Chapter 1
Introduction

1.1 Introduction

Spelling competence is important at all levels of written communication and an important skill for school children. Whereas writing has become relatively automatic for normally achieving students by the upper elementary grades, being a tool to generate ideas and for educational output, this skill is lacking in children with spelling difficulties.

Written expression is problematic for children with spelling difficulties, with respect to quality and quantity of work produced (Williams, 2002) and effort required to produce it. The spelling errors of children with spelling difficulties can be severe enough to make written work virtually or actually unreadable (Newell, Booth, Arnott & Beattie, 1992). The mechanics of writing (handwriting, spelling and punctuation) still dominate in writing activities, interfering with and inhibiting higher level composition processes (message construction and idea generating thoughts) (MacArthur, 1999). MacArthur suggests that the mechanics of writing are less important than the message that needs to be communicated. If the mechanics of writing can be supported, then the higher level processes of writing can proceed with less interference from lower level deficiencies.

1.2 Word prediction as a tool for spelling support

In the past, spelling problems in the classroom have been addressed with drill-and-practice type instruction, or through phonological remedial training. In more recent years, computer tools have proved useful to support children with spelling difficulties to write more fluently and accurately. Word processors, spelling and grammar checkers, sound feedback and speech recognition have been used as writing support with some success (MacArthur, 1999).

Word prediction, although originally designed to support the writing needs of people with physical disabilities, has also been offered as a medium of writing support for children with spelling difficulties (MacArthur, 1999). Word prediction has provided children with spelling difficulties a prosthetic tool for writing output, so that they can produce written output along with their peers, even without the spelling skills usually required to do so. The main function of assitive technology after all is to allow people with disabilities to participate on an equal basis with people without disabilities (Anson, 1993).

Case studies, although mostly anecdotal in nature (Tam, Reid, Naumann & O'Keefe, 2002), among people with a range of disabilities, including learning disabilities and language delays, have shown that students with severe spelling problems may benefit from using word prediction (MacArthur, 1999; Newell et al., 1992; Williams, 2002).

However, the effectiveness of word prediction has produced a wide range of results in the literature. The literature lists a number of studies, with results which range from improvement of the rate and quality of text production, to no difference in the rate of text entry when comparing conditions with and without word prediction, to reports of lower rates with word prediction (Horstman & Levine, 1994; Tam et al., 2002). Mixed results are attributed to differences in research methodologies, study populations, computer access devices, word prediction programs, user characteristics (Tam et al., 2002) and training interventions (Horstman & Levine, 1992). The interactions between these aspects of word prediction research are complex, and it is difficult to know what effect the various combinations will have on overall performance (Horstman & Levine, 1992). Williams (2002) suggests that the conflicting evidence provided by research may be more a function of the nature of instruction and instructional feedback that accompanies the use of word prediction, than a reflection of the efficacy of the technology itself.

The various research methodologies applied to word prediction studies have varied from repeated-measures, single-subject, alternating research designs, to group studies, across a variety of subjects and tasks etc. (Tam et al., 2002), all contributing to the growing body of knowledge and understanding of the complex word prediction issue. The study populations of the various word prediction studies from which the research groups are drawn can also have a significant bearing on results pertaining to the efficacy of word prediction for writing support where there are writing difficulties. Many experimental studies on word prediction are with participants without disabilities and in laboratory settings and they caution against generalising results from one population to another. Different users adopt a wide range of strategies to cope with the individual nature of their disabilities and it is inappropriate to use quantitative data derived from able-bodied people in models that attempt to predict the performance of disabled people (Newell, Arnott & Waller, 1992). Research projects on able-bodied people may sometimes find little benefit in word prediction usage in terms of rate enhancement (Horstmann & Levine, 1990), but Newell et al. (1992) argue that it would be unfortunate if therapists with little practical experience of using predictive systems decided against investigating predictive software as a means of improving text entry, based on results such as those of Horstmann and Levine (1990). Williams (2002) suggests that decisions regarding the selection of assistive technologies for students must be made on a case-by-case basis. Some students may benefit greatly from the use of word prediction software, but other students with different written expression needs may find little use for this software, and production rate may even be hindered by it (Horstman & Levine, 1992; Williams, 2002).

The computer device that is being used also has an impact on word prediction effectiveness, mostly due to the ease of access variations and the capabilities of the computer hardware and software itself. Technological advances in hardware sophistication have been able to accommodate software advances, mostly related to increasing the keystroke savings that can be effected. Predictive software can allow for predictions based on previous words or syntactical rules, but requires powerful hardware. However, attempts to improve performance through system design often do not work as well in practice as they do in theory, because of the complexities of the interactive process between the user and the system (Koester & Levine, 1996).

The parameters to which a word prediction system can be set are many. The number of permutations of parameter sets are extensive. This alone has made quantifying the effectiveness of word prediction programmes difficult. Many parameters of word prediction have been isolated and their effect investigated (see Table 2.1), but the interplay between the effects of combining the various word prediction parameters still challenge researchers.

The specific characteristics or skills of the user impact very significantly on the apparent efficacy of word prediction usage (Koester & Levine, 1998; Szeto, Allen & Littrell, 1993; Venkatagiri, 1994). The motor skills of the user have a direct relationship to speed enhancement in word prediction. There is a significant increase in visual-cognitive demands in word prediction usage as much cognitive processing is required to use and guide any activity incorporating word prediction use (Klund & Novak, 1995). Word prediction effectiveness also depends heavily on the motivation of the user to use it and the perception that it does in fact help (MacArthur, 1999).

Although many case studies have reported how word prediction has aided spelling, there is little evidence in the literature to attempt to investigate, quantify or correlate the impact of specific literacy skills on the effectiveness, or lack thereof, of a word prediction program to facilitate written work, with respect to rate enhancement or accuracy improvement. Gillette and Hoffman (1995) recognise that a significant number of literacy skills are essential before a user can be expected to be successful with word prediction, but these skills are neither specified nor quantified.

Word prediction requires spelling skills. The user must begin to spell a word, primarily through his knowledge of phonetic rules. Words can (generally) not be accessed without the ability to begin to spell them. However, as words are predicted, the user must search for the required word from a list of words, and recognise the correctly spelt one, if it has appeared. It can therefore be seen that by generating text through the aid of word prediction, not only spelling is required, but sight reading skills too.

A lack of spelling competence is limiting in writing, and thus in most aspects of literacy development. Spelling knowledge is foundational to both efficient reading and writing (Ganske, 1999). It is the purpose of this research project to investigate the role that spelling competence has on the ability to use a word prediction programme to spell words, using Grade 4 - 6 children with known spelling difficulties from a school for children with special needs.

The results of this research project will serve as additional information in facilitating decisions regarding the use of word prediction in the context of special education. The results will also serve to aid in the decision of which children will benefit from word prediction support, and at what stage of literacy development (in terms of spelling competency) the most advantageous time to introduce it into their educational program would be.

1.3 Outline of the Chapters

The research is presented in five chapters. Chapter 1 provides the basic introduction and motivation for the research. Word prediction is proposed as a tool to aid children with spelling difficulties. In Chapter 2 the features of word prediction that serve to support spelling difficulties are discussed. Chapter 3 describes the methodology of the research project that was designed to investigate the effectiveness of word prediction to improve spelling accuracy and increase typing time. Chapter 4 presents and discusses the results of the research project. In Chapter 5 the study is critically evaluated in terms of its results, limitations and strengths.

1.4 Abbreviations

The following abbreviations are used in the text

Age	Chronological age
Approx	Approximations
Diffage	Difference between chronological age and spelling age
Diffscore	Difference between the accuracy scores in the word prediction subtest and typing only subtest
Difftime	Difference between the time scores in the word prediction subtest and the typing only subtest
GWST	Graded Word Spelling Test
MKR	Minimum keystrokes required
Spage	Spelling age
TO	Typing Only subtest
TO score	Accuracy scores in Typing Only subtest
TO time	Time scores in Typing Only subtest
WP	Word Prediction subtest
WP score	Accuracy scores in Word Prediction subtest
WP time	Time score in Word Prediction subtest

1.5 Definition of terms

Approximations	Words chosen from the prediction lists that were correctly spelled words but were not the target word.
MKR	Minimum keystrokes required to spell a word. In conventional typing, the MKR of a word is the number of letters in the word plus the keystroke required for the space after the word. When typing with the use of word prediction, the MKR is the number of letters that have to be typed before the word appears in the prediction window, plus the keystroke required to select the word from the prediction list. No keystroke is required for the space after the word, as the space is inserted automatically. However, in this study, the automatic space that word prediction inserts after a word is entered, was not utilised.
Spelling ability	This term is expressive of a child's spelling development relative to his age or grade. The GWST scores converted into spelling age is an indication of spelling ability.
Spelling knowledge	A term used to express the amount of spelling knowledge a child has, or the extent to which he has learnt phonetic and spelling rules. It does not reflect how much the child *should* know relative to his grade or age. The GWST raw scores give an indication of spelling knowledge.

Note : It is important to distinguish between spelling knowledge and spelling ability, because a younger and older child scoring at the same GWST raw score, will not necessarily respond to word prediction in the same way, as their developmental lags are different.

Target word	The word that the child was required to spell.
The task	The experimental task the children were required to do, which consisted of training in on-screen keyboard use, training in word

	prediction and two subtests. The subtests consisted of typing out 30 words, with and without the use of word prediction.
Word prediction parameters	The features of a word prediction programme. These features can be customised to user requirements, or are related to the sophistication of the word prediction program.

1.6 Summary

Spelling difficulties impact on the quality and quantity of written work, a skill essential for academic progress. In recent years, word prediction has been offered as support for children with spelling difficulties to produce written work. However, a literature survey has shown wide-ranging results on the efficacy of word prediction to improve the accuracy and speed of typing in users with spelling difficulties. The mixed results are possibly due to a number of factors, the most influential being the parameters of the word prediction program itself and the specific skills of the user. This study investigated the relationship between word prediction usage and the user skill of spelling knowledge.

Chapter 2

Word prediction and spelling difficulties

2.1 Introduction

In this chapter, word prediction is described and its primary features analysed for their potential to support children with spelling difficulties. The support is through word completion and keystroke reduction. The manner in which word prediction is uniquely placed as a skill that integrates spelling and reading skills provides the theoretical base of why word prediction could benefit children with spelling difficulties. The spelling benefits gained through word prediction are understood against the cognitive and time costs associated with that benefit, as well as the possibility of the benefits being enhanced through training and practice.

2.2 Word prediction described

Word prediction operates by suggesting words that are in the process of being typed, and completing the word if it is selected. As letters are pressed on the keyboard, the predicted words appear in a prediction window. The predicted words are always words with the same letters that have already been typed, but are chosen based on lists of most commonly used words, general frequency of use of those words in that language, specific frequency of usage by the user, recency of use and grammar formulae. Word prediction completes words that the user is typing, as though entered in full. The prediction list starts suggesting words from the first letter typed, and gets closer to the target word with every subsequent letter typed (if it is in the vocabulary list). At any stage from its appearance in the prediction list, the target word may be selected, and the word being typed will be completed. The prediction list is cleared after either a prediction is selected or a word is entered fully and terminated with a space or punctuation symbol. The process begins again with the beginning of a new word. Word prediction is actually a misnomer. What is usually referred to as word prediction, is actually word completion (Klund & Novak, 1995). Prediction really only comes into play with more sophisticated word prediction software, where the next word can be predicted before a letter of the next word is typed.

2.3 Spelling accuracy support through word completion

It is word prediction's word completion feature that makes word prediction a powerful tool to facilitate spelling. Word prediction can offer spelling support by providing cues for spelling words correctly for individuals who have not yet achieved good skill in spelling (Gillette & Hoffman, 1995). Zordell (1990) found that if a student could type in the first few letters of a word, the word prediction software identified the word and eliminated the need to spell it in its entirety. Klund and Novak (1995) noted how a user could often recognise and select a word from a list that he could not spell. In Newell et al.'s (1992) study, children with severe spelling difficulties were often able to produce the first letters of words, although they could not spell the word correctly. It is much easier to produce the initial letters of a word and then recognise the correctly spelt whole word, than it is to spell the word in full independently (Newell et al., 1990). In spelling, every letter counts, and in the correct order too, but in reading, a word may be accurately retrieved from memory on the basis of just some of its letters (Ganske, 1999). The essence of the spelling support offered by word prediction appears to be that the spelling skill of encoding words letter by letter is being replaced by the possibility of selecting the word rather than forming it.

It can be seen then that in word prediction use, there is an interaction between spelling and reading skills. A word is begun by application of spelling rules, but can be completed by recognising and selecting it from a list of suggested words offered by the word prediction program. Word prediction spelling support is characterised by changing a spelling task to a spelling + sight word reading task, where there is a continual interplay between the two skills.

2.4 The interaction between reading and writing skills

The interaction between writing and reading is well supported in the literature. Shanahan and Lomax (1986) see reading and writing as a constellation of interrelated processes that utilise a number of knowledge bases. Their interactive model proposed influences between the reading and writing dimensions, with reading ability preceding writing ability, and writing informing reading. If reading skills are more developed than writing skills, as is usually the case in early literacy development, a child would then more likely be able to identify a word than to spell it, making the word prediction tool a useful one in aiding the child to produce words.

According to Berninger (2000), spelling and reading are not separable skills and continually interact in the language development of a child. He suggests that people are multilingual, with four functional language systems - an aural language system (language by ear), an oral language (language by mouth) system, a written form of language or reading (language by eye) and a written form of language or writing (language by hand). These four language systems develop in overlapping, parallel waves rather than in discrete, sequential stages. As each of these language systems develops, connections are created with each of the others. In Berninger's language development, writing is the last language system to develop into a functional language (children can read messages before they can write them). The other languages can thus be accessed to aid in the development of writing.

For Ganske (1999) the relationship between reading and writing is reciprocal. As children learn to read, graphic images of sounds and words, as well as their phonological, semantic, and syntactic identities are stored in memory in a kind of mental dictionary, linking the graphemes to phonemes and letters to sounds. The more a child knows about letter-sound relations, the easier it is to read words; and the more words a child knows by sight, the stronger the child's orthographic knowledge.

There is thus much support in the literature for the idea that practice at spelling or producing words would improve reading skills, and that seeing the word (reading it) frequently, would improve the ability to spell it. By using spelling and reading together in the interactive way required by word prediction, both modalities could benefit from each other, as the user is constantly drawing on both his spelling and reading skills in his efforts to produce words.

In fact, there is a support for a more integrated approach to the teaching of writing and reading based on the strong association between the two skills (Shanahan & Lomax, 1986;

Ganske, 1999). Word prediction can thus be seen not only as a prosthetic tool for spelling difficulties, but also as a training aid (Newell et al, 1992) that is educationally valid.

There are, however, some fundamental differences between spelling and reading. According to Allred (1990), whereas reading is a decoding process, spelling is an encoding process. Readers must *say* or *think* what they *see*. Spellers must *write* what they *hear* or *think*. Thus, although reading and spelling are both important aspects of literacy, they are inherently different processes. Although there is a high correlation between the performance of individuals between reading and spelling skills, there are good spellers who are not good readers, and good readers who are not good spellers. Reading and writing draw on shared knowledge, but are separable skills (Berninger, 2002)

Discrepancies between reading and spelling skills may have an important influence over the effectiveness of word prediction. If a child's reading is a weaker skill than his spelling, it may seriously limit the chances of word prediction effecting a speed increase, or even a spelling accuracy increase. This is because word completion operates on the principle of identifying words by recognising them (a reading skill). If a child struggles to identify words in the prediction list, he will be unable to access the potential benefits of word prediction. On the other hand, if a child with poor spelling skills has stronger reading skills, word prediction could be expected to have a strong positive influence in improving the child's spelling speed and accuracy. His weaker spelling skills are compensated for by his stronger reading skills. Or, to put it another way, he is able to divert the task of word formation from his weaker skill to that of his stronger skill.

2.5 Speed enhancement through keystroke reduction

Word prediction's word completion feature also exploits redundancy in natural language to reduce the number of character entries necessary to produce a piece of text (Newell et al., 1992), thereby improving input efficiency (Anson, 1993). This reduction in typing entries is called keystroke savings. On average, 2-4 characters are required to select a word. The average word length in the English language is approximately 6 characters long. Word prediction software therefore offers a 50% theoretical reduction in keystrokes (Anson, 1993).

Word completion and keystroke reduction would appear to suggest a corresponding decrease in input time compared to the input time required to type a word in full. However, this has proved to not be the case. Researchers have long recognised that reduction of keystrokes does not translate into equivalent improvements in rate. Venkatagiri (1993) noted that even with a nearly 50% reduction in keystrokes, typing with word prediction required as much time to write a message as it did to type the message without word prediction. The benefit in keystroke savings was usually offset or even exceeded by the cost of making each selection (Koester & Levine, 1994). Where the speed of writing increased when using word prediction, mild visual and/or motor impairments were usually present (Zordell, 1990). Word prediction seldom sped up writing for those with near normal motor ability (Venkatagiri, 1993). The greater the presence of motor difficulties, the greater the positive impact on word acceleration (MacArthur, 1999). For this reason, it was important to establish that none of the subjects for the current research project had motor impairments, as this would significantly impact on the timing aspects.

Although keystroke savings have not resulted in the rate improvements expected, many researchers still found word prediction useful in improving the quantity of work produced (Klund & Novak, 1995; MacArthur, 1999). In Newell et al.'s study (1992), word prediction was found to be very successful in all but one of the 17 special education needs children, in improving the quantity of written work which they produced. Williams (2002) reported the length of her subjects' papers to increase. The increased quantity could probably be explained more by the decrease in frustration and increase in motivation that can accompany more success at writing, than by speed of typing increases.

According to the literature, speed of typing increases could not be expected to be found in the current research project, as the subjects would not have motor difficulties. However, it is difficult to dissociate accuracy from speed. Two important indicators of the efficiency of any device-user interaction are speed and accuracy of message generation, i.e. the operational competence (Szeto, Allen & Littrell, 1993). Significant gains in spelling accuracy would lose their value if there were significant costs in speed decreases and effort increases.

2.6 The cost of cognitive demands in word prediction usage

The main reason for the reported lack of speed increase in word prediction (where speed increase would be expected due to the keystroke savings), is explained by the additional visual-cognitive demands that word prediction introduces over regular keyboard typing. The need to keep shifting the attention from the text that is being typed to the prediction list places a significant burden on the user. There is also the constant changing of the point of gaze to and from the keyboard (Klund & Novak, 1995). Visual cognitive demands are particularly limiting when the typing task is a copy one (Anson, 1993), as attention also has to shift to and from a text source external to the screen. Visual searching of the prediction list has been shown to be a time-consuming and energy expending task, placing a burden on the user. The user must constantly and carefully attend to the word list because it changes after each new letter typed. Decision making is a constant factor while typing and places heavy demands on working memory. There are the decisions of when to search, deciding if the word is there and deciding what to do next if the word is not there (Tam et al., 2002).

