's human-computer interfaces? Annual Review of Information Science and Technology (ARIST), 28, 153-185.

Laurillard, D. (1993). Rethinking university teaching: A framework for the effective use of educational technology. London: Routledge.

Lepper, M. R., & Chabay, R. W. (1988). Socializing the intelligent tutor: Bringing empathy to computer tutors. In H. Mandl & A. Lesgold (Eds.), Learning Issues for Intelligent Tutoring Systems (pp. 242-257). New York: Springer Verlag.

Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J.-L. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), Computers as Cognitive Tools (pp. 75-105). Hillsdale, NJ: La[illegible]e F[illegible]baum Associates.

Lewin, C. (1997). 'Test driving' CARS: Addressing the issues in the evaluation of computer assisted reading software. Journal of Computing in Childhood Education, 8(2/3), 111-132.

Lewin, C. (in press-a). Talking book design: What do practitioners want? Computers and Education.

Lewin, C. (in press-b). Evaluating talking books: Ascertaining the effectiveness of multiple feedback modes and tutoring techniques. Forty-sixth yearbook of the National Reading Conference.

Mackay, D.,Thompson, B., & Schaub, P. (1978). Breakthrough to literacy teacher's manual. London: Longman for Schools Council.

Macromedia Director 4 [Computer software]. (1994). San Francisco, CA: Macromedia Inc.

Macromedia SoundEdit 4 [Computer software]. (1994). San Francisco, CA: Macromedia Inc.

Malone, T., & Lepper, M. (1987). Making learning fun: A taxonomy of instrinsic motivations for learning. In R. Snow & M. Marshall (Eds.), Aptititude, Learning and Instruction: Conative and Affective Process Analyses (pp. 223-253). London: Lawrence Erlbaum.

Mason, R. D. (1995). Evaluating technology-based learning. In B. Collis & G. Davies (Eds.), Innovating Adult Learning with Innovative Technologies (pp. 191-199). Amsterdam: Elsevier.

Matthew, K. I. (1997). A comparison of the influence of CD-ROM interactive storybooks and traditional print storybooks on reading comprehension. Journal of Research on Computing in Education, 29(3), 263-275.

McIntyre, D., & Macleod, G. (1978). The characteristics and uses of systematic classroom observation. In R. McAleese & D. Hamilton (Eds.), Understanding classroom life (pp. 111-129). Slough, UK: National Foundation for Educational Research.

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. Reading Research Quarterly, 30(4), 934-955.

McKenna, M. C., Reinking, D., Labbo, L. D., & Watkins, J. (1996). The effects of electronic trade books on the decoding growth of beginning readers. Paper presented at the National Reading Conference, Charleston, USA, Dec 1996:

Medwell, J. (1992). A school policy for reading. Reading, 26, 3-7.

Medwell, J. (1996). Talking books and reading. Reading (April), 41-46.

Meek, M. (1982). Learning to read. London: The Bodley Head Ltd.

Milheim, W. D. (1993). How to use animation in comuter assisted learning. British Journal of Educational Technology, 24(3), 171-178.

Milheim, W. D. (1995-96). Interactivity and computer based instruction. Journal of Educational Technology Systems, 24(3), 225-233.

Miller, L., Blackstock, J., & Miller, R. (1994). An exploratory study into the use of CD-ROM storybooks. Computers in Education, 22(1/2), 187-204.

Miller, L., & Olson, J. (1994). Putting the computer in its place: a study of teaching with techology. Journal of Curriculum Studies, 26(2), 121-141.

Mostow, J., Roth, S. F., Hauptmann, A. G., & Kane, M. (1994). A prototype reading coach that listens. In proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94), (pp. 785-792). Seattle, WA.

Murphy, P., Issroff, K., Scanlon, E., Hodgson, B., & Whitelegg, E. (1995). Gender effects of group work in UK primary schools. Paper presented at Science Education Research in Europe, University of Bath, 7-11 April.

Oldfather, P., & Wigfield, A. (1996). Children's motivations for literacy learning. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), Developing engaged readers in school and home communities (pp. 89-114). Mahwah, NJ: Lawrence Erlbaum Associates.

Parham, C. (1988). Computers that talk. Classroom Computer Learning (March), 26-31.

Parham, C. (1993). CD-ROM storybooks: New ways to enjoy children's literature. Technology & Learning (January), 34-44.

Perfetti, C. A. (1983). Reading, vocabulary, and writing: Implications for computer-based instruction. In A. C. Wilkinson (Ed.), Classroom computers and cognitive science (pp. 145-166). New York: Academic Press.

Pumfrey, P. (1979). Which test? In M. Raggett, C. Tutt, & P. Raggett (Eds.), Assessment and Testing of Reading: Problems and Practices (pp. 25-42). London: Ward Lock Educational.

Pumfrey, P. D. (1985). Reading: Tests and assessment techniques. London: Hodder & Stoughton.

Rayner, K., & Pollatsek, A. (1989). The psychology of reading. Englewood cliffs, NJ: Prentice-Hall

Reinking, D. (1988-9). Misconceptions about reading and software development. The Computing Teacher (Dec/Jan 1988-9), 27-29.

Reinking, D., & Bridwell-Bowles, L. (1991). Computers in reading and writing. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), Handbook of reading research: Volume 2 (pp. 310-340). New York: Longman.