It is also important to keep in mind the interference effect that searching for the letter on the computer keyboard may have on the actual spelling process in word prediction usage. Whether typing on a conventional keyboard or an on-screen keyboard, the typing process of spelling is different to the handwriting process of spelling. In handwriting, letters are formed motorically from an internal memory bank, without any external visual reference to their form. In typing, letters are selected from an array of letters. Every letter in the spelling of a word must be scanned for before it can be selected.

All these visual-cognitive demands of word prediction usage impact on the speed at which words can be generated and can increase fatigue while typing. There is the hope that with more accurate predictions in more sophisticated word prediction software, the negative effect of cognitive load which often characterises word prediction use can be more than compensated for.

One legitimate concern about introducing word prediction is that the program itself may be difficult to learn and manage, thus adding an additional burden to the writing process. Since the purpose of prediction software is to reduce the mechanical demands of writing, it has to be relatively easy to learn and use by the students. No record in the literature has been found of subjects struggling to learn how to use word prediction. If a subject can use a word processor, word prediction appears to not be a difficult concept to learn. Williams (2002) reports that her subjects only needed a brief introduction to the software and its features as an initial practice session before they felt comfortable to use it. Her subject, J.T., used the software after one brief training session, and experienced little difficulty from then on. Newell et al., (1992) found that most of the children required only a limited amount of keyboard training, and soon became competent on the program. After 4 one-hour sessions, the subjects were left to use their computers and the word prediction program functionally in the classroom.

2.7 Additional spelling benefits resulting from word prediction usage

There are other spelling related benefits accredited to word prediction. Word prediction not only supports spelling difficulties through word completion, but by helping students to correct errors as they occur. By realising that their predicted word did not look like the word they wanted, some of the subjects would experiment with other possibilities, manipulating the letters produced so far, until the word 'looked right' (1990). Williams (2002) reported a similar tendency in her subject. If he did not find his target word in the prediction list within three strokes, he would 'back up' and try different letters until the word appeared. Word prediction also helps determine the correct word through speech output when it can't be selected visually (Gillette & Hoffman, 1995).

Word prediction has been found useful in the language development of children with spelling difficulties. Zordell (1990) noted an increase in the variety of words used and an improvement in the use of sentence structure, primarily through word prediction's grammar capabilities (Gillette & Hoffman, 1995; Zordell, 1990). Newell et al. (1992) noted improvements in 'stimulation of language', vocabulary size and basic literacy.

Word prediction, with its motivating and effort reduction writing support, may be able to provide the child with the significantly increased amounts of spelling practice he requires to develop the automaticity that comes with long term use (van der Leij, 1999). For successful writing, spelling has to become automatic. As children's understanding and assimilation of the spelling system becomes more developed, their encoding of words becomes more automatic. This frees attention for the real business of writing – the what and how of expression.

A number of additional benefits to using word prediction where there are spelling difficulties have been cited in the literature. Zordell found an improvement in the students' attitude to writing (1990). Klund and Novak (1995) noted improved confidence, improved independence in writing and increased attention span. Williams noted improvement in the speed of being able to produce the sounds necessary to spell words and increased enjoyment in writing (2002). Newell et al. (1992) noted significantly improved independence, confidence, motivation and learning, and, some increase in vocabulary size, concentration span and presentation. MacArthur (1999) noted improved motivation to write, primarily because of the neater, more readable output the child is producing and the less effort required to do so. When some of the mechanical struggles of writing are eliminated, there can be more focus on the content of the writing (MacArthur, 1999; Williams, 2002). Instead of a downward cycle of failure and demotivation, success at written output may result in increased output, with increased mastery of the skill as it is practised and functionally used.

Although fatigue *increases* have been noted due to the increased cognitive demands associated with word prediction usage, fatigue *reduction* has also been shown as an important benefit of word prediction. The source of fatigue in computer text entry may be due to the physical demands of using the keyboard. Keystroke savings reduction is on average 50% in word prediction use, offering a very significant effort reduction (Anson, 1993). If the reduction of effort in using the keyboard is a major consideration for the user, then word prediction can add significant support for the user, provided that speed is not an essential factor (Anson, 1993). Fatigue reduction may also be a benefit from the reduced effort and frustration decrease that comes from word formation support. The cognitive effort in spelling words where there are spelling difficulties can be significant (Gillette & Hoffman, 1995).

2.8 The value of instruction and practice in word prediction usage.

There is sufficient support in the literature to support the benefits of ongoing instruction, and practice, which are considered paramount to the successful usage of a word prediction program. Practice at word prediction appears to have a very significant impact, with users improving significantly as they become familiar with using the program. Practice can be relied on as one potential mechanism for improvement (Koester & Levine, 1998). Practice effects could be seen within a research session (Klund & Novak, 1995), even during the very first trial (Venkatagiri, 1993).

Of the 17 subjects in Newell et al.'s study (1992), only one did not improve in the quantity and quality of written work. Their subjects used word prediction for 6-9 months, were able to use the program for all written work, could take their computers home for homework, had individualised work programs, and had teachers who had been trained in word prediction use. Such generalisation of use is seldom seen in research projects on word prediction.

2.9 Word prediction parameters

The literature survey of the use of word prediction indicated the importance of careful selection of the various parameters to which a word prediction program can be set. *Every* word prediction activity presupposes a selection of parameters, which include factors such as window size, layout and placement location of the prediction list, vocabulary list that the predictions are chosen from, input method, the task chosen for the prediction exercise, the strategy taught, the method of selecting the prediction word of choice, sophistication of the prediction software (recency features, handling of suffixes and prefixes, grammatical prediction etc) and use of auditory feedback. The settings of the word prediction program can impact greatly on the apparent efficacy of the program. Table 2.1 summarises the various settings to which a word prediction program can be set and the impact of those settings. It also indicates the choice of settings that were made for this proposed study, chosen to reflect 'middle-of-the-road' word prediction usage that would neither prejudice nor overestimate the effectiveness of word prediction to support spelling difficulties.

Table 2.1 : Word prediction parameters

Parameter	Description	Literature Survey			Research Project
		Author	Date	Findings	
1. Window size	The number of words predicted at the same time in the prediction list.	Venkatagiri	1994	A balance could be achieved between saving keystrokes and minimising visual cognitive loads by using a five-word list.	A five-word prediction word list was used.
		Swiffin, Arnott, Pickering & Newell	1987	Text generation rate plateaus at a word list length of about 5 words, giving a useful keystroke saving while keeping cognitive load and list scanning to a minimum. The predictions can be seen as a whole, minimising the scanning necessary.	
2. Orientation	The orientation of the predicted words on the screen eg vertical, horizontal	Swiffin, Arnott, Pickering & Newell	1987	A vertical layout of the prediction list will reduce search time. A large amount of head and eye movement is necessary in a prediction list in the horizontal orientation.	The prediction word list appeared in a vertical layout.
		Tam, Reid, Naumann & O'Keefe	2002	With a vertical layout, search time is reduced because the head and eye movements is kept to a minimum.	
		Klund & Novak	1995	A vertical word list may be easier to scan visually than a horizontal word list, as well as easier to process cognitively, although this depends on word order.	
3. Vocabulary list	The list of words from which the predicted words are chosen, usually based on frequency of words used.	MacArthur	1999	Younger students may find it easier to use versions of word prediction with limited vocabularies.	The Default Lexicon of the prediction software, which is made up of the 2000 most commonly used words, was used. As the aim of the research was not to maximise typing input speed, but to correlate typing speed and accuracy with spelling ability, the most suitable vocabulary was a list from which two equivalent word lists with a wide range of spelling words could be chosen.
		Klund & Novak	1995	A relatively small dictionary containing the highest frequency words is effective.	
		Swiffin, Arnott, Pickering & Newell	1987	• A relatively small number of different words make up 'written text' eg, in Brown's corpus frequency list, top 10 words account for 24.6% of written text, and the top 1000 words account for 69.5% of written text. • Different contexts use different vocabularies. It may thus be more appropriate to store a number of small dictionaries rather than a large one. • A pre-built dictionary can increase the keystroke saving from about 50% to 70%.	
		Higginbotham	1992	A fixed dictionary, having a stored vocabulary that never changes, may be more predictable to the user, and become more automatic over time.	
		Zordell	1990	Word selection was facilitated by customising the vocabulary list for each student. They had their own most frequently used words on their personal copy of the prediction vocabulary list.	
		Anson	1993	Lists for younger children of 500 words can provide the words for 86% of their vocabulary needs in any one day. With a single environment the vocabulary needs of a person seldom exceed 2000 words.	
		Koester & Levine	1997	A parametric analysis predicted that a dictionary of 1000 words would result in poorer performance than a dictionary of 500 words because the increase in visual search and scanning times for longer word lists affected the improvement in word prediction usage.	
		Venkatagiri	1993	• No single dictionary can adequately serve the needs of a variety of users. • A system based on word prediction must create a unique dictionary for each user to be maximally effective.	

Parameter	Description	Literature Survey			Research Project
		Author	Date	Findings	
4. Adaptive dictionaries	Vocabulary can be 'learned' from the user and incorporated into the vocabulary bank of the prediction software. New words are absorbed and new frequency ratings are assigned to the vocabulary list corresponding to the user's word usage.	Higginbotham	1992	An adaptive dictionary may be able to generate the desired word sooner since it may be capable of learning the user's typing style.	To control the equivalence of the two wordlists for the research project, the feature of automatic adaptation of new words, although clearly very useful in word prediction usage, was switched off. It would reduce equivalence of the two word lists if the prediction lists introduced new words and shifted words into new positions in the predicted lists.
		Koester & Levine	1996	Adaptive lists may increase keystroke savings but they reduce the user's ability to learn the lists' contents, which in turn increases the search time.	
5. Frequency and recency promotion	All words in the vocabulary list have a frequency rating which is automatically modified during use, so that words more frequently used by the user are placed at the top of the dictionary. As any word is typed it is entered into the vocabulary lists with the maximum value of recency. Recency takes precedence over frequency.	Swiffin, Arnott, Pickering & Newell	1987	The balance between recency and frequency may significantly affect performance for long term users.	Frequency ratings change as words are entered during the research task. Because of the recency feature of the software, any words typed appear at the top of the prediction list for future words with the same initial letter. Subsequent users would find the target word at the top of the prediction list for every word the previous user spelled correctly, which is clearly unacceptable for the research task. Each subject was therefore given a 'clean' lexicon at the beginning of the research task, so each subject was presented with the predicted words in the same order, and wrongly spelled words or new words were not absorbed into the vocabulary lists.
6. Location	The placement location on the screen of the prediction list eg. top/bottom, left/right	Tam, Reid, Naumann and O'Keefe	2002	Suggested a reconsideration of the new trend in prediction software to have the list follow the cursor, as it was in this position that their subjects scored the lowest accuracies. The most effective position appeared to not be definitive, but to be correlated to the most favoured position of the individual user.	A column of large, easily accessible cells held the predicted words – to the right of the screen (so as to follow the pattern of left-right progression in word construction). This type of layout for the predicted words has not been met in the literature review.
7. Word order	Word lists can be ordered alphabetically, by word length, or statistically.	Swiffin, Arnott, Pickering & Newell	1987	Ordering the words in descending recency and frequency is beneficial.	Words were automatically ordered in descending recency and frequency by the word prediction software.

Parameter	Description	Literature Survey			Research Project
		Author	Date	Findings	
8. Text entry method	Text entry can be through a conventional keyboard or through an on-screen keyboard.	Anson	1993	• Typing speed varied a great deal between subjects typing on the conventional keyboard, suggesting the impact of familiarity on typing speed. • On-screen keyboard typing was much slower than conventional keyboard typing. • Typing on a conventional keyboard with word prediction was always slower than typing without word prediction, irrespective of the user's typing rate on the conventional keyboard. • Typing on an on-screen keyboard with word prediction was a little faster than typing without word prediction. (Note: Anson's subjects did a copy task).	An on-screen keyboard was used for text entry. The range of familiarity with the conventional keyboard was expected to be large in the sample group, and would have a significant impact on input times. All the subjects were equally unfamiliar with the on-screen keyboard. Also, subjects with slow typing speeds were excluded from the research sample.
		Newell, Arnott & Waller	1990/ 1992	For very slow conventional keyboard operators (ie less than 5 words per minute) there is a doubling of speed.	
		MacArthur	1999	Typing could prove to be a barrier to the speed of written output if the child is not practised at typing.	
9. Task	The writing task, which can range from single word entry, through to copy writing, dictation, sentence writing or free composition.	Tam, Reid, Naumann & O'Keefe	2002	The copy task introduced much eye gaze, shifting, forgetting where they were and errors as they skipped words.	As the focus of this research project was on the effectiveness of word prediction in relation to spelling ability, a single word entry task was chosen, so as to minimise the impact of other cognitive demands that may be required in other writing tasks.
		Szeto, Allen & Littrell	1993	Speed depends on the task performed.	
		MacArthur	1991	The effectiveness of results depended, amongst other things, on the design of the task and the match between the tool and the writing task.	
10. Text	The material which has to be typed.	Lesher, Moulton & Higginbotham	1998	• Typing texts too similar or dissimilar to the statistical source corpus may provide skewed estimates of performance. • No single text is representative of all written communication.	The text for the research project was single words. All the words in the task's wordlists appeared in the Default Lexicon of the prediction software, except for 3 words (out of 30) which were introduced to simulate the 'never there' condition in the prediction lists, and 7 words that were added into the Default Lexicon.
		Venkatagiri	1993	It must be ensured that the material is not so esoteric as to render the test useless, or so biased in favour of the dictionary used for the prediction lists as to invalidate the test. The linguistic corpus used in the research test should be random and yet possess face validity.	
11. Strategy	The strategy taught of when searching the prediction list should begin eg. before the word is typed (if there is grammar prediction), after the 1st letter of the word is typed, or after the 2nd letter of the word, or after the 3rd letter is typed.	Koester & Levine	1996	The strategy of searching the word prediction list only after typing the first two letters was shown to be the most effective generalised strategy, having a lower cost with respect to item selection rate. If the user's search time is slow, a strategy that requires fewer searches is more advisable.	As training in the research project was limited, the subjects were only given rudimentary strategy training. They were shown how to look at the list after each letter was typed, but it was pointed out that the target word could appear at any stage during the typing of the word, so they should not spend too much time searching the list if the word does not appear to be there.
		Klund & Novak	1995	A good 'all purpose' strategy is to type one letter and then search the word list. However, depending upon the speed of keypress time and list search time, another strategy may be more efficient.	
		Koester & Levine	1998	• Strategy is one of the few factors that a user or clinician can control. • The system should be configured to support this strategy (by only presenting the prediction list after the 2nd letter is typed)	

Parameter	Description	Literature Survey			Research Project
		Author	Date	Findings	
12. Grammar prediction	More powerful prediction software with grammar prediction capabilities will offer the most likely next words and words in the most likely form corresponding to the previous words in the sentence, through application of grammar rules and word pair frequencies.	Koester & Levine	1998	Word prediction which incorporates grammar prediction capabilities can increase keystroke savings by about 5% and could lead to a noticeable increase in text generation rate if the list search time remained constant.	The grammar prediction capabilities were set to zero. Grammar prediction and handling of suffixes and prefixes were irrelevant in the single word entry task of this research project, where each word was independent of the previous word, and where the subject was required to type or select the exact form of the requested word. The various forms of a word were offered as independent words, or the target word could be manipulated into the correct form required in the research task.
13. Space savers	An automatic space is added after a word is inserted from the predicted list. Automatic capitalisation occurs after a fullstop. Proper nouns are stored in the vocabulary lists with their capitals. Abbreviation expansion allows phrases to be stored in abbreviated form.	Hunnicutt & Carlberger	2001	• All these features save keystrokes and thus increase input rate. • Abbreviation expansion requires memory skills to remember the abbreviation codes, but can be used to bypass specific spelling difficulties.	None of these features could be utilised in the research project with its single word entry task.
14. Auditory feedback	Auditory feedback can be received for every letter, word or sentence that is typed.	Tam, Reid, Naumann & O'Keefe	2002	Word prediction in combination with auditory feedback was more effective in the reduction of spelling errors than was word prediction alone. However, the rate of text entry was slower than when word prediction was used without auditory feedback.	Auditory feedback was switched off, as spelling skills were investigated. Use of auditory feedback would distort the results, as it could bypass the skills required to form and recognise a word. However, it was recognised that auditory feedback could be a powerful strategy to increase the number of correctly spelled words.

Parameter	Description	Literature Survey			Research Project
		Author	Date	Findings	
15. Font size	The size, and thus ease of recognition and/or access, of the predicted words. Most often, the window has predicted words in a font size similar to that of the typing font.	Paucity of research. No studies were found that specifically addressed these aspects.			The words on the prediction list were very large and readily accessible, thus minimising co-ordination skills required to access the words and maximising ease of visual location of the words.
16. Method of selection	The method of selecting a word from the prediction list - by code or direct click	Paucity of research. No studies were found that specifically addressed these aspects.			Direct selection through a mouse click was the selection method for the research project. Although no comment on this aspect was found in the literature, word entry by code would almost certainly increase cognitive load. The user must note the code next to the target word, remember the code, find the code key (which is found on the keyboard) and press the relevant code key. Also, the use of an on-screen keyboard and a mouse for text entry, the method of text entry in the research task, makes mouse clicking directly on the chosen word very natural.

2.10 Summary

In this chapter it was shown how the features of word prediction are suited to support spelling difficulties. Spelling accuracy is primarily enhanced through the word formation support that word prediction, through its word completion feature, offers. Spelling support is also through the unique interactive process combining spelling and reading skills, a process which finds much theoretical support in the literature. Typing speed is increased through the keystroke reduction feature of word prediction, although this aspect of word prediction usage has not realised the apparent potential. The use of word prediction is often at the cost of increased visual-cognitive demands, resulting in time and fatigue increases.

There are other spelling related benefits that word prediction offers, such as spelling development, language development, increased motivation for writing, increased writing output and fatigue reduction.

Although there are significant limitations to the research project, even modest increases in spelling accuracy and typing speed in the research task would lend support to offering word prediction to children with spelling difficulties, as the literature supports a marked improvement in word prediction usage with training and practice.

However, there was no indication in the literature of the level of spelling literacy required to benefit from word prediction or whether young children even have sufficient cognitive resources to operate word prediction. The necessity of supporting young children with spelling difficulties in their writing requirements, and the potential benefit of word prediction to provide the tool, together with the lack of information available upon which to make informed decisions, motivated this study.

A literature survey of the parameters to which a word prediction can be set, and which impact on it effectiveness across a variety of user characteristics and activities, are noted and applied to the proposed study.

Chapter 3

Research Methodology

3.1 Introduction

In this chapter, the methodology of the research project is explained. This project was designed to investigate the relationship between spelling ability and the efficacy of using word prediction to improve spelling accuracy and on-screen keyboard typing speed.

3.2 Research Question

Can word prediction improve the on-screen keyboard typing spelling accuracy and increase the on-screen keyboard typing speed of Grade 4 – Grade 6 children with spelling difficulties?

In order to answer this research question, 3 research questions were formulated.

Sub-question 1	Does the use of word prediction improve spelling accuracy?
Sub-question 2	Does the use of word prediction increase typing speed?
Sub-question 3	Is there a relationship between spelling ability and improvement in spelling accuracy and typing speed with word prediction?

3.3 Research Design

The research project was a cross-over within-subject counterbalance design using multiple subjects. Counterbalance designs are often used where there are more than one tests and when the order of participation varies (DePoy & Gitlin, 1994). The task for this study consisted of two subtests – Typing Only subtest (TO) and Word Prediction subtest (WP), and two Wordlists – Wordlist A and Wordlist B. The sample was divided into four groups, each group executing the research task in a different combination of order of subtest and wordlist used. Table 3.1 outlines the differences between the four groups.

Table 3.1 : The four sample groups

Group 1	**Group 2**
Typing Only subtest (TO) + Wordlist A followed by Word Prediction subtest (WP) + Wordlist B	Typing Only subtest (TO) + Wordlist B followed by Word Prediction (WP) subtest + Wordlist A
Group 3	**Group 4**
Word Prediction subtest (WP) + Wordlist A followed by Typing Only subtest (TO) + Wordlist B	Word Prediction subtest (WP) + Wordlist B followed by Typing Only subtest (TO) + Wordlist A

The cross-over design was chosen to counter-balance the possible effect of differences between the two wordlists and the effect of the presentation order of the two methods of text entry (with and without word prediction). The wordlists could be logically equated on many criteria in their formulation, but, especially with respect to word prediction, it was not known if all the criteria for equality had been identified. Also, it was not known if the experience the subjects gained with the on-screen keyboard during the first subtest would impact on the results of the second subtest.

To form the four groups, all the children who passed the selection criteria were sorted in ascending order, within each grade, on the basis of their scores on the Graded Word Spelling Test (GWST). They were systematically divided into the four groups, starting with the child scoring the lowest score in Grade 4 and divided off one by one into the four groups ending with the child with the highest score in Grade 6 (see Appendix A).

3.4 Pilot Study

3.4.1 Subjects

Most of the subjects for the pilot study were drawn from a private mainstream school in March 2003. The headmaster and parents of the school were informed of the nature of the project, and their permission obtained to conduct the pilot study (see Appendix B and

Appendix C). The choice of this school was a practical one, the school being a local one where contacts were in place.

The children in the pilot study were from the Grade 2 to Grade 5 classes, although the sample group for the main study would be drawn from the Grade 4 to Grade 6 children i.e. children in the first three grades after completion of the Foundation Phase of schooling (see section 3.4.2.1). Some of the children from the main study sample were expected to have a spelling ability significantly below that expected of their grade, even possibly having a spelling grade level of a Grade 1 child. However, it was inappropriate to pilot the performance of Grade 1 children, as they had not yet been introduced to the entire alphabet at the stage of the pilot study. It was unnecessary to pilot the performance of children spelling higher than a Grade 5 level, as the sample group's Grade 6 spellers with a spelling ability of higher than a Grade 5 level would not have a spelling difficulty.

In addition to the 14 private school children, 5 other private mainstream school children between Grade 2 to Grade 5 were also used. In total, there were 19 children in the pilot study.

3.4.2 The development of the wordlists

Before the pilot study could begin, 2 equivalent wordlists had to be compiled. All the words, except for 7 words (for exceptions, section 3.5.4.4.v) were chosen from the Penfriend Default Lexicon of children's 2000 most commonly used words (Spooner, 1999). See Appendix D for the full vocabulary list, on which is indicated the words chosen for the research project.

Two equivalent word lists of 40 words were drawn up. One list was to be used with the Typing Only (TO) subtest and one for the Word Prediction (WP) subtest. The wordlists were designed to maximise equivalence with respect to the range of phonetic principles and the number and length of the words.

3.4.2.1 Range of phonetic principles

At the end of the Foundation Phase (the first three grades of schooling in South African education), a learner is expected to have writing skills that can be used for functional purposes of message generation (see Appendix E). That is, most of the phonetic principles needed to form words have been introduced to the children. However, a significant difficulty was met in attempting to define 'Grade 3 literacy' in terms of phonetic development. There was no national or provincial set curriculum with the South African Education's Outcomes Based Education (OBE) approach to education. Every school devised their own curriculum, based on the national guidelines as set out in the Specific outcomes of the Foundation Phase.

The process followed to select the words for the two wordlists, as well as the words chosen for the wordlists is outlined in Appendix F. It involved a compilation of the syllabi of three schools and the selection of words to cover all the major phonetic principles covered by the end of Grade 3.

Note: At the beginning of 2004, after the completion of the test phase of this study, a new curriculum was introduced into South African schools, the Revised National Curriculum Statement (RNCS). Although still based on OBE principles, the new curriculum is a little more formalised than the old one.

3.4.2.2 The number and length of words

Table 3.2 indicates the 40 word pairs that were selected for the pilot study. They are arranged in columns of word length, to indicate the distribution of the size of the words. The average word length of the words in the wordlists is 5.6 characters. The average word length in text is between 5 and 6 characters per word (Anson, 1993). A weight of long words would overestimate the influence of word prediction, and a weight of short words would underestimate the influence of word prediction.

Table 3.2: The number and length of the words – pilot study

3 letter words		4 letter words		5 letter words		6 letter words		7 letter words		8 letter words		9 letter words	
Word List A	Word List B	Word List A	Word List B	Word List A	Word List B	Word List A	Word List B	Word List A	Word List B	Word List A	Word List B	Word List A	Word List B
cry	dry	tall	ball	fairy	hairy	letter	ladder	married	worried	horrible	terrible	opera-tion	injec-tion
tie	pie	fire	wire	white	wheel	yellow	follow	picture	mixture	polished	finished		
eye	one	talk	walk	chief	thief	garden	carpet	witches	matches	furious	curious		
		hurt	turn	fruit	blood	rocket	packet	weather	feather	mountain	fountain		
		push	pull	light	right	string	strong	quietly	quickly	sleeping	sweeping		
		tail	jail	alone	awake	bridge	fridge	village	message				
		cake	coke	noise	voice	circle	circus						
				crown	clown								
				dress	cross								
				knock	knife								
				break	great								

3.4.3 The objectives, findings and recommendations of the pilot study

The data results of the pilot study are presented in Appendix G. A summary of the objectives, findings and recommendations of the pilot study follows.

Note : In the research task, the children would do one subtest with Wordlist A and one subtest with Wordlist B. However, in the pilot study, different combinations were tested. Some children were used to test the equivalence of the two word lists – they did the TO subtest with both lists. Some children were used to assess the difference between the TO subtest and the WP subtest – they were given the same wordlist for both subtests. Some of the children, who were willing and did not appear tired, completed *both* word lists using *both* the typing only and word prediction subtests. Through this method, a fairly comprehensive view of the relationship between the two word lists and the two subtests was gained.

3.4.3.1 To develop two equivalent wordlists for use in the main study with respect to word difficulty

The number of words and word length were precisely defined and equivalence in these areas can be seen in Table 3.2. Word difficulty equivalence was somewhat less definitive, and the pilot study was used to investigate this area, by highlighting the most unmatched pairs. Only scores where the children completed both wordlists in the typing only method of text entry were considered in the assessment of equivalence of the two wordlists.

Although, on average, there was little difference in the accuracy scores between the two word lists (Wordlist A elicited 319 correct responses, whereas Wordlist B elicited 324 correct responses), there were significant individual differences between accuracy scores on the two wordlists. Children scored from between –4 to +6 correct score differences between the two lists. See Appendix H for a table of all the pilot study accuracy scores.

It was considered that the source of the variation between the scores that subjects showed on the two lists could be due to *unmatched pairs*. Unmatched pairs are words that, although chosen to be equal in terms of spelling difficulty level, may not be as closely matched as they appear to be at face value. Some children would score some word pairs equally (that is, both correctly or both wrongly) whereas some children would spell one of the words correctly and one of the words wrongly. The variance between individual children's scores on the two word lists ranged from 0 differences to 12 differences (see Appendix I). The record form for the children's responses during the pilot study highlighted any unmatched scores. An example of the record form, with the scoring details of the child who scored 12 unmatched pairs can be seen in Appendix J. Examples of unmatched pairs from the pilot study were word pairs such as *cake/coke*, *tie/pie* and *dress/cross* where subjects did not apply the same phonetic principle to both of the words. Some subjects spelled *tie* correctly and *pie* incorrectly, and some spelled *tie* incorrectly but *pie* correctly. The results of the pilot study warned that it could not be assumed that words using the same phonetic principle would be spelled similarly. The reason for this variation was not obvious but was sufficiently meaningful to be taken into consideration in the compilation of the final two wordlists.
Each of the 40 word pairs used in the pilot study was ranked according to the number of times a word pair was unmatched. To increase the equivalence of the two lists, the words that received the 10 highest unmatched scores were removed in the compilation of the final word lists (see Appendix I).

It was noted that between the two wordlists there was a greater score difference in terms of the order of presentation, than there was in terms of the wordlists themselves. The wordlists presented first elicited 328 correct responses and the wordlists presented second elicited 313 correct responses. This suggested that fatigue could negatively influence the performance on the second list.

3.4.3.2 To achieve a wide and well-distributed variance range in the spelling results.

The number of times each word in the two wordlists was scored correctly was recorded (see Appendix K) and plotted onto a distribution graph – Figure 3.1. Only 1 word was so difficult that only 1 of the 13 subjects who spelled both Wordlist A and Wordlist B spelled it correctly. Likewise, only 1 word was scored correctly by all 13 the subjects. The other words were of a range of difficulty that approximated the normal curve. The words then appeared to represent an acceptable range of difficulty that could elicit a wide range of correct scores in the main study. Also, the floor and ceiling of the word lists appeared to be sufficiently represented by easy and difficult words.

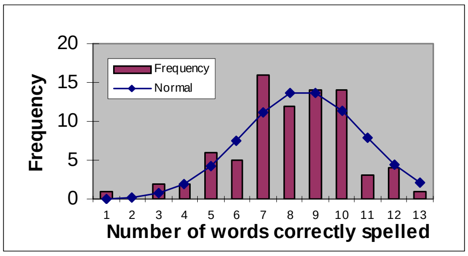

Fig 3.1 : Frequency graph of correctly spelled words – pilot study

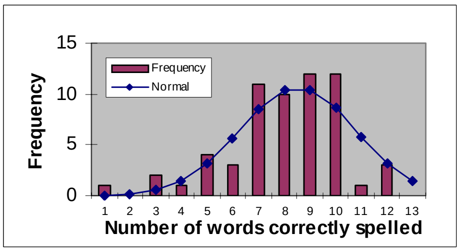

Fig 3.2 : Frequency graph of correctly spelled words – reworked after removal of 10 words

It was important that the balance in the words seen in the above graph be retained after the removal of the 10 word pairs to form the final wordlists for the main study. Figure 3.2 indicates the new distribution graph that resulted after the removal of the 10 most unmatched words. It can be seen in the new distribution graph that the ceiling of the wordlists would not be significantly impacted by the above changes, but a possible area of concern was that there was not a stronger representation of the ‘easier’ words. A number of the unmatched pairs that had to be removed represented the easier words.

3.4.3.3 To test the Word Prediction subtest, with respect to spelling accuracy improvements and speed enhancements that could result from the use of word prediction, the equivalence of the wordlists and the length of the task

There was a strong indication that word prediction could have a significant impact on spelling accuracy. The greatest score improvement from typing without word prediction to typing with word prediction was from 4 correct scores to 15 correct scores, an improvement of 11 correct scores (see Appendix G).

The impact of word prediction on typing speed enhancements was not so clear, as time results for subjects who were tested with both TO and WP on the same wordlist were available for only 2 subjects. However, it was noted that the weaker spellers worked very slowly through the WP subtests.

The target of 30 words for the main study appeared a suitable choice, as the 40 words in the pilot study tended to tire the children.

Two significant elements of poor list design and non-equivalence between the two lists were uncovered and were addressed in the final compilation of the word lists for the research task. Firstly, Wordlist A had a significant advantage over Wordlist B in terms of the minimum number of keystrokes possible to spell the words. Words from Wordlist A could be typed using a minimum of 153 keystrokes, whereas the minimum number of keystrokes required to type Wordlist B was 173. The reason for the discrepancy lay in that Wordlist B had a disproportionate number of words that did not appear in the prediction list, and therefore could not be accessed for word completion before the word was fully spelled letter by letter. The wordlists for the research task in the main study would have to be carefully selected to equate the minimum keystrokes required exactly, through systematic analysis of each word of how many letters were required to be typed before the word appeared in the prediction window. Secondly, the words in the word lists were not well distributed in terms of spreading the easier and more difficult words out from one another, and the unpredicted words appeared in a cluster, rather than being well spread throughout the word lists. It was necessary to have the easier words, the more difficult words and the unpredicted words more evenly distributed throughout the word lists for the research task. Appendix L indicates all the above features of the two wordlists as they presented in the pilot study.

3.4.3.4 To test the keyboard and word prediction training programs

The five words given for the training of the on-screen keyboard appeared more than sufficient. Most subjects were fully confident with the on-screen keyboard after five trials. However, all five of the words would be retained for rudimentary keyboard practice.

The word prediction training was of primary importance in the pilot study. This was because the use of word prediction was essential to the success of the research project. The pilot study had to ascertain if there was reasonable validation for the assumption that children as young as Grade 4 – 6 would be able to use word prediction independently (there was a paucity of research in this regard).

It was decided to retain the word prediction training program as initially designed. The pilot study indicated that there was reasonable expectation that most of the children in the main study would be able to grasp the principle of word prediction. The few children in the pilot study that did not cope with word prediction represented the weakest of Grade 2 and Grade 3 children in a mainstream school. Although the sample children would have spelling difficulties, they would be older children and in Grade 4 or above. There was the concern, however, of what to do should a subject in the main study appear to not understand the strategy of word prediction after the training was completed. It was decided to introduce a small test at the end of the word prediction training, using 5 very simple words. If any of the subjects failed to find at least 4 out of the 5 words in the prediction list correctly, they would be disqualified from the sample.

3.4.3.5 To investigate the time aspects of the task

The pilot study gave an indication of the time allocation to allow for the various sections of the research test.

1.	Welcoming and orientation to research task	2 mins
2.	On-screen keyboard use training	2 mins
3.	TO subtest	About 15 mins
4.	Closing	1 mins

5.	Word prediction training	About 10 mins
6.	WP subtest	About 15 mins
7.	Rest	10 mins
	Total	About 55 mins

The 15 min time allocation for the 2 subtests appeared to be a generous one. The average time for the subtests was expected to be closer to 11-12 minutes, but it was recognised that there would be significant differences between the time taken by the children to complete the research task.

A period of rest and refreshments between the two subtests needed to be introduced because of the fatigue some children indicated. The score difference between the first and second tests administered indicated higher accuracy scores for the first test administered (see Appendix H)

That the time scores decreased markedly between the time scores on the first and second lists administered was a strong indication that there may be practice influences to consider (probably related to the on-screen keyboard practice) (see Appendix H).

3.4.3.6 To test the procedural aspects of the test administration

The SKin keylogging program planned for the recording of the children's test data proved unsuitable. It worked well during the Typing Only subtest, but was incompatible with the word prediction software. It was decided to replace the internal computerised timing of the Skin keylogging program with manual timing with a stopwatch. However, manual timing introduced new levels of concentration and accuracy on the part of the researcher. The test procedure for the researcher would have to incorporate additional timing instructions and tester reliability issues would have to be addressed. Each test would have to be video taped and a professional would have to assess the integrity of the testing procedure.

The research task was originally planned to run automatically, with the next spelling word presented immediately upon entering the current word. A digital voice would call out the next word to be typed and a picture of the word would appear to ensure rapid comprehension of which word was required. The digitised voice resulted in some word identification problems, as did some of the pictures. Although those problems could be addressed through recordings of the voice and different pictures, it was decided to do away with the digital voice and the pictures. The reason was because the program of interlinked grid formats proved unstable in the word prediction subtest, due to the keylogging program; the screen would freeze unpredictably. Such instability was an unacceptable risk in the test administration process. The test program would need to be redesigned without the interlinked grid format. It was decided that the words would be called out by the researcher, just as in a conventional spelling test, a format well known to the children.

3.4.3.7 To test word comprehension and familiarity

Some occurrences of word confusion were noted with homonym words. It was decided to introduce descriptive phrases after the calling out of the words, to increase the comprehension of exactly which word was required by placing the word in a context. This procedure would be exactly the same as that used in the Graded Word Spelling Test (see section 3.5.4.2).

Only words that were considered familiar to the children were chosen for the word lists. All the words, except for few that were added (see section 3.5.4.4.v), were words drawn from a list of the 2000 most frequently used words (Spooner, 1999). No words were requested to be explained by the children during the pilot study. It can thus be reasonably expected that there were no familiarity difficulties with the words in the wordlists. However, a difference between the pilot study and main study sample group is acknowledged, as the pilot study sample was drawn from a middle to upper social group, whereas the main study sample would be drawn from a wide range of socio-economic groups (see section 3.5.1).

3.5 Main Study

3.5.1 Context

All the subjects were drawn from the pupils of a school for learners with special needs. Most of these pupils had been withdrawn from mainstream schooling because of poor academic progress in one or more of the following areas of educational performance – reading, writing or mathematics. Many had unspecified diagnoses although some had the specific diagnosis of Learning Disability. Some of the children, who had been in special education since the beginning of their schooling, had Cerebral Palsy. Cerebral Palsy is primarily characterised by motor difficulties (with visual and hearing problems often accompanying the motor difficulties), but often has associated poor academic progress. It can thus be seen that the children did not represent a uniform group. The common factor however was significantly poor educational progress (in one or more areas of educational performance).