Reinking, D. (1995). Reading and writing with computers: Literacy research in a post-typographic world. In K. A. Hinchman, D. J. Leu, & C. K. Kinzer (Eds.), Perspectives in literacy research and practice. Forty-fourth yearbook of the National Reading Conference (pp. 17-34). Chicago, IL: National Reading Conference.

Reiser, R. A., & Kegelmann, H. W. (1994). Evaluating Instructional Software: A Review and Critique of Current Methods. ETR&D, 42(3), 63-69.

Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-while-listening, and independent reading with computer-based speech feedback. Reading Research Quarterly, 23(2), 219-235.

Roth, S. F., & Beck, I. L. (1987). Theoretical and instructional implications of the assessment of two microcomputer word recognition programs. Reading Research Quarterly, 22(2), 197-218.

Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), Attention and performance (pp. 573-603). New York: Academic Press.

Schofield, J. W., & Anderson, K. (1987). Combining quantitative and qualitative components of research on ethnic identity and intergroup relations. In J. S. Phinney & M. J. Rothera (Eds.), Children's ethnic socialization: Pluralism and development (pp. 252-273). Newbury Park, CA: Sage Publications.

Schofield, J. W. (1993). Increasing the generalisability of qualitative research. In M. Hammersley (Eds.), Educational Research: Current Issues (pp. 91-113). London: Paul Chapman Publishing Ltd.

Schofield, J. W., Eurich-Fulcer, R., & Britt, C. L. (1994). Teachers, Computer tutors, and teaching: The artificially intelligent tutor as an agent for classroom change. American Educational Research Journal, 31(33), 579-607.

Scoresby, K. J. (1996) The effects of electronic storybook animations on third graders' story recall. Unpublished PhD thesis, Brigham Young University, USA.

Scottish Council for Research in Education (1976). The Burt Word Reading Test (1974 revision). London: Hodder & Stoughton.

Sewell, D. F. (1990). New tools for new minds - A cognitive perspective on the use of computers with young children. Hemel Hempstead, UK: Harvester Wheatsheaf.

Smith, F. (1971). Understanding reading: A psychoanalytic analysis of reading and learning to read. New York: Holt, Rinehart & Winston.

Smith, F. (1985). Reading. Cambridge, UK: Cambridge University Press.

Soloman, G., Globerson, T., & Gutterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognitions from a reading partner. Journal of Educational Psychology, 81, 620-627.

Southgate, V. (1984). Reading: Teaching for learning. London: Macmillan Education.

Spaii, G. W. G., Ellermann, H. H., & Reitsma, P. (1991). Effects of segmented and whole-word sound feedback on learning to read single words. Journal of Educational Research, 84(4), 204-213.

Squires, D., & McDougall, A. (1994). Choosing and using educational software. London: The Falmer Press.

Standish, D. (1992). The use of CD-ROM based books to improve reading comprehension in second grade students. ERIC Document Reproduction Service No. ED 352 605.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 16(1), 33-72.

Terry, P., Samuels, S. J., & LaBerge, D. (1976). The effects of letter degradation and letter spacing on word recognition. Journal of Verbal Learning and Verbal Behaviour, 15, 577-585.

Todman, J., & File, P. (1990). A scale for children's attitudes to computers. School Psychology International, 11, 71-75.

Topping, K. (1985). Parental involvement in children's reading: Theoretical and empirical background. In K. Topping & S. Wolfendale (Eds.), Parental involvement in children's reading (pp. 333). Beckenham, UK: Croom Hill Ltd.

Topping, K. (1992). Parents and reading. In C. Harrison & M. Coles (Eds.), The reading for real handbook (pp. 170-185). London: Routledge.

Turner, J. (1995). The influence of classroom contexts on young children's motivation for literacy. Reading Research Quarterly, 30(3), 410-441.

Underwood, J., & Underwood, G. (1990). Computers and learning: Helping children acquire thinking skills. Oxford, UK: Basil Blackwell Ltd.

Underwood, G., & Underwood, J. (1996). Gender differences in children's learning from interactive books. In B. Robin, J. D. Price, J. Willis, & D. A. Willis (Eds.), Technology and Teacher Education Annual. Charlottesville, USA: AACE.

Underwood, J., & Brown, J. (1997). Conclusions. In J. Underwood & J. Brown (Eds.), Integrated Learning Systems: Potential into practice (pp. 190-193). Oxford, UK: Heinemann for National Council of Educational Technology.

van Daal, V. H. P., & Reitsma, P. (1990). Effects of independent word practice with segmented and whole-word sound feedback in disabled readers. Journal of Research in Reading, 13, 133-148.

Venezky, R. L. (1983). Evaluating computer-assisted instruction in its own terms. In A. C. Wilkinson (Ed.), Classroom computers and cognitive science (pp. 31-49). New York: Academic Press.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Walker, R. (1993). The conduct of educational case studies: ethics, theory and procedures. In M. Hammersley (Ed.), Controversies in classroom research (pp. 163-195). Buckingham, UK: Open University Press.

Wepner, S. B., & Feeley, J. T. (1993). Moving forward with literature: Basals, books and beyond. New York: Maxwell Publishing Co.