It is important to note that the Grade levels at this special school are not an accurate reflection of mainstream grades. There is a wide range of educational maturity seen in each of the grades. For example, the Graded Word Spelling Test raw scores of the Grade 6s ranged between 6 and 46 correctly spelled words, representing a spelling age range of 6-02 yrs to 12-10 yrs. This is partly explained by the National Education Policy (1996) which discourages grade repetition, despite lack of educational progress.

> In principle, learners should progress with their age-cohort. Repetition of grades seldom results in significant increases in learning attainment and frequently has the opposite result. The norm for repetition is one year per school phase where necessary. Multiple repetition in one grade is not permissible.

There was also a wider age-range than is typically the case in mainstream schools. For example, at the time of the study, the Grade 4s ranged in age from 9-03 yrs to 12-06 yrs, an age range of 3-03 yrs. This could be due to factors such as late school starting ages because of lack of school readiness at the appropriate age, or lack of school facilities in the child's area (South Africa is addressing a massive backlog of children who did not previously attend schools), and grade repetitions that bypassed the official policy.

The classes from which the sample for the current research were drawn also reflect a wider range compared to mainstream schools of factors such as the educational standards of the previous schools from which the children came, mother tongue differences, cultural differences, economic levels, parental educational levels and computer experience.

3.5.2 Subjects' selection criteria

Table 3.3 outlines the selection criteria for the subjects for the study, and indicates how the criteria were assessed. The table also includes some comments that explain the importance of the criteria to the study.

Table 3.3 : Subjects' selection criteria

	Selection Criteria	Method	Comment
1	School Grade 4-6	School registers	Grade 4 presupposes a minimum literacy level of Grade 3 (the Foundation Phase of literacy) where there has been exposure to most of the basic spelling rules.
2	Reported spelling difficulty	Graded Word Spelling Test (Vernon , 1998)	Spelling difficulty was the criterion by which children in the above classes were selected (see section 3.5.4.2 for a discussion on this test and section 3.5.5.2 for the criteria by which spelling difficulty was judged).
3	English schooling experience of at least the past 2 years	Children questioned as to previous school attendance	It would be unfair to test spelling ability on a child who has limited understanding of English vocabulary and experience in the spelling of English words.
4	No previous experience with word prediction	Children were asked, *"Have you ever used a computer program where you don't have to type the whole of a word because the computer helps you by guessing what you want to type and then completes the word for you?"* Their yes or no answer was accepted as fact.	Although some of the children might have had experience with the word prediction found in cell phone SMS technology, experience in it did not result in disqualification from the project, as SMS technology operates on a different system to computer word prediction.

	Selection Criteria	Method	Comment
5	Absence of specific difficulties that may impact on the execution of the research task.		The origin of and specific nature of the children's spelling difficulties were not under scrutiny in this project, so all the children, irrespective of their diagnosis, were considered eligible for selection. Only children with problems limiting their *execution* of the research task were disqualified from selection. These were functional hearing, attention, visual and motor difficulties.
	a. Hearing difficulties	The children with significant hearing difficulties were identified by the Speech and Hearing Therapists at the school. For the purposes of this research project, functional hearing was considered adequate if the child was able to hear verbal directives from the teacher without lip reading. The teacher was requested to judge if the child had functional hearing.	The research task required the researcher to sit next to the subject and call out words to be spelled while the subject focused on the screen. Hearing difficulties would result in time wastage as words would have to be repeated, or there could have been misinterpretation of the words.
	b. Attention difficulties	Information was gathered from the school psychologist, nurse and teachers to ascertain which of the children had uncontrolled attention difficulties (inability to independently complete concentration demanding school tasks taking longer than 10 mins).	As the research task was timed, and required maintained concentration over about 15 mins, it was important that the subjects be able to maintain focus on the task for the duration of the task. Children with uncontrolled attention difficulties were not eligible for selection.
	c. Motor and visual difficulties	Subjects scoring at speeds significantly longer than their peers on the Mouse Control Screening were considered to have co-ordination skill difficulties and/or insufficient visual skills for the research task. (see Appendix M for the test and Figure 3.3 for the results of the test)	Motor and visual difficulties could impact on functional mouse control and speed of letter identification on the screen. As timing was an issue in the research project, it was considered important to control for. Although ease of access to the on-screen keyboard of the research task was maximised and visual demands minimised (see section 3.5.4.1), a mouse control screening was administered to all the children.
6	Parental permission	Signed reply slip by parents was required.	Ethical obligations required that parents be informed of the research project and sign permission for their child's participation.

Table 3.4 summarises the number of subjects who qualified for the research sample group. The table indicates the number of subjects who were available for selection, how many subjects were disqualified on the basis of the selection criteria, and for what reason they were disqualified.

Table 3.4 : Summary of subjects qualifying for the research sample group

Current Grade	Actual number available	Number disqualified on the grounds of …..							Total	Total sample group
		Parental permission denied	No spelling difficulty	Non-English schooling experience	Word prediction experience	Specific difficulties				
						Hearing	Attention	Motor and visual difficulties		
Grade 4	29	0	5	0	0	0	2	0	7	22
Grade 5	32	1	3	0	0	(1)	0	0	4	28
Grade 6	38	1	3	0	(1)	2	0	1	7	31
Total	99	2	11	0	0	2	2	1	18	81

Note: The numbers enclosed in brackets represent subjects who were disqualified on more than one ground. Numbers in brackets were thus not included in the various totals.

3.5.3 Subjects' description criteria

3.5.3.1 Gender distribution

The classes from which the subjects were chosen were weighted with boys relative to girls. Although it has been shown that girls, in general, spell better than boys (Allred, 1990), it was decided to not make gender a selection criterion for this study. Each child would be compared to himself with respect to speed and accuracy, and not to group norms. For reference value, the relative number of boys to girls in the classes and in the final sample group are indicated in Table 3.5. The gender ratio (boys to girls) in both the class and the sample was 2:1. The sample group can be seen to be comparable and representative with respect to gender.

Table 3.5 : Number of boys and girls

Grade	Boys		Girls	
	Actual in class	Sample	Actual in class	Sample
Grade 4	22	16	7	6
Grade 5	21	20	11	8
Grade 6	24	21	14	10
Total	67	57	32	24

3.5.3.2 Age distribution

The ages of the children in the three grade classes at the school used for the research study have more variance than is usually found in mainstream grades. The average national age variance as stipulated in the South African School's Act (1996) is :

> The statistical age norm per grade is the grade number plus 6.
> Example: Grade 1 + 6 = age 7

Table 3.6 indicates the age variance in the three grades from which the subjects were drawn, and compares that to the average age variance as stipulated in the South African School's Act. The age range in the grades in the school used for the research study is clearly far wider than the national average.

Table 3.6 : The ages of the subjects

Grade	Youngest (yr – mth)	Oldest (yr – mth)	Variance (yr – mth)	Average national age variance (yr – mth)
Grade 4	9-03	12-06	3-03	10-00 to10-11
Grade 5	10-07	13-04	2-09	11-00 to 11-11
Grade 6	11-07	13-10	2-03	12-00 to 12-12

Gender, age and other subject characteristics, are probably important variables in the successful use of lexical prediction (Venkatagiri, 1994). As subject characteristics other than the characteristic of spelling ability were not the focus of the present study, they were not included in the selection criteria, but are important for descriptive purposes. The impact of gender and age on the results was investigated in the analysis of the results (see section 4.3)

3.5.4 Materials and Equipment used in the study

3.5.4.1 Computer

A laptop computer with Windows 98 (1.2 GHz) was used for the Mouse Control Screening and the main study.

i. On-screen keyboard

The keyboard for the experimental task was designed using Clicker 4.1.72 software, a program that can 'send' information from customised cells into a word processor from an on-screen keyboard. The on-screen keyboard was chosen for the research task because of the varying experience with keyboard typing that may exist among the sample and the possible influence of those varying levels of familiarity (see parameter 8 in Table 2.1). The subjects were drawn from a wide range of economic backgrounds, some being exposed to computers from a young age in their homes, others only experiencing computers in the school context (see section 3.5.1).

No punctuation was included in the experimental design, as that skill goes beyond simple word formation. The task was designed for maximum ease of operation and minimum room for error or disturbance to the computer program. The letters were arranged alphabetically rather than in QWERTY formation. The only keys available were the 26 Roman alphabet letters, the Enter icon and the Delete icon. The colours were chosen to maximise visual clarity of the letters and words. The font was clear, large, bold, black, Comic Sans, font size 36.

The screen layout was divided into two parts – a very large on-screen Clicker grid keyboard filling the lower 4/5ths of the screen and on the upper 1/5 of the screen, a blank space to type into. For the TO subtest, the keyboard was on the left, with the 26 alphabet letters arranged in 4 rows of 5 letters, and one row of six letters. Next to the keyboard was a large 'delete' icon and an even larger 'enter' icon above the 'delete' icon. The screen for the WP subtest was essentially the same, but in addition had a column of five large cells, in which the predicted words appeared. These cells were positioned between the keyboard and the icons. The flow of word completion was from left to right. Appendix N indicates how the screen formats appeared to the subjects.

ii. The word prediction software

Penfriend W3 1.04 is a word prediction program that is compatible with most word processors, but specifically also with Clicker 4. The window for the list of predicted words can be set to maximise accessibility and readability, as the prediction list can appear in customised Clicker cells. Penfriend is user-friendly for children and uses current advances in word prediction technology.

A literature survey of many research studies indicated the range of permutations that a word prediction program can be set to. What was clear was that *any* activity using word prediction is bound by the word prediction settings selected and has a significant impact on the functional use of the prediction software. Considering the goal of this research – to isolate the impact of spelling ability on word prediction usage for young Grade 4-6 children with spelling difficulties, a set of word prediction parameters was chosen. Great care was taken to choose settings that reflected 'middle-of-the-road' usage of word prediction. A table of all the parameters that needed to be considered, and which parameter settings were chosen for the research project can be found in Table 2.1. The higher level features of word prediction available in Penfriend (such as grammar prediction and sound output), were not relevant to this project and were not accessed, as they would have complicated the results.

3.5.4.2 The Graded Word Spelling Test

The Graded Word Spelling Test was used for identifying children with spelling difficulties, for arranging the sample population into four sample groups and for correlation analysis on the results of the research tests. The Graded Word Spelling Test tests the spelling of single words. There are 80 words in the test (see Appendix O).

Three important limitations in the use of the Graded Word Spelling Test for this particular research will now be described.

- The Graded Word Spelling Test tests the spelling of individual words, graded in average order of difficulty. Spelling of single words is recognised as only one facet of spelling words as there are a number of ways of assessing spelling. There are standardised proof reading-type tasks, writing out of single words in the traditional spelling test, a dictation-type test. Although each test can provide relevant information with respect to spelling performance, and although spelling tests correlate quite highly, they do measure different skills.

Considering all the factors related to the various spelling test forms, the single word spelling test was chosen for the research project – both for the research task and for the correlation of the performance of the children on the research task to their current spelling knowledge. This was because it provided a greater possibility of being highly selective and productive in the choice of words, and because basic spelling competence (internalising of spelling rules and their exceptions) is the main skill under scrutiny in this project, not general writing skills. It is a recognised limitation of this project that spelling competence on the research task can not be directly extrapolated to spelling competence in more functional writing, which requires the integration of a number of higher level skills.

- A second difficulty arising out of the use of the Graded Word Spelling Test for this particular research is that it is a British test, referenced to the national norms of progression and attainment of spelling skills in Britain. No modern single entry type spelling test standardised for South Africa that would serve the same purpose as The Graded Word Spelling Test could be identified. The Graded Word Spelling Test has become the most frequently purchased spelling test of its kind in South Africa.

As the test progresses, the words become more complex, and so become increasingly unfamiliar to the children. However, this is a feature of all spelling tests. There is a link between vocabulary and spelling ability, but that relationship is beyond the scope of this project.

- Thirdly, the Graded Word Spelling Test provides standardised scores in terms of Spelling Ages and Spelling Quotients. Children in Britain begin school at the age of 5-06 yrs in Aug/Sept of the calendar year. However, in South Africa the children begin school in the calendar year they turn 7 (see section 3.5.3.2). In June of the school year, the average age of the South African children should be exactly 10 years in Grade 4, 11 years in Grade 5 and 12 years in Grade 6.

  This age discrepancy in school starting ages results in considerable difficulty in applying the age norms of The Graded Word Spelling Test to South African children. The average 1½ year age difference in school starting age between South African and British children has a significant impact on the scores, especially in the earlier years. A direct reading of scores in the Spelling Age Norm table would result in South African children appearing to score at spelling age levels lower than they are in fact scoring, considering their shorter period of schooling. However, it is recognised that the effect of age difference between South African school children and British school children would probably decrease as the number of years of schooling increase.

In addition to the difficulties in applying spelling test age norms to South African children in general, there are additional difficulties in applying the age norms to the special needs education classes of South Africa, with their wide age range of children. An older child functioning at the same grade level as a younger child may have had as many years of schooling as the younger child, and should not be considered developmentally delayed just because he has not had the educational experience of his age peers.

For the method selected to determine which subjects had spelling difficulties, see section 3.5.5.2.

3.5.4.3 Mouse Control Screening

Mouse control is a fundamental requirement in the execution of the research task. Mouse control can be impacted by motor difficulties, visual difficulties and unfamiliarity of use.

The Mouse Control Screening served to screen the functional mouse ability of the subjects, and not to define the specific limitations of the subject. The design of the Mouse Control Screening aimed to cover for all the above requirements of mouse control.

The Mouse Control Screening took the form of timing the subjects as they clicked on a small object (a cat) with the mouse for 30 mouse clicks. As the cat was clicked on, it appeared in another position on the screen, and the subject had to keep on clicking on the cat until the end of the task, indicated by a red light. The task was pleasurable for the children and took on average less than a minute to execute. The exact sequence of steps of the Mouse Control Screening, and all the positions on the screen that the cat appeared in can be seen in Appendix M. The Mouse Control Screening was designed by the researcher, so is not a standardised test.

The procedure for administering the results is indicated in section 3.5.5.3. The Mouse Control Screening was designed using Clicker software. Each new screen was a single cell, inter-linked, grid. It operated in a continuous, automatic manner. Once the subjects started, they could continue to the end without any input from the administrator. The size of the cat was chosen to be smaller than any of the letter sizes in the keyboard that the subjects had to operate in the research task itself. The positions of the cat were chosen to cover the full extent of the screen and to have large variations in position relative to the previous position.

3.5.4.4 The wordlists

For the rationale and development of the wordlists for the research task, see section 3.4.1.

i. The final wordlists

In preparing the final list for the research task, it was necessary to first remove the 10 most unmatched pairs, spread out the 3 unpredicted word pairs through the list, separate out the 3 most difficult words which appeared in a cluster in the pilot study (furious/curious, chief/thief and witches/matches), switch some words of a pair from one list to another to balance the keystrokes, and correct minimum keystroke figures because of changed positions in the list. The final wordlist is seen in Table 3.7.

Table 3.7 : The final wordlists

Number of keystrokes required to spell words using typing only	Minimum number of keystrokes required to spell word using word prediction (first appearance of the word in the prediction list plus the keystroke required to select the word)	Word List A	Order of presentation	Word List B	Minimum number of keystrokes required to spell word using word prediction (first appearance of the word in the prediction list plus the keystroke required to select the word)	Number of keystrokes required to spell words using typing only
4	3+1	tall	1	ball	3+1	4
4	2+1	fire	2	wire	2+1	4
3	2+1	cry	3	dry	2+1	3
5	5	fairy	4	hairy	5	5
8	3+1	horrible	5	terrible	3+1	8
4	2+1	talk	6	walk	2+1	4
8	3+1	mountain	7	fountain	3+1	8
6	3+1	letter	8	ladder	3+1	6
6	3+1	yellow	9	follow	3+1	6
7	3+1	married	10	worried	3+1	7
6	2+1	garden	11	carpet	2+1	6
5	3+1	white	12	wheel	3+1	5
7	3+1	picture	13	mixture	3+1	7
6	6	rocket	14	packet	6	6
5	3+1	chief	15	thief	3+1	5
5	3+1	fruit	16	blood	3+1	5
4	2+1	tail	17	jail	2+1	4
5	1+1	light	18	right	1+1	5
6	3+1	bridge	19	fridge	3+1	6
4	2+1	hurt	20	turn	2+1	4
4	2+1	push	21	pull	2+1	4
5	2+1	alone	22	awake	2+1	5
5	2+1	noise	23	voice	2+1	5
5	3+1	crown	24	clown	3+1	5
7	3+1	furious	25	curious	3+1	7
7	7	witches	26	matches	7	7
8	3+1	polished	27	finished	3+1	8
8	2+1	sleeping	28	sweeping	2+1	8
7	3+1	weather	14	feather	3+1	7
5	4+1	string	30	strong	4+1	5
169	115				115	169

(Shaded) Word does not appear on prediction list

Note : Keystroke counts do not include the ‘enter’ keystroke required after *all* words are typed, in both the Typing Only and the Word Prediction subtests.

ii. Minimum keystrokes required

Adjustments to the recency and frequency statistics of the word prediction software program to ensure equal keystroke minimums for each word pair in the two wordlists had to be made for a few words.

It is important to note that minimum keystrokes of the words is not a constant. The recency and frequency statistics constantly change during use. The minimum keystroke figures seen in Table 3.7 were based on perfectly correct typing of each word, or choice of word from the prediction list, in the order presented. Wrong words selected from the prediction lists could result in changes in minimum keystrokes required of subsequent words.

The automatic space facility, which would decrease by 27 keystrokes the minimum number of keystrokes used to type the above 30 words with word prediction software in functional writing, is not utilised in the research task.

iii. First appearance of words in the prediction list

It was also important to have the predicted words appearing at an average frequency of chance. It would give a positive but distorted view of word prediction effectiveness if the target words mostly appeared for prediction after the first letter was typed, and a negatively distorted view if the target words mostly appeared for prediction only after the fourth letter was typed. The majority of the 2000 words in Penfriend's Default Lexicon appear after the second or third letter is typed. The words chosen for Wordlist A and Wordlist B reflect this frequency, as can be seen in Table 3.8.

Table 3.8 : Predicted words' first appearance

Never predicted		Prediction after 1st letter typed		Prediction after 2nd letter typed		Prediction after 3rd letter typed		Prediction after 4th letter typed	
Wordlist A	Wordlist B	Wordlist A	Wordlist B	Wordlist A	Wordlist B	Wordlist A	Wordlist B	Wordlist A	Wordlist B
rocket	packet	light	right	push	pull	letter	ladder	string	strong
witches	matches			garden	carpet	tall	ball		
fairy	hairy			alone	awake	married	worried		
				noise	voice	yellow	follow		
				talk	walk	picture	mixture		
				tail	jail	mountain	fountain		
				hurt	turn	crown	clown		
				sleeping	sweeping	horrible	terrible		
				cry	dry	fruit	blood		
				fire	wire	furious	curious		
						bridge	fridge		
						polished	finished		
						chief	thief		
						weather	feather		
						white	wheel		

iv. Familiarity of the words

The words on the wordlists have high familiarity to the children. All the words, except for 3 word pairs (see Table 3.9), appear in Penfriend's Default vocabulary list, which represents the 2000 most commonly used words by children (Spooner, 1999) (see Appendix D).

Seven words (crown, clown, feather, fridge, fountain, jail and sweeping) were added to the vocabulary list, because of the difficulty of finding words from the available list that met all the above criteria. None of the added words appeared to result in unfamiliarity issues in the pilot study.

v. Unpredicted words

Three word pairs were added to introduce words into the test task that were not reflected in the vocabulary list and would never be predicted. This reflected the natural use of word prediction, where it would be a common experience to find a word not predicted. Also, if the subjects understood that the target words *could* always be found, it may alter the way they work through the test task. However, if they were given to understand that sometimes the word would not appear, they would not be as perturbed when their target word did not appear (due to incorrect spelling or due to the word not being in the prediction list).

These three word pairs which were added, although they did not appear in the prediction list, had their root words appearing in the list. Table 3.9 list the 3 word pairs that are not predicted, and the root words that do appear in the prediction lists.

Table 3.9 : Unpredicted words

Words in wordlist that do not appear in prediction list	Root word appearing in prediction list
fairy	fair
hairy	hair
rocket	rock
packet	pack
witches	witch
matches	match

vi. The contextual phrase for each word

The words were placed in a contextual phrase. This was to ensure that the subject was fully aware which word had to be typed, and because there were a few homonyms in the wordlists. (The Graded Word Spelling Test uses a similar approach). After the required word was called out, the tester would say a sentence with the required word in it. Table 3.10 lists the phrases associated with each word. The phrases were chosen to be as short as possible, so as not to distract the subject, who may already have started typing the target word when the phrase is called out. An example of how the target word would be called out is: *"Picture. Draw a picture. Picture."*

Table 3.10 : The contextual phrase for each word

Wordlist A			Wordlist B		
No	Word	Phrase	No	Word	Phrase
1	tall	A tall man.	1	ball	Throw the ball.
2	fire	A veld fire.	2	wire	Electric wire.
3	cry	Babies cry a lot.	3	dry	Dry your hands.
4	fairy	A fairy godmother.	4	hairy	The hairy dog.
5	horrible	It tastes horrible.	5	terrible	A terrible thing happened.
6	talk	Talk softly.	6	walk	Don't walk on the grass.
7	mountain	Climb a mountain.	7	fountain	A water fountain.
8	letter	Write a letter.	8	ladder	Climb a ladder.
9	yellow	The colour yellow.	9	follow	Follow the leader.
10	married	A married woman.	10	worried	You look worried.
11	garden	A flower garden.	11	carpet	Sit on the carpet.
12	white	The colour white.	12	wheel	Turn the wheel.
13	picture	Draw a picture.	13	mixture	Cough mixture.
14	rocket	A space rocket.	14	packet	A packet of chips.
15	chief	An Indian chief.	15	thief	The police caught the thief.
16	fruit	Fruit juice.	16	blood	Blood is red.
17	tail	The dog wags his tail.	17	jail	The thief is in jail.
18	light	Switch the light on.	18	right	Right or wrong.
19	bridge	Cross over on the bridge.	19	fridge	A fridge keeps things cold.
20	hurt	I hurt my arm.	20	turn	Turn the tap on.
21	push	Push the trolley.	21	pull	Pull the suitcase.
22	alone	He is all alone.	22	awake	Wide awake.
23	noise	Don't make a noise.	23	voice	A lovely singing voice.
24	crown	A queen wears a crown.	24	clown	A circus clown.
25	furious	Mom was furious with me.	25	curious	The curious cat.
26	witches	Witches and wizards.	26	matches	Don't play with matches.
27	polished	Polished shoes.	27	finished	This is nearly finished.
28	sleeping	Sleeping Beauty.	28	sweeping	Sweeping the floor.
29	weather	Rainy weather.	29	feather	A bird's feather.
30	string	A ball of string.	30	strong	Strong muscles.

3.5.4.5 Video Camera

A video camera on a tripod stand was used to record the entire research task event.

3.5.4.6 Administration equipment

The researcher also required

1. Table (large enough for laptop, mouse, and researcher's notes), at suitable height for the subjects.
2. Two chairs, one for the subject and one for the researcher, well suited to the height of the table and to the subjects' size.
3. Stop watch for timing during the test.
4. "Procedure and record sheets", one for each child.

A "Procedure and record sheet" (see Appendix P) was available to guide the researcher in the procedural order to follow for each subject, as each subject had been assigned to 1 of 4 sample groups, which indicated which subtest and which word list the subject had to do first. The words for the training procedures and the words of the wordlists appeared in the "Procedure and record sheet". This helped the researcher in keeping the administration of the test consistent, thereby increasing reliability of data collection.

3.5.5 Data Collection procedures

3.5.5.1 Introduction

Consent for the project was obtained in writing from the Gauteng Department of Education, headmaster of main study school and the parents of the subjects (see Appendices Q, R and S). Letters were sent informing them of the purpose of the study, what the research task involved and what would be required of the children.

The teachers of the relevant classes were informed verbally of the nature of the project and their support secured, as the experimental phase was conducted during school hours. The permission of the Nursing Sister at the main school was also verbally requested, as the experimental phase was executed in the medical wing of the school, utilising their facilities. Discussions with the Psychologist, Speech and Hearing Therapists and Nursing Sister concerning the children's attention and hearing difficulties also took place.

The subjects' anonymity in the results was assured – to the parents and to the subjects themselves. No child was forced or pressurised to participate. Test anxiety was reduced through the assurance of anonymity, familiarity with the researcher as a staff member at the school, enjoyable, unthreatening training, reward of juice and chocolate, and the expression of gratitude for their help. All the relevant interest groups were informed that a copy of the final report would be available to any of them on request.

3.5.5.2 Graded Word Spelling Test

The Graded Word Spelling Test was administered to *all* the Grade 4 - Grade 6 pupils at the main study school in June 2003. Routine testing for spelling was done by the teachers every year and the results formed part of the ongoing records of the children's progress. Permission was requested and granted from the headmaster to use the results for the current project. However, the test usually administered proved to be unsuitable – there were too few test items to obtain a large variance across the pupils results, especially among the higher grade children. Permission was then requested and granted to administer another test, the Graded Word Spelling Test.

The test was administered by the researcher as a group test in each of the seven classes used in the study. The Graded Spelling Test is not usually administered in its entirety, because each grade would be given only the range of the words suitable for that grade. However, because there was reasonable expectation that some of the children might score significantly below their age level, the test was always started from Word 1. In all classes, all the children completed up to word 40. The researcher then paused to look at the children's work, and a few children (the few who were still scoring words correctly) were requested to continue for another 10 words. The others were given the option of continuing or not. This procedure continued until each child had made at least 10 errors. Allowing the children to stop if they were no longer coping with the spelling words was considered necessary, as the weaker spellers tended to become frustrated when word after word was clearly beyond their abilities. Some, however, chose to continue and they were allowed to.

In the administration of the Graded Word Spelling Test, all the words are used in a sentence to put them in context and for easy identification (Vernon, 1998). The words of The Graded Word Spelling Test are familiar to South African children, except for word 29 – *Shovel*, which is partially equated with the South African *spade.* A few minor changes were made for the South African situation. For shovel, an additional sentence of "A shovel is like a spade" was included. Also, a minor change was made in the sentences paired with words 9 and 32. Instead of '*Down. Go down to the cellar. The word is down.'*, "Go down the *steps* was used" as a cellar is an unfamiliar concept to many South African children. And instead of '*Million. The building cost a million pounds. The word is million.'*, the word *pounds* was replaced by *rand*.

The test sheets were then scored by the researcher according to the criteria of the manual. Scoring discontinued after 10 errors, even if, as in a few cases, children spelt a word correctly after the cut-off.

Using the scoring criteria on the Graded Word Spelling Test, a spelling age for each subject was determined from his raw score of correctly spelled words. The raw score at which a child 1 year younger than the subject would score at the 36.9 percentile rank (1 Standard Deviation from the mean) was compared to the subject's raw score. If the subject's raw score was higher, he would be disqualified from the sample, as he would have a spelling age equal (or above) that expected for his age.

The above method offered a full year compensation for the earlier school starting age of British children compared to South African children. The 36.9 percentile was chosen to compare the results against, as a child a little below average (50 percentile) could not be considered to have a spelling difficulty. The method above was considered a conservative basis for determining which children had spelling difficulties.

3.5.5.3 Mouse Control Screening

The Mouse Control Screening was administered to every child in the potential sample group in a quiet room in a therapy department. The same laptop computer used in the research task was used for the Mouse Control Screening. The mouse was positioned on the right hand side of the laptop for right-handed children and on the left for left handed children. For the instructions given to the children, see T. Timing started as the children clicked on the GO and stopped as they clicked STOP. Timing was measured in seconds. The Mouse Control Screening took on average less than 1 minute per child to administer.

The times each subject took to complete the task were plotted on a graph. See Figure 3.3.

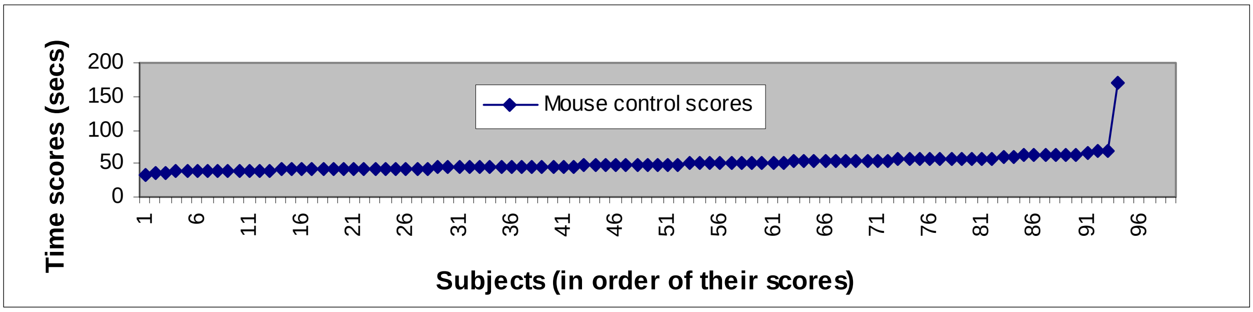

Figure 3.3 : Mouse control screening scores

3.5.5.4 The Research Task

i. Location of the research task

The research task was administered in a private therapy room. In front and to the right of the subject was the video camera on a tripod stand. On a single table was the laptop computer and a record form for the researcher. The record form had the detailed procedure requirements for the researcher, and space where the researcher could record all relevant information.

ii. Sequence and timing of events

After initial preparations by the researcher, the subject entered into the testing room. He was welcomed and orientated as to what the project was about, what would happen and what was expected of him. The subject was then taught how to use the on-screen keyboard.

At this point, the four sample groups of the project followed a different sequence of events through the rest of the research task. Table 3.11 is an outline of the sequence of events in the test administration. The estimated times involved are indicated.

Table 3.11 : Procedure overview

<table>
<tr><th>Group 1</th><th>Group 2</th><th>Group 3</th><th>Group 4</th></tr>
<tr><td colspan="4">Initial preparations (about 1 min)</td></tr>
<tr><td colspan="4">Welcoming and orientation (about 2 mins)</td></tr>
<tr><td colspan="4">On-screen keyboard training (about 2 mins)</td></tr>
<tr><td rowspan="2">TYPING ONLY
(TO) SUBTEST
WORDLIST A
(about 15 mins)</td><td rowspan="2">TYPING ONLY
(TO) SUBTEST
WORDLIST B
(about 15 mins)</td><td colspan="2">Word prediction training (about 10 mins)</td></tr>
<tr><td rowspan="2">WORD PREDICTION
(WP) SUBTEST
WORDLIST A
(about 15 mins)</td><td rowspan="2">WORD PREDICTION
(WP) SUBTEST
WORDLIST B
(about 15 mins)</td></tr>
<tr><td colspan="2">Rest (about 10 mins)</td></tr>
<tr><td colspan="2">Word prediction training (about 10 mins)</td><td colspan="2">Rest (about 10 mins)</td></tr>
<tr><td>WORD PREDICTION
(WP) SUBTEST
WORDLIST B
(about 15 mins)</td><td>WORD PREDICTION
(WP) SUBTEST
WORDLIST A
(about 15 mins)</td><td>TYPING ONLY
(TO) SUBTEST
WORDLIST B
(about 15 mins)</td><td>TYPING ONLY
(TO) SUBTEST
WORDLIST A
(about 15 mins)</td></tr>
<tr><td colspan="4">Closing (about 1 min)</td></tr>
</table>

The total time required to test each subject was approximately 55 mins, including a 10 min break.

iii. Initial preparations

Before the next subject entered the room, the researcher had to

- get the subject's individual record form ready

- activate the word prediction program with a ‘clean’ lexicon, so that the words entered by the previous subject did not affect the recency and frequency ratings of the words of the wordlists and thereby change the number of keystrokes that were required to activate the target words.

iv. Welcoming and orientation

The subject was welcomed and made to sit comfortably in front of the computer. For the detailed words of introduction, see Appendix T.

v. On-screen keyboard training

The words for the on-screen keyboard training were chosen to be simple enough not to intimidate even the weakest spellers and to cover a wide range of letters on the keyboard. The calling out of the words followed the same procedure as that in the Graded Word Spelling Test. The word was said, followed by a short phrase using that word, and then the word was repeated. For the procedure that was followed in the training, and the instructions that were given, see Appendix T.

If the subject spelled the word incorrectly, the researcher gently corrected the error. If the child requested feedback, it was given, to ensure that comprehension of the task was maximised. However, they were told that in the real test, no help would be given. If the children forgot, reminders were given to 'enter' as soon as the word was entered. It was important that the child saw the ‘enter’ as part of the task of completing the word, so as not to lose time.

vi. Word prediction training

Orientation to the word prediction subtest was more extensive. This was because word prediction was a new skill that had to be taught before it could be used. It was also an interactive learning experience as choices were made throughout the typing of a word. The subject was made aware that the same result of a correctly typed word could be achieved through a few different options.

The pilot study used a 20-word training program. The pilot study indicated that by 15 words of training, most children understood the principle of word prediction and could find easy words in the prediction list. For the main study, a 15-word training was given, but the last five words now served as a test to determine if the children could use word prediction independently. If they found the predicted word in the prediction list for 4 out of the 5 words, they were considered able to use word prediction. If they were unable to find at least 4 of the 5 words, their results would not be used.

The words for the 5-word word prediction test were chosen for maximum spelling ease and to minimise the risk that children would fail because of spelling errors rather than because of not understanding word prediction as a typing strategy. Only regular letter-name sounds, and only three or four letter words, were used. Some letters were avoided because of the spelling difficulties often associated with them, namely

- ‘b’ and ‘d’, because of the reversal difficulties that some of the subjects might have.
- the short vowel ‘i’, because of the vowel confusion between ‘i’ and ‘e’ often seen in children whose mother tongue is not English.
- ‘c’, because of the confusion of which ‘c’ sound to use, ‘c’ or ‘k’.

All children followed the same training program. For the procedure that was used in the word prediction training see Appendix T.

The following principles guided the researcher’s interactions with the subjects.

- The children were made aware that the two subtests would be timed so they had to work quickly. They were also urged to work carefully. However, if the children appeared to be fixated with a word, searching for it in the list but not progressing, they were urged to keep moving. *“If you can’t find the word, don’t waste too much time looking for it, just keep on typing.”*

- If the children did not know how to spell a word, they were helped to spell it correctly. This could be done by indicating which letter had to be typed next if the children hesitated for a long time, or by indicating the error immediately after a wrong letter was typed. The important issue was that the children learnt how to operate the program and experienced the learning principle that each word in the training gave. The target word would not appear in the prediction list if it was incorrectly spelt.

- If the children missed an opportunity to use a word, it was pointed out, even if it meant deleting their previously typed letter so that they could see the missed opportunity (as letters were deleted, the words that had been in the prediction list would reappear). However, they were told it was OK to not use the word. They could choose whether or not to use the word.

- If the children appeared to be unable to read the words in the prediction list, they were helped to find the target word by giving clues, such as telling him, *"It is there. Look carefully and try to find it."*

- At all times it was the researcher's task to encourage the children and to reduce anxiety. If they succeeded in selecting a word from the prediction list, encouraging comments such as, *"Well done"*, *"You found it!"*, *"Good"*, were given.

vii. The two subtests

At the beginning of the first subtest, the subject had a blank screen above his on-screen keyboard. The subject typed the words of Wordlist A and Wordlist B into the same document, which was opened, saved and closed for each subtest.

The researcher called out the first word of the first subtest. As the researcher called out the first word, timing on the stopwatch began. As the subject entered the word, a digitally spoken 'next' was heard. The researcher then immediately called out the next word. The words of the two lists were always presented to the subjects in the same order. As the subject clicked 'enter' for the last word on the Wordlist, the researcher stopped timing. The time was recorded on the "Procedure and record sheet", rounded to the nearest second.

The subject was permitted to alter the word as much as he wished in the formation of the word, but as he clicked ‘enter’, the word disappeared off the screen. If the subject needed clarification or repetition of the word, the tester repeated the word. The words were read out clearly, with ordinary pronunciation and speed, and without accenting or spacing out the syllables so as to provide clues on the spelling of the word (Vernon, 1998).

The instructions given to the children before the subtests were administered can be seen in Appendix T.

viii. The rest

After the first subtest, the subjects were given juice and a chocolate. They either went back to class (enjoying the refreshments later), or waited nearby (enjoying the refreshments then) until the researcher called them back for the second half of the test procedure.

ix. The closing

On completion of the test, the child was thanked for his participation and asked to return to class. The subject’s document then had to be saved and copied to a disk.

3.5.5.5 The video recording

The entire research event, except for the rest period, was filmed, for all the subjects. The video camera was positioned to the front of the subjects, so that it could capture the face and hands of the subjects and the researcher at all times. It was activated after the subject had entered the room but before the welcoming began. The subject was informed of its presence and function during the welcoming and orientation phase of the task. It was switched off only after the closing thank-you had been given.