Wetzel, K. (1991). Speaking to Read and Write: A Report on the Status of Speech Recognition. The Computing Teacher (Aug/Sep 1991), 6-10.

Wise, B., Olson, R., Ansett, M., Andrews, L., Terjak, M., Schneider, V., Korstuch, J., & Kriho, L. (1989). Implementing a long term computerised remedial reading program with synthetic speech feedback: Hardware, software and real-world issues. Behaviour Research Methods, Instruments and Computers, 21, 173-180.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17, 89-100.

Wray, D., & Medwell, J. (1990). Literacy and language in the primary years. London: Routledge.

A. Talking Book Titles used in the Pilot Study

Table A.1: Stage 2 Talking Stories

Title	Author	Illustrator
What a bad dog!	Roderick Hunt	Alex Brychta
The go-kart	Roderick Hunt	Alex Brychta
The dream	Roderick Hunt	Alex Brychta
The toys' party	Roderick Hunt	Alex Brychta
New trainers	Roderick Hunt	Alex Brychta
A new dog	Roderick Hunt	Alex Brychta
Floppy's bath	Roderick Hunt	Alex Brychta
The babysitter	Roderick Hunt	Alex Brychta
The water fight	Roderick Hunt	Alex Brychta
Kipper's balloon	Roderick Hunt	Alex Brychta
Spots	Roderick Hunt	Alex Brychta
Kipper's birthday	Roderick Hunt	Alex Brychta

Table A.2: Stage 3 Talking Stories

Title	Author	Illustrator
On the sand	Roderick Hunt	Alex Brychta
The dolphin pool	Roderick Hunt	Alex Brychta
Nobody wanted to play	Roderick Hunt	Alex Brychta
A cat in the tree	Roderick Hunt	Alex Brychta
The rope swing	Roderick Hunt	Alex Brychta
By the stream	Roderick Hunt	Alex Brychta

Table A.3: Naughty Stories Volume 1

Title	Author	Illustrator
Billy the Bothersome Bully	Bill & Lou Bonham	Paul Hutchinson
Doris the Dotty Dog	Bill & Lou Bonham	Paul Hutchinson
Edwina's Energetic Elephant	Bill & Lou Bonham	Paul Hutchinson
Terry's Tricky Trainers	Bill & Lou Bonham	Paul Hutchinson
Tina's Terrible Trumpet	Bill & Lou Bonham	Paul Hutchinson
Toby the Troublesome Tractor	Bill & Lou Bonham	Paul Hutchinson

B. Records of Participant Interviews (Pre-Test and Post-Test)

B.1 Pre-Test Interview Form

Date: _____ Name: _____ Class: _____

Section 1: Books

What kinds of books do you like best?
true stories/facts ❑ stories ❑ comics ❑

Why?_____

What is your favourite book? Why do you like it?

_____

What kinds of books don't you like?
true stories/facts ❑ stories ❑ comics ❑

Why?_____

What kinds of books do you have at home?
true stories/facts ❑ stories ❑ comics ❑

Do you go to the library? Which library do you go to? (e.g. school, town)

_____

What kinds of books do you get from the library?
true stories/facts ❑ stories ❑ comics ❑

What kinds of books do other people read at home?
true stories/facts ❑ stories ❑ comics ❑ newspapers ❑ magazines ❑

Section 2: Reading

How do you feel about reading?
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

Who do you read to at home?
grown-up ❑ brother/sister ❑ friend ❑ yourself ❑

What kinds of books do you read?

_____

What kinds of books do you like reading best?

_____

What kinds of books don't you like reading?

_____

Do you think reading is hard or easy?
very hard ❑ hard, but ok sometimes ❑ ok, but hard sometimes ❑ really easy ❑

Which is easiest:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Which is hardest:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Which do you like best:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Which do you like the least:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Section 3: Computer at Home

Do you have a computer at home?
yes ❑ no ❑

Who uses it?
grown-up ❑ sister ❑ brother ❑ you ❑ other ❑

How often do you use it?
everyday ❑ weekend ❑ once/twice a week ❑ not very often ❑

What do you use it for?
games ❑ learning ❑ writing ❑ other ❑

Do you help anyone at home with the computer?

_____

Section 4: Computer at School

Do you use the computer at school?
yes ❑ no ❑

Do you use it mostly on your own or with others?
alone ❑ one other ❑ more than one ❑

How often do you use it?
every day ❑ once a week ❑ once a month ❑ not often ❑

What do you use it for?
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Section 5: Computer attitude

How do you feel about using the computer?
don't like it ❑ not bad but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

Why?_____

Do you find it easy or hard to use the computer?
very hard ❑ hard, but ok sometimes ❑ ok, but hard sometimes ❑ really easy ❑

Does using the computer worry you?
always ❑ sometimes ❑ not often ❑ never ❑

Would you like to spend more, less or the same amount of time using the computer in school?
more ❑ the same ❑ less ❑

Section 6: Comparing other reading methods

How do you feel about:
reading alone
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

reading to a friend/parent
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

reading to the teacher
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

reading with a computer
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

What is good about reading with the computer?

_____

What is bad about reading with the computer?

_____

B.2 Post-Test Interview Form

Date: _____ Name: _____ Class: _____

Section 1: Reading

How do you feel about reading?
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

Who do you read to at home?
grown-up ❑ brother/sister ❑ friend ❑ yourself ❑

What kinds of books do you read?