3.5.5.6 The output document

The document of spelled words created on the screen was prepared for printing (by decreasing the font to 12 pt and inserting headings). They were then printed out and marked for spelling accuracy. A score of the number of words spelled correctly was obtained for each of the subtests and entered onto the subject's "Procedure and record sheet" (see Appendix P).

Also calculated in the marking of the document were the number of approximations used (see section 4.2.4 for a discussion on approximations). Approximations were not initially planned for in the research task but emerged on reviewing the results.

3.5.5.7 Record of all the raw data

The following raw data on all the subjects, accumulated by the end of the project, was entered on an Excel spreadsheet master information table for processing (see Appendix U).

1. Respondent number
2. School grade
3. Gender
4. Sample group allocation
5. Chronological age
6. Raw score of the Graded Word Spelling Test
7. Spelling age
8. Number of words spelled correctly on the TO subtest
9. Time taken to complete the TO subtest, in minutes and seconds
10. Number of words spelled correctly on the WP subtest
11. Time taken to complete the WP subtest, in minutes and seconds.
12. Number of approximations

Added to the master information table were the calculations

1. Difference between Chronological age and Spelling age
2. Difference between WP and TO accuracy scores
3. Difference between WP and TO time scores

3.5.5.8 Tester reliability assessment

The video material was passed on to an independent rator, who was commissioned to watch the video. She then analysed and assessed the researcher's administration of the research task for a random sample of 25% of the subjects. The independent assessor (a registered Speech and Hearing Therapist) recorded her findings on a copy of the same "Procedure and record sheet" used by the researcher (see Appendix P).

The assessor's task was to

- check that all units of the procedure had been executed, and in the correct order.
- time the research test to ascertain accuracy of time scores
- note any irregularities in the research protocol
- make any other comments deemed relevant

The assessors' findings were collated into a table (see Appendix V). Also in the table are the researcher's comments on the independent assessor's findings.

In conclusion, it was seen that

- there was a high degree of consistency between the researcher's and the assessor's timings. Sometimes there were a few seconds difference between the two times, but this was considered to not be significant over the total time, the mean of which across the sample was 118.36 secs for the TO subtest and 147.63 for the WP subtest.
- there were a few procedural irregularities such as premature calling of words and interruptions. These were compensated for by the researcher, as indicated in the table.
- there was a procedural error of wrong group placement. This error is unfortunate, but it is considered to not have had a marked effect on the results.

3.5.6 Data analysis

Table 3.12 outlines the data analysis procedures that will be applied to the data, in order to interpret the results.

Table 3.12 : Data analysis procedures

	Aim	Test	Reference
1	To calculate the mean and range of • the TO and WP accuracy and time scores, as well as the differences between them, to assess the mean and range of test responses across the sample. • the GWST and TO accuracy scores, to assess the degree of equivalence of the two spelling tests.	Means, ranges and standard deviations of scores to compare one group of results with another.	McMillan & Schumacher, 2001.
2	To calculate the mean number of approximations used by age and grade groups of the sample, to assess which groups tended to use approximations more.		
3	To calculate • the mean age, GWST and spelling age of each of the four sample groups, to assess the degree of equivalence of the four sample groups with respect to sample distribution. • the mean TO and WP accuracy scores and times of each of the four sample groups, to assess the degree of equivalence of the four sample groups with respect to test results		
4	To correlate • TO and WP accuracy scores, to ascertain if spelling scores improved with the use of word prediction. • TO and WP time scores, to ascertain if on-screen typing times decreased with the use of word prediction. • scores on the Graded Word Spelling Test with those of the TO subtest, to ascertain if there was a high correlation between the two tests. • scores on the Graded Word Spelling Test with those of the difference between the WP and TO subtests, to ascertain if there was a correlation between the children's spelling knowledge and the degree to which they benefited from the use of word prediction. • the number of approximations used with TO and WP and GWST accuracy scores, to ascertain if the tendency to use approximations decreased with increasing spelling knowledge.	Pearson's Correlation Coefficients, to indicate the size and degree of the relationship between the two variables.	
5	To calculate the variance between the dependent variables (TO and WP accuracy and time scores) and • the subject characteristics (grade, gender and chronological age), to ascertain if there was a significant degree of difference between the results with respect to these subject characteristics. • the research design characteristics (wordlists used and presentation order of the subtests), to ascertain if there was a significant degree of difference between the results from each of the four sample groups.	LS (Least Squares) Means were used for analysis of variance. (The ANOVA resulted in some discrepancies in the results). LS Means is one of the pairwise comparison methods that are used in such cases, the advantage being that it corrects for unbalanced designs when there are not equal numbers in the cells (the sample groups varied from 19 – 22). Duncan's Multiple Range Test was applied to test which of the means were significantly different from one another (used when more than two groups are being compared to each other).	

3.6 Summary

A research plan was drawn up to investigate the benefit of using word prediction with children who have spelling difficulties, and the relationship between children's performance with word prediction and their spelling ability. The sample of Grade 4 – 6 children with spelling difficulties was drawn from a special needs school. A standardised spelling test (Graded Word Spelling Test) was administered to the children to obtain a baseline of their current spelling ability and for correlation with their performance on the research task. A computerised single-word input task was designed to test the difference between the children's performance when typing with and without the aid of word prediction. A pilot study tested many aspects of the study, the results and recommendations of which were incorporated into the final design of the research task.

Chapter 4

Results and discussion

4.1 Introduction

The data is first analysed with respect to the three sub-questions. The first two sub-questions investigate spelling accuracy improvements and typing speed enhancements with the use of word prediction software. The third sub-question investigates the relationship between a child's spelling ability and the influence of word prediction on spelling accuracy and typing speed. Arising from the study, the unexpected influence of 'approximations' will also be discussed.

The findings however, can not be fully interpreted without relating the results to the specific characteristics of this study. The most important variables that will be addressed are the subject characteristics (gender, grade and chronological age) and the features of the research design itself (the four different groupings of the participants, the two wordlists used and the order of presentation of those wordlists).

4.2 Addressing the research sub-questions

Table 4.1 is a summary of the correlation data between the variables that are being considered in this section, and will be referred to throughout the discussion on the results. Only correlation coefficients that are statistically significant will be discussed in this chapter.

Table 4.1 : Correlations between dependent variables

	Graded Word Spelling Test scores (GWST scores)	Difference between chronological age and spelling age (Diffage)	Typing Only subtest accuracy score (TO score)	Typing Only subtest time (TO time)	Word Prediction subtest accuracy score (WP score)	Word Prediction subtest time (WP time)	Difference between WP and TO accuracy scores (Diffscore)	Difference between WP and TO times (Difftime)
Graded Word Spelling Test scores (GWST scores)		-0.70	0.91	-0.27	0.88	-0.41	0.02	0.38
		<0.0001	<0.0001	0.0438	<0.0001	0.0001	0.8505	0.0004
Difference between chronological age and spelling age (Diffage)	-0.70		-0.56	0.05	-0.63	0.20	0.13	-0.26
	<0.0001		<0.0001	0.6667	<0.0001	0.0784	0.2456	0.0193
Typing Only subtest accuracy score (TO score)	0.91	-0.56		-0.18	0.83	-0.37	0.28	0.37
	<0.0001	<0.0001		0.1045	0.0001	0.0007	0.0135	0.0008
Typing Only subtest time (TO time)	-0.27	0.05	-0.18		-0.17	0.79	-0.02	0.31
	0.0438	0.6667	0.1045		0.1327	<0.0001	0.8704	0.0057
Word Prediction subtest accuracy score (WP score)	0.88	-0.63	0.83	-0.17		-0.32	-0.32	0.31
	<0.0001	<0.0001	0.0001	0.1327		0.0035	0.0043	0.0057
Word Prediction subtest time (WP time)	-0.41	0.20	-0.37	0.79	-0.32		-0.07	-0.60
	0.0001	0.0784	0.0007	<0.0001	0.0035		0.5276	<0.0001
Difference between WP and TO accuracy scores (Diffscore)	0.02	0.13	0.28	-0.02	-0.32	-0.07		0.01
	0.8505	0.2456	0.0135	0.8704	0.0043	0.5276		0.9068
Difference between WP and TO times (Difftime)	0.38	-0.26	0.37	0.31	0.31	-0.60	0,01	
	0.0004	0.0193	0.0008	0.0057	0.0057	<0.0001	0.9068	

Note :
1. Pearson's Correlation Coefficients were used for correlation data.
2. The upper number is the correlation figure and the lower number is the p-value.
3. Shaded areas indicate the correlation coefficients that are statistically significant at a 1% confidence level (p-value ≤0.0001).

4.2.1 Research sub-question 1

Does typing with the use of word prediction improve spelling scores compared to typing without the use of word prediction?

The data of all the subjects was sorted in ascending order of TO scores. A graph was plotted to determine how WP scores related to TO scores (see Figure 4.1). From the graph it is evident that for the majority of the subjects (88.75%), WP scores were higher than TO scores. This indicates a general improvement across the sample of accuracy scores with the use of word prediction in the research task. For a few subjects, however, word prediction can be seen to have had a negative influence on accuracy scores. TO and WP scores have a correlation coefficient of 0.83 (see Table 4.1). As spelling improved, so the performance of the subjects on WP tended to improve.

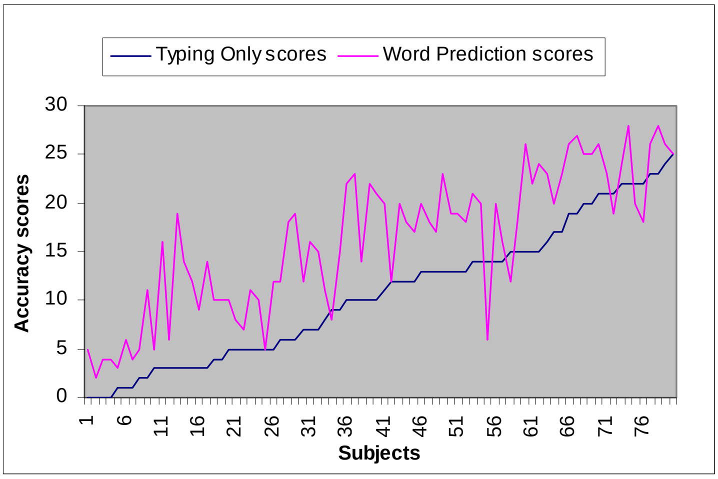

Figure 4.1 : The relationship between TO and WP accuracy scores

Table 4.2 documents the mean TO and WP scores across the sample group, as well as the mean difference between the WP and TO scores. Whereas the average spelling accuracy score *without* word prediction was 10.73, the average spelling accuracy score *with* word prediction was 16.10. Word prediction increased the average accuracy scores by 5.38 correctly spelled words, which represents an average 50% improvement in spelling accuracy.

Table 4.2 : Summary of TO and WP scores in terms of spelling accuracy

Number of subjects N = 80	Typing Only (TO) accuracy scores		Word Prediction (WP) accuracy scores		Difference between WP and TO scores (Diffscore)	
Mean	10.73		16.10		5.38	
Range	0	25	2	28	-8	+16
Standard deviations	7.05		7.14		4.19	

Table 4.2 includes the range of scores seen in the TO and WP subtests and the standard deviations of the scores. Across the sample, there was an increase in correct scores from –8 to +16 when using word prediction, indicating that there is great range in individual responses to word prediction usage.

The wide range of subjects' responses to word prediction must be a seen as cautionary when introducing word prediction as writing support where there are spelling difficulties, without greater awareness of the individual's response to word prediction. The literature has indicated that the specific characteristics or skills of the user impact on the apparent efficacy of word prediction usage (Koester & Levine, 1998; Szeto et al, 1993; Venkatagiri, 1994). The influence of some subject characteristics on word prediction effectiveness in this study are addressed in section 4.3. In addition to the influence of subject characteristics, a study of some of the influences of this research design may also explain some of the extremes in the range of scores (see Section 4.4).

4.2.2 Research sub-question 2

Does the use of word prediction increase text entry time?

The data of all the subjects was sorted in ascending order of TO times. A graph was plotted to determine how WP times related to TO times (see Figure 4.2). From the graph it is evident that for the majority of the subjects (91.25%), WP times were higher than TO times, indicating a general increase across the sample of typing time when typing with word prediction. TO and WP times have a correlation coefficient of 0.79 (see Table 4.1).

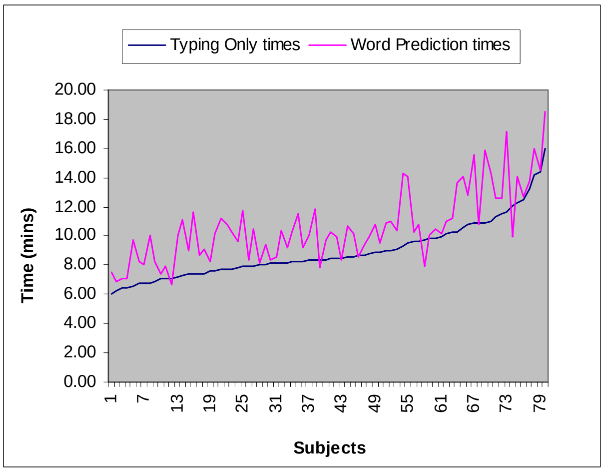

Figure 4.2 : The relationship between TO and WP times

Table 4.3 documents the average TO and WP times across the sample, as well as the average difference between the TO and WP times. Whereas the average time to complete the TO subtest was 8.98 mins, the average time to complete the WP subtest was 10.61 mins. The average increase in time with the use of word prediction was therefore 1.64 mins. The time difference between typing without and with word prediction relates to an average 18.2% increase in typing time across the sample.

Table 4.3 : Summary of TO and WP text entry times

Number of subjects n = 80	Typing Only (TO) Times		Word Prediction (WP) times		Difference between WP and TO times (Difftime)	
Mean	8.98 mins (538 secs)		10.61 mins (637 secs)		1.64 mins (99 secs)	
Range	6.03 mins 362 secs	15.98 mins 959 secs	6.70 mins 402 secs	18.53 mins 1112 secs	-2.07 mins	+5.53 mins
Standard deviations	118.36 secs		147.63 secs		89.75 secs	

Table 4.3 includes the range of times seen in the TO and WP subtests and the standard deviation of the times. Across the sample, there was an increase in times from -2.07 mins to +5.53 mins when using word prediction, indicating that there is great range in individual responses to word prediction usage with respect to typing times.

Again the caution is raised. Failure to recognise individual differences in capabilities and responses to word prediction when spelling difficulties exist can be counter-productive. Increases in accuracy scores must be measured against time costs.

4.2.3 Research sub-question 3

Is there a relationship between spelling knowledge and the influence of word prediction on spelling accuracy and typing speed?

The purpose for establishing if there is a relationship between spelling knowledge and improvements in spelling accuracy and typing speed with word prediction is a practical one. If such a relationship exists, word prediction intervention for young children with spelling difficulties would be better understood, and decisions of when and to whom word prediction should be offered would be based on a better foundation.

Before the relationship between spelling ability and the influence of word prediction on spelling accuracy and typing speed can be discussed, it is important to establish if there is a good match between GWST accuracy scores (the independent standardised test, administered before the research task) and TO accuracy scores (the spelling list designed for the research project). The TO scores form the base scores to establish if word prediction improves accuracy scores. The GWST scores are correlated against the research results so that Question 3 can be answered. It is thus important that there be a high correlation between the two tests.

The data of all the subjects was sorted in ascending order of GWST accuracy scores. A graph was plotted to determine how TO accuracy scores related to GWST accuracy scores (see Figure 4.3). The graph shows that the two tests had a similar distribution of scores across the sample.

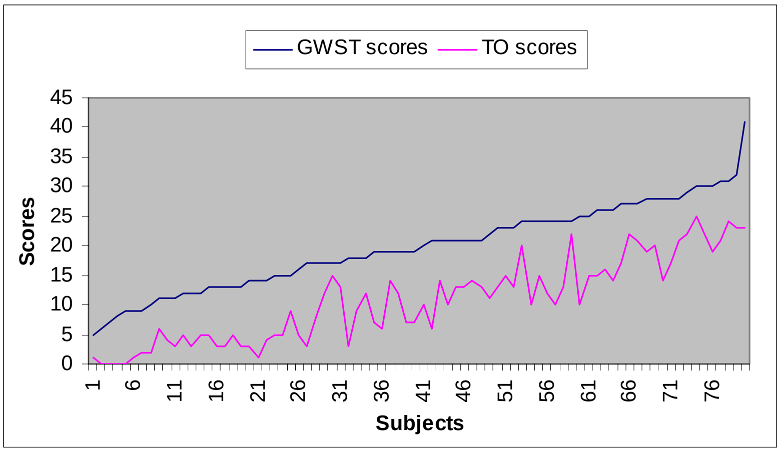

Figure 4.3 : The relationship between GWST and TO accuracy scores

Table 4.4 presents the averages, ranges and standard deviations of the two tests. The subjects scored an average of 19.78 correct words for GWST and 10.73 correct words for TO. The standard deviation for GWST was 7.22 and for TO it was 7.05. A weakness of the TO test was that the subjects who scored the highest accuracy scores on the TO subtest had very little room for improved accuracy scores in the WP subtest, as they scored close to the ceiling of the test (25 out of 30).

Table 4.4 : Comparison of GWST and TO accuracy scores

	Graded Word Spelling Test scores (GWST)	Typing Only scores (TO)
Mean	19.78	10.73
Lowest	5	0
Highest	41	25
Ceiling	80	30
Standard deviations	7.22	7.05

Table 4.1 indicates the correlation figures for the four primary variables under consideration in this research project - TO score, WP score, TO time and WP time. There was a high correlation between GWST accuracy scores and TO accuracy scores (0.91) and between

GWST accuracy scores and WP accuracy scores (0.88). This indicates that accuracy scores, for both TO and WP, tended to increase with increasing spelling knowledge.

There was however not a strong correlation between GWST scores and Diffscores. That is, there was little correlation between GWST scores and the *degree* of benefit in improved scores with word prediction use. This indicates that spelling knowledge can not be used as a strong indicator of *which* children will benefit *more* from word prediction than others.

The correlation between GWST scores and TO and WP typing times were not as high as for accuracy scores, being -0.27 for TO time and -0.41 for WP time. However, it can be seen that there was a negative relationship between GWST and TO and WP typing times. That is, as spelling knowledge increased, both TO and WP typing times tended to decrease.