_____

What kinds of books do you like reading best?

_____

What kinds of books don't you like reading?

_____

Do you think reading is hard or easy?
very hard ❑ hard, but ok sometimes ❑ ok, but hard sometimes ❑ really easy ❑

Which is easiest:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Which is hardest:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Which do you like best:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Which do you like the least:
reading ❑ writing ❑ science ❑ maths ❑
other ❑ _____

Section 2: Computer attitude

How do you feel about using the computer?
don't like it ❑ not bad but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

Why?_____

Do you find it easy or hard to use the computer?
very hard ❑ hard, but ok sometimes ❑ ok, but hard sometimes ❑ really easy ❑

Does using the computer worry you?
always ❑ sometimes ❑ not often ❑ never ❑

Would you like to spend more, less or the same amount of time using the computer in school?
more ❑ the same ❑ less ❑

Section 3: Comparing other reading methods

How do you feel about:
reading alone
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

reading to a friend/parent
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

reading to the teacher
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

reading with a computer
don't like it ❑ not bad, but don't enjoy it ❑ like it a bit ❑ enjoy it a lot ❑

What is good about reading with the computer?

_____

What is bad about reading with the computer?

_____

Section 4: Self-evaluation of Talking Book Software

Which was your most favourite talking book and why?

_____

Which was your least favourite book and why?

_____

Having used the computer to practice reading, do you enjoy reading:

more ❑ the same ❑ less ❑

Would you like to continue using the computer to practice reading? YES/NO

Why?_____

Since using the computer to practice, how well do you read:

better than before ❑ the same ❑ worse than before ❑

Why do you think this is?

_____

C. Criterion Referenced Tests

C.1 Basic Sight Vocabulary

Test C.1.1

to	and	the	it	of	he
that	I	is	was	in	a

Test C.1.2

not	so	one	have	his	be
at	with	said	as	you	him
are	on	but	for	all	they
we	had				

Test C.1.3 (Only administered if score from Test 1 + Test 2 > 16)

about	did	now	like	some	were
an	do	little	off	their	what
back	down	first	old	been	look
only	them	when	from	made	or
then	where	get	make	our	big
which	there	more	other	this	go
who	by	has	me	will	out
two	call	her	much	over	up
came	your	if	right	must	want
here	can	into	my	see	well
come	just	no	could	new	she
went	before				

C. 2 Basic Phonics

Test C.2.1 Decoding CVC Words

cat	sun	dig	net	rib	lap
kit	cob	fap	bim	rel	gup

Test C.2.2 Blending

r-a-t	b-i-n	t-o-p	l-a-g	w-e-b	p-u-n
t-e-b	p-o-g	s-u-t			

C.3 Further Phonics (Testing terminated if child able to read less than 25% of words in a sub-test)

Test C .. CCVC (Initial Blends)

step	flat	drip	skim	fret	glum
plug	bren				

Test C.3.2 CVCC (Final Blends)

dust	pump	sink	hand	loft	hint
desk	bond				

Test C.3.3 Consonant Digraphs

chop	thud	whip	shot	pith	rash
much	when				

Test C.3.4 Double-vowel Digraphs

roof	peep	boom	leek

Test C.3.5 Vowel Digraphs

leaf	wait	loud	coil	goat	tea
rail	pie				

Test C.3.6 Vowel-consonant Digraphs

town	coy	ray	pawn	cart	fern
hurt	pork				

Test C.3.7 "Magic-e"

mane	hope	ride	tube

Test C.3.8 Common Endings

wanted	fishing	happily	ration	farmer	cuddle

D. Staff Questionnaire (Pilot Study)

Date: _____ Child's Name: _____ Class: _____

Section 1: Social background and characteristics of child

Please note below any social/medical/behavioural factors which may have contributed to the child's low reading age prior to participating in the study:

_____

_____

Section 2: Reading

Please briefly describe any additional teaching the child has received during the study which may have contributed to an improvement in reading ability:

_____

_____

Please indicate briefly your personal assessment of the child's' reading ability and particular problems:

_____

_____

Since participating in the study has the attitude of the child to reading:

greatly improved ❑ improved ❑ not changed ❑ deteriorated ❑ greatly deteriorated ❑

Any specific examples from the classroom or other comments relating to this:

_____

_____

Since participating in the study has the child's' confidence in reading:

greatly improved ❑ improved ❑ not changed ❑ deteriorated ❑ greatly deteriorated ❑

Any specific examples/comments relating to this:

_____

_____

Section 3: Use of the computer

How often has the child used the computer:

less than once a term ❑ once a term ❑ once a month ❑ once a week ❑
more than once a week ❑

Please indicate which types of software they have used:

word-processor ❑ spreadsheet ❑ database ❑ adventure game ❑

Somerset Talking Computer materials ❑ maths applications ❑ science applications ❑

E. Full Results from Pre-Tests and Post-Tests in Pilot Study

Table E.1: Pre-Test and Post-Test Results for Reading Ages and Common Word Recognition

Child	Pre-S Y, M	Post-S Y, M	Gain Y, M	Pre-B Y	Post-B Y	Gain Y	Pre-C Words	Post-C Words	Gain Words
Aaron	7, 2	7, 6	0, 4	7.1	7.6	0.5	83	88	5
Tony	6, 1	6, 2	0, 1	5.8	6.3	0.5	17	34	17
Trevor	6, 11	7, 0	0, 1	7.1	7.3	0.2	75	73	-2
Laura	6, 10	6, 6	0, -4	7	7.5	0.5	74	73	-1
James	-------	--------	--------	5.1	5.1	0.0	5	4	-1
Karim	6, 5	6, 6	0, 1	6	7	1.0	52	62	10
Leonard	7, 2	7, 6	0, 4	7.5	7.8	0.3	74	84	10
Suraiya	6, 6	6, 9	0, 3	6.9	7.5	0.6	78	80	2
Kunal	6, 6	6, 9	0, 3	6.6	7.3	0.7	51	75	24

Pre-S Pre-test score for Salford Sentence Reading Test, Years and Months
Post-S Post-test score for Salford Sentence Reading Test, Years and Months
Pre-B Pre-test score for Burt Word Reading Test, Years
Post-B Post-test score for Burt Word Reading Test, Years
Pre-C Pre-test score for Common Words Recognition Test (100 words), Words
Post-C Post-test score for Common Words Recognition Test (100 words), Words

F. All Aboard Book Titles replicated as Talking Books

Table F.1: Stage 3 Sam and Rosie Titles (sight recognition)

Title	Author	Illustrator
Tilak's Tooth	Miriam Simon	Julie Park
A Book for Jack	Miriam Simon	Julie Park
Home Time	Miriam Simon	Julie Park
The Butterfly Sale	Miriam Simon	Julie Park
The Greedy Guinea-pig	Julia Jarman	Julie Park
The Ghost in the Castle	Julia Jarman	Julie Park
Scat Cat!	Julia Jarman	Julie Park
Lizzie and the Car Wash	Julia Jarman	Julie Park

Table F.2: Stage 3 Sam and Rosie Titles: Booster Books (sight recognition)

Title	Author	Illustrator
Lost at the School Fair	Monica Hughes	Julie Park
Mo's Photo	Amanda Cant	Julie Park
The Den	Leonie Bennett	Julie Park
Poor Bobby	Leonie Bennett	Julie Park

Table F.3: Stage 3 Pattern and Rhyme Titles (phonological awareness)

Title	Author	Illustrator
Something in the Fridge	Julia Jarman	Tony Ross
The Giant Sandwich	Julia Jarman	Jonathan Langely
I Fell Out of Bed	Miriam Simon	June Goulding
The Accident	Miriam Simon	Bethan Matthews

Table F.4: Stage 3 Pattern and Rhyme Titles: Traditional Tales (phonological awareness)

Title	Author	Illustrator
Counting Chickens	Julia Donaldson	Jeffrey Reid
Town Mouse and Country Mouse	Julia Donaldson	Nick Schon
The Magic Porridge Pot	Amanda Cant	Phil Garner
The Big Turnip	Michelle Clare	Michael Brownlow

G. Questionnaire Sent to Practitioners

Section 1: Software

1. Please indicate which of Sherston's talking books are currently used in your classroom:

❑ Oxford Reading Tree Stage 2 ❑ Oxford Reading Tree Stage 3
❑ Naughty Stories

2. Do you use any of the BrØderbund Living Book Series? Yes/No

3. Do you use any other talking book software? Yes/No

If yes, please specify: _____

Section 2: Who is using it

4. Please indicate which year group you teach:

❑ Reception ❑ Year 1 ❑ Year 2 ❑ Year 3

If other, please specify:_____

5. Please indicate which group of children use talking book software:

❑ All children ❑ Those experiencing difficulties learning to read

If other, please specify:_____

Section 3: How is it being used

6. Please indicate how often the children use talking book software:

❑ Once a term ❑ Weekly ❑ Several times a week ❑ Daily
❑ Depends on an individual's needs

If other, please specify:_____

7. Please indicate how many children generally use the talking book software:

❑ One child at a time❑ Pairs ❑ Three or more at a time

8. We have found that teachers use talking book software in many different ways. Please indicate below that which most closely resembles your teaching style with regard to the use of talking book software:

❑ The children use the software completely independently
❑ The children use the software independently, and support is given if requested
❑ The children use the software independently, and support is offered periodically
❑ An adult works with the children for the duration of a session on a regular basis
❑ An adult works with the children at all times

9. Below are a number of reasons why talking book software may be used in a classroom. Please select any that you feel apply to you:

- ❑ as an integral part of reading instruction (i.e. regular use supported by other activities)
- ❑ as supplementary practice for all children
- ❑ as supplementary practice for those that need it
- ❑ to practice reading books already used as part of a reading scheme or reading instruction
- ❑ to contribute towards satisfying the national curriculum requirements for IT
- ❑ to contribute towards general IT skills
- ❑ to contribute towards general reading skills (e.g. familiarity with books, words etc.)
- ❑ to make stories accessible to those experiencing problems learning to read
- ❑ to build up the confidence of those children with low self-esteem with regard to reading
- ❑ to read stories aloud to early readers, demonstrating the reading process
- ❑ to motivate children and expose them to the pleasures of reading

Section 4: Features of Current Software
Please circle the most appropriate answer: Never / Sometimes / Usually / Always