The correlation between GWST scores and Difftime was not high (0.38), but appears to indicate that as spelling knowledge increased, the difference between WP and TO times tended to increase. As spelling knowledge increased, the children were more able to enjoy the time benefits offered by word prediction. It may be that as spelling knowledge increased, the time cost of word prediction use decreased.

It can therefore be seen that there was a high correlation between GWST scores and TO and WP scores, but not between GWST scores and Diffscores. It appears that spelling knowledge does not have a direct correlation with the degree to which a child with spelling difficulties will benefit from word prediction use. However, there was a moderate correlation between time benefits and spelling knowledge, indicating a possible decrease in time cost as spelling knowledge increases.

4.2.4 An additional cost to consider in word prediction usage – Approximations

4.2.4.1 Approximations described

An additional cost to consider in word prediction usage was highlighted during the test administration and marking – the tendency to make selection of approximate words.

The apparent increase in spelling scores with the use of word prediction must be seen against this tendency to make selection of approximate words. The selection of an approximate word was a specific error type seen in the subjects' responses during word prediction. Approximations were words that had been chosen by the subjects from the prediction list, and were therefore correctly spelled, but were not the target words the subject was required to spell. Approximate words included the selection of homonyms, where these appeared in the prediction lists. Only occasionally was the selection of an approximate word a random response. The subjects who chose an approximate word were generally under the impression that they had chosen the target word. The error was related to incorrect word recognition.

In functional writing and reading, the increase in the number of words correctly spelled when using WP must be measured against the confusion caused by the introduction of correctly spelled but approximate words into the text. The tendency to misread target words and insert approximate words was a frequent response in the research task and serves as a serious caution in the introduction of WP to improve spelling. In written text, an inappropriate word may be even more difficult to read than an incorrectly spelled word.

The average number of approximations used across the sample was 2.94. The range of approximations used was 0 - 22, and the standard deviation of scores was 3.47.

4.2.4.2 Which subjects tended to use approximations?

An analysis was done to see which subjects tended to use approximations the most. Table 4.5a highlights the relationship between approximations and the subject characteristics of chronological age and grade by indicating the means for each of four age groups and the three grades. Table 4.5b gives the correlation coefficients between approximations and other variables.

Table 4.5a : Approximations and chronological age and grade

		Mean number of approximations selected	Standard deviations
Chronological age	10 yrs	3.13	2.75
	11 yrs	3.29	2.77
	12 yrs	2.73	4.42
	13+ yrs	2.79	3.09
Grade	4	4.74	4.74
	5	2.75	2.76
	6	1.8	2.38

Table 4.5b : Correlating approximations and other variables

	Chrono-logical age	Graded Word Spelling Test scores (GWST)	Spelling age (Spage)	Typing Only score (TO score)	Word Prediction score (WP score)	Typing Only time (TO time)	Word Prediction time (WP time)
Approx-imations	-0.05	-0.57	-0.55	-0.50	-0.65	0.19	0.20
p-value	0.6704	<0.0001	<0.0001	<0.0001	<0.0001	0.0927	0.0728

Note : Pearson's Correlation Coefficients were used for correlation data.
The upper number is the correlation figure and the lower number is the p-value.
Shaded areas indicate the correlation coefficients that are statistically significant at a 1% confidence level (p-value ≤0.0001).

It can be seen that there was a decrease in the tendency to use approximations as grades increased. Grade 4 children used on average 4.74 approximations, the Grade 5s used 2.75 and the Grade 6s used 1.8. This is probably due to the progressive spelling maturity of each increasing grade.

The relationship between approximations and chronological age, however, was weak – the correlation was –0.05. Although increasing age is usually well related to increasing grades, for the participants of this study the relationship was not a direct one. There was a wide age range in each grade of the sample group (see section 3.5.3.2).

Age and grade appear to relate differently to word prediction efficiency. Increasing grades (not not age) resulted in a decreasing tendency to use approximations, whereas increasing age (not not grade) resulted in decreasing benefit in improving accuracy scores.

The tendency to use approximations did not correlate well to time scores but there was a significant correlation with accuracy scores. TO scores had a correlation coefficient of -0.50 with the number of approximations chosen, and WP scores a correlation coefficient of -0.65. This negative correlation suggests that the higher subjects scored on TO and WP accuracy scores, the less they tended to use approximations. What can be understood from this is that as children's spelling knowledge improved, they were probably less likely to select incorrect words from the prediction list. It could be that as their ability to form words increased, so did their ability to recognise words.

Approximations correlate to GWST scores and spelling age in a similar way as to TO (-0.57 for GWST scores and -0.55 for spelling age), supporting the indication that as spelling knowledge and spelling ability increases, so the tendency to use approximations decreases.

4.2.4.3 The words that most frequently elicited approximate words

Appendix W lists all the words in the two wordlists of the research task, and the approximate words that were chosen. In general word prediction use, the words that tend to elicit approximations would be directly related to the specific words appearing together with the target word in the prediction window and the vocabulary of the word prediction software. It is probable that the larger the word prediction vocabulary, the more likely approximations will be selected because of the increasing chance of visually similar words appearing.

4.3 Analysis of variance in the results between subject characteristics

It has been shown that, on average, a significant increase in spelling accuracy scores can be achieved with the use of word prediction, but that for most subjects there is an increased time cost associated with that increase. It would however be interesting to note whether there were significant differences in the results associated with the variations in the sample characteristics. An analysis of variations in the results relative to gender, grade and age follows, the data summarised in Table 4.6. Only sections highlighted in grey, which indicate statistical significance, will be discussed.

Table 4.6 Analysis of variance for grade, gender and chronological age

				Dependent variables													
				Typing Only Score (TO score)		Typing Only Time (TO time)		Word Prediction score (WP score)		Word Prediction Time (WP time)		Difference between TO and WP scores (Diffscore)		Difference between TO and WP times (Difftime)		Approx-imations (Approx)	
Independent variables	Grade	4	Mean	C	5.55	A	576	C	11.05	B	693	A	-5.50	A	-118	B	4.73
		5		B	11.68	A	528	B	16.93	AB	622	A	-5.25	A	-95	A	2.75
		6		A	13.63	A	520	A	19.03	A	609	A	-5.40	A	-89	A	1.80
	Gender	Boy	Mean	A	10.23	A	553	A	15.36	A	656	A	-5.13	A	-103	A	3.04
		Girl		B	11.88	A	503	B	17.83	A	592	A	-5.96	A	-90	A	2.71
	Chronological age	10	Mean	C	7.60	A	592	B	13.93	B	723	B	-6.33	A	-131	A	3.13
		11		C	7.95	A	526	B	14.33	B	638	B	-6.38	A	-112	A	3.19
		12		B	12.47	A	528	A	17.60	AB	603	A	-5.13	A	-76	A	2.73
		13+		A	14.50	A	519	A	17.86	A	613	A	-3.36	A	-95	A	2.79

Note : Duncan's Multiple Range Test was applied to test which of the means were significantly different from each other (at a 5% level of significance). Where the groups have a different symbol, they are significantly different. Significantly different groups have been shaded in grey.

4.3.1 Grade

Table 4.6 gives an indication of how the scores of the three grades of the sample group related to each other. The main study indicated a significant increase in both TO and WP accuracy scores as the grades increased from Grade 4 to Grade 6. Increased spelling knowledge is to be expected across the increasing grades. However, all three grades shared similar Diffscores between WP and TO accuracy scores i.e. all three grades benefited from word prediction usage to a similar degree.

There is a strong relationship between Grade level and the tendency to use approximations when choosing words from the prediction window. Grade 4 subjects chose on average 4.73 approximate words each, the Grade 5 subjects chose on average 2.75 and the Grade 6 subjects chose on average only 1.80 approximate words. This is probably due to maturation in sight word recognition (see section 4.2.4.2).

The differences between the results of the Grade 4s and the Grade 5s (for accuracy scores and approximations) are greater than the differences between the results of the Grade 5's and the Grade 6s. This may indicate that there is more spelling developmental progress between Grade 4 and 5 than there is between Grade 5 and Grade 6.

Similarly, the time scores between the three grades indicated that the younger grades may pay a heavier time cost than the children in the higher grades.

4.3.2 Gender

The girls scored significantly higher than the boys in both TO and WP accuracy scores. This supports the finding that girls spell better than boys (Allred, 1990).

There were no significant differences between the time scores or the tendency to use approximations relative to gender (see Table 4.6).

4.3.3 Chronological age

It can be seen in Table 4.6 that TO and WP scores increased as age increased. This trend is to be expected as increasing age is associated with higher school grades and therefore increasing spelling knowledge. However, the difference between WP and TO scores (Diffscore) is seen to decrease as age increases (significantly different). Word prediction appears to have resulted in greater increases in spelling accuracy scores in the younger subjects. This difference between WP and TO scores was not noted between increasing grades.

As children's ages increased they tended to type faster.

4.4 The research design analysis

The data resulting from the research project has been analysed with respect to the research questions, and has also been related to some of the characteristics of the sample group. However, it is still necessary to analyse the effect that the research design itself had on the results. As will be seen, additional insights into the application of word prediction for children with spelling difficulties can be gained from an analysis on the research design itself. Four aspects were investigated, namely

1. The equivalence of the four sample groups
2. Variations in the results of the four sample groups
3. Wordlists
4. Presentation order

4.4.1 The equivalence of the four groupings of subjects

Table 4.7 indicates the distribution of the sample across the four sample groups, with respect to the number of subjects, age, GWST scores, spelling age, grade and gender. It can be seen that the 80 subjects were fairly well represented across the four sample groups with respect to the independent variables.

Table 4.7 : The distribution of the subjects in the four sample groups with respect to number of subjects, age, GWST scores, grade and gender

Sample Group	Number of subjects	Mean Age	Mean GWST score	Mean Spelling age	Number of subjects – grade			Number of subjects - gender	
					4	5	6	Boy	Girl
1	22	11.96	20.50	8.31	6	8	8	15	7
2	19	11.79	20.26	8.26	6	6	7	13	6
3	20	12.07	19.05	8.10	5	7	8	12	8
4	19 (1) *	12.20	19.21	8.08	5	7	7	16	3
Total	80 (1) *	12.00	19.76	8.18	22	28	30	56	24

* Subject deceased (see section 4.4.1.1)

Duncan's Multiple Range Test was applied to the 4 sample groups. No significant differences between the groups with respect to age, GWST score or spelling age were found.

4.4.1.1 Number

The study started with 81 subjects, divided into 4 groups. Group 1 had 21 subjects each and Groups 2 - 4 had 20 each. The discrepancies noted above were due to one procedural error (a subject from Group 2 was administered the Group 1 procedure) and the death of one subject from Group 4 (after the groups had been formed but before he did the research task).

4.4.1.2 Age

From Table 4.7 it can be seen that the subjects' average ages were well distributed across the four sample groups.

4.4.1.3 GWST scores and spelling age

The average GWST scores and spelling ages can be seen to follow a decreasing tendency across the four sample groups. This bias was a direct result of the sample group allocation procedure, where the subjects were divided into equal groups after being sorted according to their GWST scores. Group 1 would always have the highest GWST score of each successive group of 4 subjects, and Group 4 the lowest. In retrospect, the group allocation method should have been more controlled.

4.4.1.4 Grade

Although there were unequal numbers of each grade in the sample, the grades were well distributed across the four sample groups.

4.4.1.4 Gender

An exception to the equivalent distribution of the four sample groups is the gender distribution of groups 3 and 4, where the girls were distributed with an unbalanced weighting – whereas Groups 1 and 2 had 32% girls, Group 3 had 40% and Group 4 had 16% (the group average was 30%).

4.4.2 Variations in the results of the four sample groups

Although the four groupings of subjects have a high degree of equivalence, a very different picture emerges when the results of the research task are compared across the four sample groups.

Table 4.8 summarises the results of the research study with the means of each of the dependent variables for each of the four sample groups. Variations in the results of the four sample groups can be seen across all the dependent variables.

Table 4.8 : Test results across the four sample groups

Sample Group	Typing Only score (TO score)	Typing Only time (TO time) in secs	Word Prediction score (WP score)	Word Prediction time (WP time) in secs	Difference between Word Prediction and Typing Only score (Diffscore)	Difference between Word Prediction and Typing Only time (Difftime) in secs	Approxi-mations (Approx)
1	12.41	535	17.27	629	4.86	95	2.86
2	9.79	607	17.74	675	7.95	68	2.68
3	9.15	514	15.65	631	6.50	117	2.20
4	11.37	498	13.58	614	2.21	116	4.50
Average	10.73	538	16.10	637	5.38	99	2.94

The variations seen in the results of the research task are worthy of analysis. The research design factors – the wordlists used and the order of presentation - will be analysed and discussed. It will be shown that many of the variations seen in the results of the four sample groups can be explained by this analysis. Table 4.9 summarises all the results in terms of the wordlist used and the order of presentation.

Table 4.9 : Analysis of variance - Wordlist used and order of presentation

				Dependent variables													
				Typing Only score (TO score)		Typing Only time (TO time)		Word Prediction score (WP score)		Word Prediction time (WP time)		Difference between TO and WP scores (Diffscore)		Difference between TO and WP times (Difftime)		Approx-imations (Approx)	
Research design characteristics	Wordlist	AB	Mean	A	11.93	A	518	A	15.56	A	622	A	-3.63	A	-105	A	3.39
		BA		B	9.46	A	559	A	16.67	A	652	B	-7.21	A	-93	A	2.46
	Order	T1	Mean	A	11.20	A	568	A	17.49	A	650	A	-6.29	A	-82	A	2.76
		T2		A	10.23	B	506	B	14.64	A	623	B	-4.41	A	-117	A	3.13

Note :

1. Duncan's Multiple Range Test was applied to test which of the means were significantly different from each other. Where the groups have a different symbol, they are significantly different. Significantly different groups have been shaded.
2. The subjects tested with TO on Wordlist A and WP on Wordlist B are referred to as the AB group
   The subjects tested with TO on Wordlist B and WP on Wordlist A are referred to as the BA group

3. Those subjects who were tested on the TO subtest first and WP subtest second are referred to as the T1 group.
4. Those subjects who were tested on the WP subtest first and TO subtest second are referred to as the T2 group.

4.4.3 Wordlists

The design required two equivalent wordlists. Great care was taken to equate the two lists on as many criteria as possible in their compilation. However, the results have shown that despite the care taken, the lists had some areas of non-equivalence that had not been anticipated. Table 4.9 highlights the differences between the two wordlists.

Although the subjects forming the sample groups for the two wordlist combinations were closely matched, their performance on the two wordlists differed significantly.

4.4.3.1 Accuracy scores

The TO scores for the two wordlists were significantly different. Wordlist A appears to have been significantly easier to spell than Wordlist B for the TO task. Subjects doing TO on Wordlist A scored an average of 2.47 correctly spelled words more than those doing TO on Wordlist B. In the WP task, Wordlist A again appears to have been easier to spell than Wordlist B, although not as markedly as was seen in the TO task. In the WP task, subjects using Wordlist A scored on average 1.11 correctly spelled words more than those using Wordlist B.

As the research design required the subjects to do the TO task on one wordlist and the WP task on the other wordlist, the differences between the two lists were compounded in the research task. The subjects who were tested on TO with Wordlist A (the easier of the wordlists with respect to TO) and WP on Wordlist B (the more difficult of the wordlists with respect to WP), would show less improvement with word prediction than the other subjects who were tested on TO with Wordlist B (the more difficult of the wordlist with respect to TO) and WP with Wordlist A (the easier of the wordlists with respect to WP). In the research test, the AB subjects achieved a mean of 3.63 more correctly spelled words with word prediction, whereas the BA subjects scored a mean of 7.21.

In Figure 4.4 all the TO and WP scores of Wordlists A and B were sorted in ascending order and plotted on a single graph. The large difference between the TO scores on Wordlist A and the TO scores on Wordlist B can be seen clearly, as can the smaller difference between the WP scores on Wordlist A and the WP scores on Wordlist B. Figure 4.4 also clearly shows the greater score difference between the BA group (large arrow) compared to the AB group (small arrow).

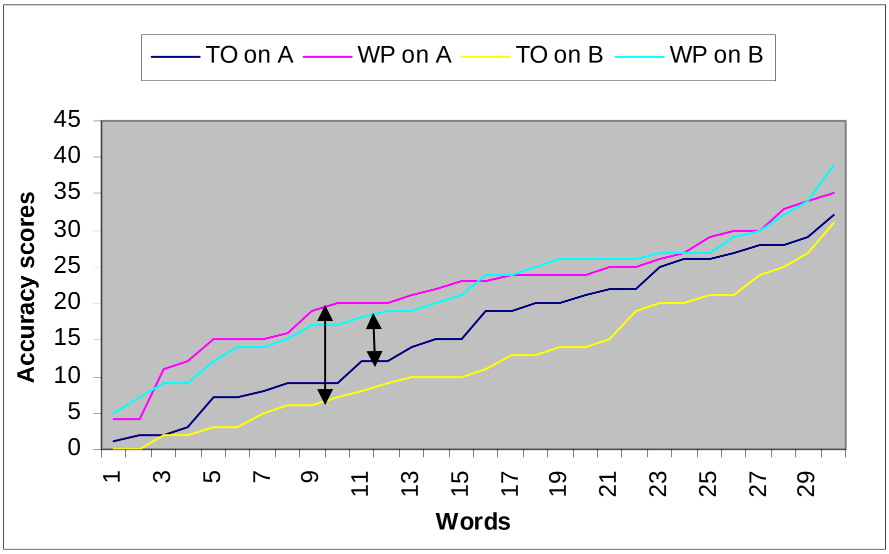

Figure 4.4: TO and WP accuracy scores on Wordlists A and B

The origin of the differences between the two wordlists was investigated.

i. Factors impacting on TO accuracy scores

The non-equivalence of the two wordlists is difficult to understand. Word pairs were carefully chosen to represent words that had the same phonetic principles, form, syllables and number of letters.

Although there were 41 subjects who did the task TO on Wordlist A compared to 39 doing the TO task on Wordlist B, the difference between the two lists would still be marked even had the 2 extra subjects received full marks for their TO test.

Most (but not all) of the unequal word pairs were in favour of Wordlist A. Wordlist A received 489 correct responses whereas Wordlist B received only 369 responses, resulting in a 120 difference between the total number of correct responses achieved on each wordlist. The differences between the TO scores between Wordlist A and Wordlist B varied from –7 to + 18.

In Figure 4.5 the TO scores of each word in Wordlist A were arranged in ascending order and the TO scores of the corresponding words in Wordlist B plotted alongside them. The graph shows the extent to which the word pairs were not well matched. A few of the words of Wordlist B were easier (received higher scores) than the corresponding word pairs in Wordlist A, but the majority of the words in Wordlist A were easier than the corresponding word pair in Wordlist B.