10. On the whole, do the children find the software easy to use? N / S / U / A

11. Do you use the facility to log the individual words that have been requested? N / S / U / A

12. If so, do you find this useful for monitoring children's progress? N / S / U / A

13. Do you use the autoplay feature to read the entire story to children? N / S / U / A

14. Do you or the children use the index feature to move to specific pages? N / S / U / A

15. Do the children complete the keyword exercises at the end of the story? N / S / U / A

Section 5: Selection of Individual Words

16. The current version of the software gives the whole word when a child selects it individually. This restriction is due to the technological constraints which must be met to ensure that the books can be used on those machines which are available in British classrooms. Which of the following features do you feel would be beneficial in future talking books:

- ❑ Feedback at the level of onset (beginning sound) and rime (remaining sound)
- ❑ Feedback at the level of syllables and/or sub-syllables
- ❑ The option to obtain a hint rather than to obtain the word immediately

17. Are there any other features you would like or comments that you would like to make:

_____

THANK YOU FOR TAKING THE TIME TO COMPLETE THIS QUESTIONNAIRE!

H. Examples of Lingo Scripts

```
--    startMovie
--        Description:  pre-defined Lingo event handler

on startMovie

  -- Global variables and descriptions
  global gPagesRead            -- Number of pages read
  global gInstructions         -- Verbal instructions (TRUE/FALSE)
  global gClickWord            -- Key word to click on in activity
  global gClickList            -- List of key words in all text
  global gOnsetAndRime         -- Is book Pattern and Rhyme strand?
  global gPictureItems         -- What could be seen in illustrations
  global gNumPages             -- Total no. of pages in this book
  global gActivityResults      -- Logging information
  global gBookTitle            -- Current book

  -- Initialise globals which require it when entering new book
  set gPagesRead = []
  set gClickWord = EMPTY

  -- Create pop down menu for Quit etc.
  installMenu cast string(#FileMenu)

  -- Set up screen buttons according to user configuration
  fSetUserState()

  -- Log information: book title and time entered
  set gActivityResults = gActivityResults & RETURN & "Reading New ¬
                         Book:" & gBookTitle && "Time:" & the time

  -- Set cursor to hand with pointing finger for text field
  set myCursor = [(the number of cast string(#cFinger)), ¬
                 (the number of cast string(#cFingerMask))]
  set the cursor of sprite 20 to myCursor

  -- Set up pop down menu, if verbal instructions selected
  if (gInstructions = TRUE) then
    fSetCheckMark()
  end if

  -- If its pattern and rime, set the global to TRUE else FALSE
  set gOnsetAndRime = FALSE

  -- List of keywords in book
  set gClickList = [#about, #all, #at, #back, #big, #but, #came,
                    "didn't", #down, #everybody, #for, #got, ¬
                   #looked, #out, #put, #some, #then, #went, ¬
                   #wanted, #gave, #made, #at]

  -- List of 6 items which can be seen in illustrations
  set gPictureItems = [#book, #cat, "Jack", "Sam", #panda, #pens]

  -- Set up number of pages in this book
  set gNumPages = 13

end startMovie
```

Figure H.1: Start-up script invoked when book entered

```
--    fWhatCanISee
--        Description:  identifies what can be seen in picture
--        Parameters:   NONE
--        Return Value: the list of what can be seen on page gFrameNo
--        Caveats:      Uses gFrameNo - current page, must be
--                      set up in calling environment

on fWhatCanISee

  global gFrameNo                    -- the frame we are on now

  if (gFrameNo = 2) then
    set lCanSee = [#book, #Sam, #panda]
  else if (gFrameNo = 3) or (gFrameNo = 9) or (gFrameNo = 10) then
    set lCanSee = [#book, #Sam]
  else if (gFrameNo = 4) or (gFrameNo = 13) then
    set lCanSee = [#book, #Jack, #Sam]
  else if (gFrameNo = 11) then
    set lCanSee = [#Sam, #pens]
  else if (gFrameNo = 12) then
    set lCanSee = [#book, #Sam, #pens]
  else
    set lCanSee = [#book, #Sam]
  end if

  return lCanSee
end fWhatICanSee
```

Figure H.2: Script called by illustration activity to identify what can be seen on a particular page of the book

```
--    fPageSentenceOK
--        Description:  checks if sentence shuffle activity valid
--        Parameters:   NONE
--        Return Value: TRUE if valid, FALSE otherwise
--        Caveats: Uses gFrameNo - current page, must be
--                      set up in calling environment

on fPageSentenceOK
  global gFrameNo

  set lReturn = TRUE

  -- Prevent sentence shuffle on pages where it doesn't make
  -- sense. For example, on page 12 the sentence reads
  -- "Sam made a book about a baby and ...". This is not a complete
  -- sentence and may be confusing for a user, so mark as invalid.
  if (gFrameNo = 4) or (gFrameNo = 11) or (gFrameNo = 12) or ¬
                      (gFrameNo = 13) then
    set lReturn = FALSE
  end if

  return (lReturn)
end fPageSentenceOK
```

Figure H.4: Script which checks if meaning and syntax activity will be valid on a specific page

I. List of Key Words

I.1 Stage 1 Key Words

Adam	Mum	a	the
Dad	Sam	and	they
Grandad	Rosie	in	to
Jack	Tilak	it	wanted
Lizzie		no	went
Mo		saw	

I.2 Stage 2 Key Words

gave	his	on	sat
had	I	one	she
he	Let's	play	was
her	made	said	yes

I. 3 Stage 3 Key Words

about	but	for	some
all	came	got	then
at	didn't	looked	
back	down	out	
big	everybody	put	

J. Staff Pre-Test and Post-Test Questionnaires

J.1 Pre-Test Questionnaire

Informal Reading Assessment

Date: _____ Childs Name: _____
Class: _____ School: _____

Section 1: Teaching

Please tick ONE box below which most closely represents the resources being used to teach the child to read:

Ginn All Aboard ❒
Another reading scheme ❒
Real books ❒
A Combination of real books and a scheme ❒

If you are using one or more other reading schemes, please state which:

_____

There are two options available in the software. An unknown word can be obtained instantly or a general hint may be given first (such as look at the initial sound etc) Do you feel that the child would benefit more from instant feedback or a hint first, please indicate your choice:

instant feedback ❒ hint first ❒

The All Aboard scheme is divided into 3 strands, two of which have been replicated in the software. The pattern and rhyme strand has been designed to support the development of phonic awareness. There are two options in these books concerning instant feedback, please indicate that which you feel will best suit the child concerned:

feedback at the level of onset and rime first, then whole word ❒

whole word feedback only ❒

Section 2: Current Strategy Use

Please tick those strategies that you feel the child already employs regularly when decoding unknown words:

Initial letter or beginning sound: ❒
Word analogy: ❒
Phonic synthesis: ❒
Meaning: ❒
Syntax: ❒
Illustration: ❒

Section 3: Motivation and Confidence:

Please indicate the child's attitude to reading and books by circling a number below where 1 represents negative attitude and 5 represents positive attitude:

1 2 3 4 5

Please indicate the child's confidence level when reading aloud to an adult or peer where 1 represents low confidence level and 5 represents a high confidence level:

1 2 3 4 5

J. 2 Post-Test Questionnaire

Post Intervention Report

Date: _____ Childs Name: _____
Class: _____ School: _____

Section 1: Current Strategy Use

Please tick those strategies that you feel the child now employs regularly when decoding unknown words:

Initial letter or beginning sound: ☐
Word analogy: ☐
Phonic synthesis: ☐
Meaning: ☐
Syntax: ☐
Illustration: ☐

Section 2: Motivation and Confidence:

Please indicate the child's current attitude to reading and books by circling a number below where 1 represents negative attitude and 5 represents positive attitude:

1 2 3 4 5

Please indicate the child's current confidence level when reading aloud to an adult or peer where 1 represents low confidence level and 5 represents a high confidence level:

1 2 3 4 5

Please comment on how you feel the child has benefitted from participating in this study and whether you feel that there have been any disadvantages for this individual:

_____

_____

_____

Section 3: General Comments

Do you feel that there have been any general problems with the project overall? YES/NO

If YES, please explain:

_____

_____

If you had the resources available, would you consider incorporating talking books in your classroom?

_____

_____

If so, how would you use it? (How often and for how long, individuals or pairs etc)

_____

_____

What kind of information would you like the system to record? (Words selected, activity results, time taken, books read etc)

_____

_____

K. Example of Modelling Mechanism

Table K.1 presents an example of the modelling mechanism in operation. The type of interaction is briefly described and the models that change as a result are indicated by shading. A variety of interactions are included.

On starting the book, the child is recorded as being proficient both at using the illustration, onset and meaning cues, and at recognising key words in the texts (models are recorded as value five or greater). She selected the word 'behind' and the models representing use of onset and meaning are decremented. A hint is given initially and on this occassion it relates to the use of meaning. As this is specific coaching, the model representing the meaning cue is incremented (indicated in bold). Another word ('tooth') is then selected for help and a hint is given in which the use of the rime is targeted. Then the child attempts an activity, the first of which provides coaching in key word recognition. She completes it successfully on the first attempt and so the key word model is incremented, indicating an increasing competence in key word recognition. Two more key word activities are undertaken with similar outcomes and then she is presented with an illustration activity. She fails the first three attempts at this activity and so the model representing the use of illustration is decremented each time. Finally, she is successful at three more activities, another illustration activity, another key word activity and a sentence shuffling activity designed to improve the use of meaning and syntax cues. Each time the associated models are incremented. When she exits the Talking Book, the representation of her levels of competence in the use of the cues indicates that she is using the meaning cue successfully and has good sight recognition of the key words, but she now needs further practice in the use of the illustration, onset, rime and syntax cues.

Table K.1: Example of modelling mechanism in operation

Interaction	Illustration	Onset	Rime	Meaning	Syntax	Keyword
Starting book	5	5	0	5	0	5
word selected: 'behind'	5	**4**	0	**4**	0	5
hint given for word selected: target cue - use of meaning	5	4	0	**5**	0	5
word selected: 'tooth'	**4**	**3**	0	**4**	0	5
hint given for word selected: target cue - use of rime	4	3	**1**	4	0	5
keyword activity: successful attempt	4	3	1	4	0	**6**
keyword activity: successful attempt	4	3	1	4	0	**7**
keyword activity: successful attempt	4	3	1	4	0	**8**
illustration activity: successful attempt	**5**	3	1	4	0	8
illustration activity: failed attempt 1	**4**	3	1	4	0	8
illustration activity: failed attempt 2	**3**	3	1	4	0	8
illustration activity: failed attempt 3	**2**	3	1	4	0	8
illustration activity: successful attempt	**3**	3	1	4	0	8
keyword activity: successful attempt	3	3	1	4	0	**9**
sentence shuffle activity: successful attempt	3	3	1	**5**	**1**	9

L. Case Study Data

Table L.1: Pre-Test and Post-Test Results for Five Selected Case Studies

Case	Jason		Justin		Mark		David		Rifat	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Salford Sentence (years)	6.42	6.42	6.50	6.75	7.17	7.42	6.08	6.25	6.33	6.67
Burt Word (years)	5.60	6.30	5.90	6.40	6.70	7.80	6.10	6.10	6.30	6.70
Common 100	15	38	37	54	72	87	15	16	50	60
Key Words (max = 44)	14	28	20	26	39	43	16	23	35	38
Phonic-CVC (max = 12)	2	2	1	2	8	6	2	0	3	8
Phonic-Blend (max = 9)	5	9	1	4	7	8	2	0	4	7
Phonic all (max = 75)	8	11	3	6	26	26	2	2	4	19
Motivation (rated 1-5)	3	4	3	3	3	5	3	3	5	5
Confidence (rated 1-5)	3	3	2	3	2	5	4	3	2	4
Initial Letter (0: no 1:yes)	0	1	1	1	1	1	1	1	1	1
Word Analogy (0:no 1:yes)	0	0	0	0	0	1	0	0	0	0
Meaning (0:no 1:yes)	0	1	0	0	1	1	1	0	0	1
Syntax (0:no 1:yes)	0	0	0	0	0	0	0	0	0	0
Illustration (0:no 1:yes)	1	1	0	0	1	1	1	0	1	1

M. Observation Records for Case Studies

M.1 Case Study 1

Table M.1: Observed behaviours during use of software

Behaviour	first observed session (day 7)	second observed session (day 20)
recognise miscues, get hint immediately	6	3
recognise miscues, solve independently	3	1
instant access	2	2
hint ignored	6	5
realises miscued after sentence feedback	2	0
miscues	1	2
repeat words	17	4
use initial letter	5	1
reread to bolster comprehension	3	3
physical signs of loss of concentration	7	16
physical signs of enjoyment (e.g. laugh)	1	0
comment on success	2	0
comment on software use	7	2
comment on low self-esteem	1	0
not finishing sentence	0	1
confident completion of activity	0	2
points to words	6	0
comment unrelated to reading	0	2

M.2 Case Study 2

Table M.2: Observed behaviours during use of software

Behaviour	first observed session (day 15)	second observed session (day 20)
recognise miscues, get hint immediately	1	3
recognise miscues, solve independently	0	4
instant access	19	9
hint ignored	18	9
realises miscued after sentence feedback	1	1
miscues	1	4
repeat words	9	3
use initial letter	0	1
reread to bolster comprehension	0	1
physical signs of loss of concentration	11	25
physical signs of enjoyment (laugh/smile)	0	2
comment on success	0	0
comment on software use	6	12
comment on low self-esteem	0	0
not finishing sentence	1	0
confident completion of activity	several	2
points to words	0	0
comment unrelated to reading	6	8
mimic computer output	8	11
comment relating to text	5	8
signs of frustration	1	0

M.3 Case Study 3

Table M.3: Observed behaviour during use of software

Behaviour	first observed session (day 8)	second observed session (day 20)
recognise miscues, get hint immediately	0	1
recognise miscues, solve independently	3	3
instant access	0	0
hint ignored	0	1
realises miscued after sentence feedback	0	0
miscues	0	1
repeat words	1	0
use initial letter	1	0
reread to bolster comprehension	4	0
physical signs of loss of concentration	7	1
physical signs of enjoyment (laugh/smile)	0	1
comment on success	0	0
comment on software use	2	1
comment on low self-esteem	0	0
not finishing sentence	0	0
confident completion of activity	3	2
points to words	30	22
comment unrelated to reading	6	0
comment relating to text	2	4
signs of frustration	0	0

M.4 Case Study 4

Table M.4: Observed behaviour during use of the software

Behaviour	first observed session (day 7)
recognise miscues, solve via another strategy	1
instant access	68
realises miscued after sentence feedback	0
miscues	7
repeat words	49
use initial letter	1
reread to bolster comprehension	0
physical signs of loss of concentration	6
physical signs of enjoyment (laugh/smile)	0
comment on success	1
comment on software use	0
comment on low self-esteem	1
not finishing sentence	0
points to words	0
comment unrelated to reading	5
comment relating to text	0
signs of frustration	0

M.5 Case Study 5

Table M.5: Observed behaviour during use of the software

Behaviour	first observed session (day 20)
recognise miscues, solve via another strategy	3
instant access	4
realises miscued after sentence feedback	0
miscues	10
repeat words	1
use initial letter	0
reread to bolster comprehension	6
physical signs of loss of concentration	3
physical signs of enjoyment (laugh/smile)	0
comment on success	0
comment on software use	0
comment on low self-esteem	0
not finishing sentence	0
points to words	0
comment unrelated to reading	0
comment relating to text	0
signs of frustration	0