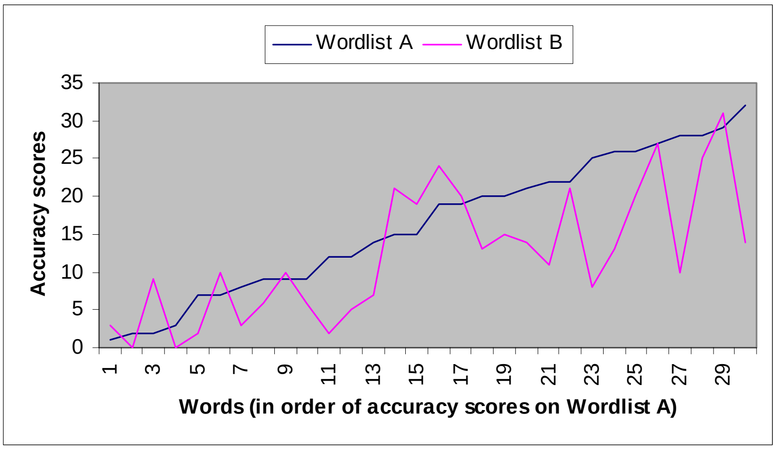

Figure 4.5: TO accuracy scores on Wordlists A and B

It would be very interesting to understand what makes two words well matched in terms of spelling, but currently there is no well-founded explanation for the discrepancy seen in the above results. It could be that in the compilation of the wordlists for the current research task, the researcher tended to form word pairs starting with the more familiar word, which would be placed in Wordlist A, following with the second word of the pair, which would be placed in Wordlist B. However, the impact of unmatched pairs is evident!

ii. Factors impacting on WP accuracy scores

The subjects also scored higher accuracy scores with word prediction on Wordlist A than on Wordlist B. The AB group scored on average 15.56 correct responses and the BA group scored on average 16.67 correct responses. Although the score difference is not as great as that with TO, it is still important to consider what the influences might have been to cause the difference in average accuracy scores.

Wordlist A resulted in a total of 650 correct responses and Wordlist B a total of 638 correct responses. The difference appears negligible but it must be remembered that the AB group had 2 more subjects than the BA group, so the difference is in fact larger.

The problem of unmatched pairs (i.e words that were not equally hard or easy to spell) may have been evident in WP, as it was in TO, but it is difficult to quantify how much it impacted on the results, as there were other influences too.

What is certainly clear is that there were extreme differences between the correct responses that words in each list elicited. In Figure 4.6, the word pairs were ordered according to the differences in scores between WP accuracy scores on Wordlists A and B (just as for Figure 4.5). Each list had words that were easier to spell with word prediction, but also words that were more difficult.

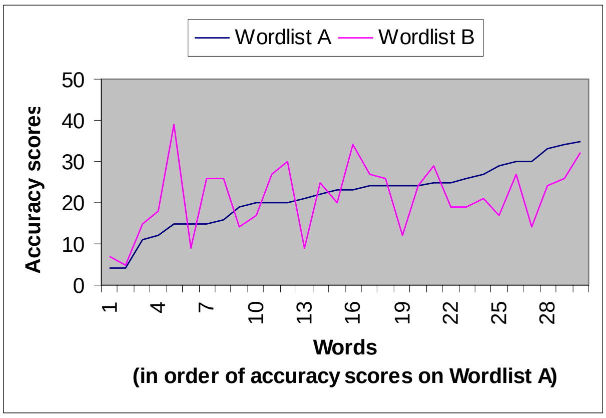

Figure 4.6 : WP accuracy scores on Wordlists A and B

The tendency to chose words that approximated the target word in the prediction list was more evident in the words from Wordlist B than those from Wordlist A, as can clearly be seen in Figure 4.7.   Wordlist A elicited 96 approximations, whereas Wordlist B elicited 139 approximations.

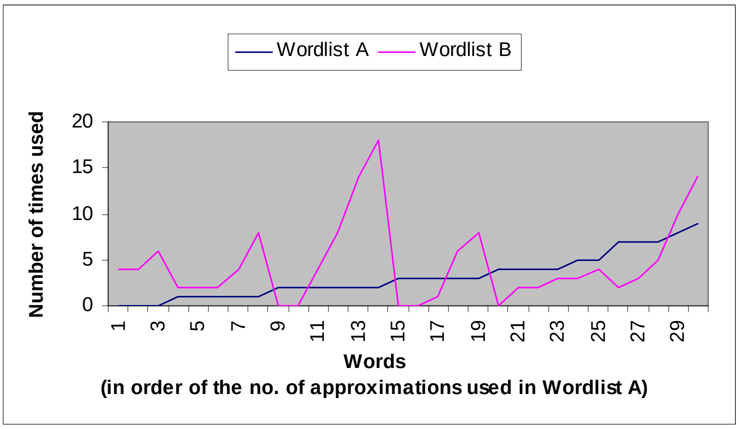

Figure 4.7 : Number of approximations used in Wordlists A and B

The factor of approximations was not considered in the compilation of the wordlists, as this problem had not been met in the literature search on the use of word prediction to improve spelling accuracy. It is clear that the effect they have on results in a study of the influence of word prediction on spelling accuracy is significant.

4.4.3.2 Time scores

The time differences between Wordlist A and B were not significant, for both TO and WP scores.

4.4.4 Presentation order

Although the subjects forming the sample groups for the two orders of presentation in the research test were closely matched, their performance on the two wordlists differed significantly (see Table 4.9). In the statistical analysis of the results, three variables were indicated that could be considered significantly different with respect to the presentation order of the wordlists – TO time, WP score and Diffscore.

4.4.4.1 Accuracy scores

In Figure 4.8, the TO scores of the two wordlists for the two presentation orders were each sorted in ascending order and plotted on the same graph to get an indication of how the two wordlists related to each other in terms of whether they were presented to the subjects first or second. Figure 4.9 is a similar graph for WP scores.

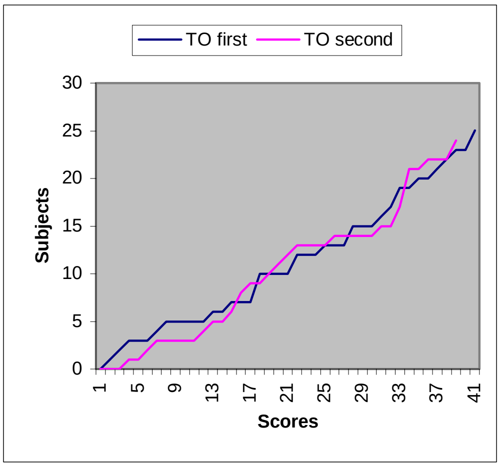

Figure 4.8 : The relationship between TO accuracy scores and presentation order

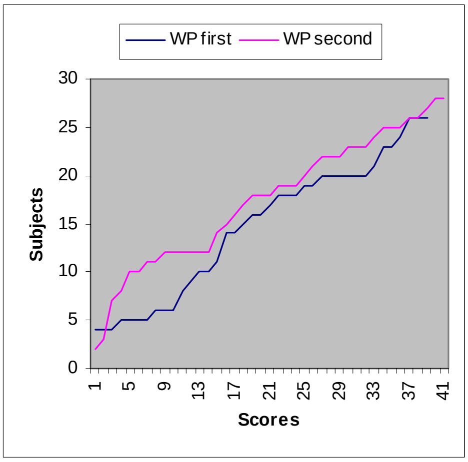

Figure 4.9 : The relationship between WP accuracy scores and presentation order

The T1 group scored significantly higher on the WP subtest, which for them was presented second in the research task (see Table 4.9 and Figure 4.9). When the WP subtest was administered second, the subjects spelled 2.85 more correct words. This may be due to the increased familiarity of the on-screen keyboard. The practice gained in spelling with the on-screen keyboard during the TO subtest appears to have more than compensated for any fatigue there may have been after the TO subtest. Anson (1993) indicated the influence that familiarity of use had in using on-screen keyboards (see parameter 8 in Appendix N). The skill required in managing a new keyboard and a new way of spelling at the same time was probably costly on accuracy scores for the T2 group.

It can thus be seen that the T1 group scored higher spelling accuracy scores than the T2 group in both the TO subtest and the WP subtest.

4.4.4.2 Time scores

The subjects typed the TO subtests significantly faster if they were tested on the TO subtest second (see Table 4.9). This indicates the influence on typing speed of practice on the keyboard. It would appear that if typing time could decrease during the test procedure due to increasing familiarity with the on-screen keyboard, extended practice could possible have a very marked effect on decreasing time.

WP time was not significantly different between those subjects who typed the WP subtest first or second (see Table 4.9). This may indicate that the visual-cognitive issues involved in word prediction were probably stronger than the influences of keyboard familiarity.

It can thus be seen that the presentation order of the wordlists had a significant impact on both the accuracy scores and time scores of the research project. Fatigue and keyboard practice may have been the reasons for the differences.

4.4.5 Variations in the results of the four sample groups

Having determined the relationships between the various subject characteristics (the independent variables) and the results, as well as the research design factors that influenced the results, understanding the result variations in the sample groups is now possible. (Refer to Table 4.8).

Why did Group 2 show the largest difference between WP and TO scores (7.95)?

- Groups 2 and 3 were administered the TO on Wordlist B and WP on Wordlist A, the wordlist test pair which resulted in the larger Diffscores (BA group).
- Groups 1 and 2 were administered TO first and WP second, the presentation order which also resulted in larger Diffscores (T1 group).

Group 2 was uniquely placed to reflect the greatest Diffscore of all the groups, being a group that had both the most advantageous wordlist combination and the most favoured presentation order.

Why did Group 4 show the smallest difference between WP and TO scores (2.21)?

- Groups 1 and 4 were administered TO on Wordlist A and WP on Wordlist B, the wordlist pair which resulted in the smallest Diffscores (AB group).
- Groups 3 and 4 were administered TO second and WP first, the presentation order which also resulted in smaller Diffscores (T2 group).

Group 4 was uniquely placed to benefit the least from the wordlist combination as well as the presentation order.

A similar process of reasoning will explain all the major variations in the above table. The large average number of approximations in Group 4 may simply be because of the unusually high score of 22 'approximations' contributed by Subject 18, who was in Group 4.

4.5 Summary

The research question was addressed by discussing each of the 3 sub-questions as well as the cost of approximations on word prediction usage. Typing with the use of word prediction improved the spelling accuracy of the majority (88.75%) of the sample by about 50%. However, a large variation in the responses of the subjects to word prediction was noted. Spelling accuracy improvement was at the cost of time. For the majority (91.75%) of the subjects, there was an increase of about 18.2% in time required to type with word prediction software. It appears that spelling knowledge does not have a direct relationship with the degree to which a child with spelling difficulties will benefit from word prediction use. However, there was a relationship between time benefits and spelling knowledge, indicating a decrease in time cost as spelling knowledge increases. Approximations were a frequent error type in word prediction use in the research study and it is proposed that it could be a significant cost in word prediction usage generally. Approximations tended to decrease as grades and spelling knowledge increased.

Some differences in the results related to subject characteristics emerged in this study. All three grades benefited from word prediction to the same degree, but a decreasing tendency to use approximations and a decreasing time cost with increasing grades was noted. Girls spelled better than the boys but did not type faster, or benefit from word prediction more. Word prediction appears to have resulted in greater increases in spelling accuracy scores in the younger subjects.

Although there was high degree of equivalence between the four sample groups in the research project, there were important variations in the results between the four groups. This was due to differences in the two wordlists used, and the impact of the presentation order of the subtests on the performance of the subjects. However, the cross-over design of the study served to counter-balance the effects of these variations.

Chapter 5

Critical evaluation of study

5.1 Introduction

This study emerged from a concern about the poor ability of school children with spelling difficulties to participate in functional writing activities, the foundational tool for all educational progress in school. It was assumed that if the mechanics of writing (such as spelling) could be supported, then the higher levels of writing could proceed with less interference from lower level deficiencies. A literature survey indicated that there were many features of word prediction that could theoretically benefit a child to spell better and to generate words faster (see Chapter 2). But it was necessary to ascertain if children with spelling difficulties could operate word prediction software with sufficient skill in order to benefit from its use in terms of improved spelling accuracy, without too heavy a cost in terms of time. This study therefore sought to investigate if word prediction could enhance the spelling accuracy and increase the on-screen typing speed of young children with spelling difficulties. Also of interest was the predictive value of understanding which children, in terms of spelling competence, would be most likely to benefit from the use of word prediction.

5.2 Summary of the results

5.2.1 Spelling accuracy improvements

Spelling accuracy in the research task increased 50% when typing with word prediction.

However, there was a large variation noted across the subjects' responses. Spelling accuracy improvements correlated to neither spelling knowledge, grade nor gender, but did show a negative correlation to chronological age. As the children's age increased so the degree of benefit with word prediction decreased.

This wide range in the subjects' responses is a caution in the decision to introduce word prediction as writing support for a child with spelling difficulties, without greater awareness of the individual's response to word prediction. The literature strongly supported this caution (see Chapter 1). There is much indication that the specific characteristics or skills of the individual, as well as the interaction between those characteristics and the specific parameters of the word prediction software itself, impact significantly on the apparent efficacy of word prediction usage. Many of these influences are still not fully understood. It is thus advised that decisions to offer word prediction to children with spelling difficulties should be made on an individual basis, and then their performance with it evaluated.

5.2.2 Typing speed enhancements

Almost all the subjects had increases in typing speed when using word prediction, the average increase being 18.2%. This is an important aspect of word prediction to consider, as potential benefit must be measured against the cost to the subject of that benefit.

As with spelling accuracy improvements, there was a wide range of responses in this study with respect to typing times in word prediction usage. There appeared to be no significant correlation between typing times and gender, age or grade, but a decrease in the time cost associated with word prediction use was seen as spelling knowledge increased.

The time cost is probably related to the more cognitively demanding process of forming and recognising words in word prediction use, and would probably be exacerbated in writing tasks more complex than single-word spelling.

5.2.3 'Predictive' value of a spelling test

No correlation could be found between spelling knowledge (as determined by the administration of a spelling test) and the degree to which the subjects benefited in spelling accuracy improvements. What the study does indicate, however, is that as spelling knowledge increases, so the children are able to benefit from time savings in word prediction usage.

The decision to offer word prediction to children with spelling difficulties as writing support can therefore not be made on the extent of the spelling difficulty alone

5.2.4 Approximations

An unexpected additional cost in the use of word prediction software was the tendency of the subjects to chose words from the prediction list that were similar to the target word, but not the correct one. Approximations were a frequent error type and had a significant impact on the results. The use of approximations reflected an immaturity in sight word recognition. The tendency to use approximations decreased with increasing grade as well as increasing spelling knowledge and ability.

5.2.5 Influence of subject characteristics

The influence of some subject characteristics on word prediction effectiveness in this study was investigated. The girls spelled better than the boys but did not benefit from word prediction more than the boys, both with respect to spelling accuracy improvements and speed enhancements.

There was increasing spelling accuracy, with and without the use of word prediction, as grades increased, and a decreasing tendency to use approximations. All grades appeared to benefit from word prediction to the same degree with respect to accuracy improvements, but there was a decreasing time cost as grades increased.

There was improvement in word accuracy (but not time enhancement) associated with increasing chronological age. However, the degree of benefit offered by word prediction use decreased with increasing age. The younger children benefited more from word prediction use than did the older children.

5.2.6 Influence of the research design

Although there were no significant differences between the four sample groups, and the subjects in the four groups were matched and comparable, the test results of the four sample groups of the main study were different from each other. The main factors causing this discrepancy were shown to be the differences between the two wordlists and the order in which the lists were presented in the task. However, it is important to note that the cross-over counterbalance design of the study was selected for the purpose of counter-balancing the effects any differences between the wordlists and the presentation order might have had on the results.

The two wordlists were not as equal as was hoped, in terms of spelling difficulty, despite the care that was taken in their compilation. The effect of the unequal wordlists was seen in both the Typing Only subtest as well as the Word Prediction subtest. Inequality between the two wordlists was mainly due to unmatched word pairs, resulting in Wordlist B being more difficult to spell than Wordlist A. Also, Wordlist B had more words than Wordlist A that elicited approximate choices. The inequality of the wordlists highlighted the impact of approximations in word prediction use and also showed the range of words that children might choose when selecting their target word in the prediction list. This range indicated the importance of sight-word reading skills in the effective use of word prediction.

The order of presentation had a significant impact on the results, indicating the influence of practice and fatigue in the performance of the subjects. The practice advantage for subjects was probably in the use of the on-screen keyboard. The literature supports the strong influence that practice at the skill of on-screen keyboard use might have (Anson, 1933). The impact on fatigue of the order of presentation of the word prediction subtest served as a caution regarding the increased cognitive demands required to operate the software. It may be that the children could not sustain the cognitive focus required to manage word prediction over a long period of time.

5.3 Critical evaluation of the study

5.3.1 Strengths of the study

The cross-over counterbalance design of this study was its primary strength. The design counterbalanced the threats to the internal validity of the study caused by the interaction of various aspects of the design (DePoy & Gitlin, 1994). This design was important as the degree of equivalence between the two wordlists used for the testing and the impact that the order of presentation would have on the two subtests was unknown prior to testing.

Although it was seen that both the wordlist used and the presentation order had a marked impact on the results, the results are still valid due to the strength of the design. In addition, information concerning unmatched pairs and approximations would not have been highlighted if it were not for the inequalities present in the two wordlists. In the same manner, information concerning the impact of on-screen keyboard training on the effectiveness of word prediction, and the important effect of fatigue in word prediction would not have been available were it not for the cross-over counterbalance design of this study.

The wordlists appeared to satisfy the requirements of providing a spelling list for the main study that accurately assessed single-word spelling, with accuracy scores over a wide range of scores. The high correlation between the Graded Word Spelling Test and the Typing Only subtests showed that the wordlists were appropriate.

That the four sample groups had a high level of equality with respect to the subject characteristics was a good foundation upon which to analyse the results after testing. The sample size of 81 subjects was also a positive feature of the study.

The word prediction training programme that was designed to familiarise the children with the principles of word prediction appeared to be adequate and suitable. Not one child failed the five-word test given to the children after the training. This also indicated that young children, even those with severe spelling difficulties, could potentially use word prediction.

The task that the children were required to execute proceeded without difficulties. All the logistical flaws of the task had been highlighted and dealt with during the pilot study. Only the video recording of a few of the testing sessions was not always successfully executed (due to insufficient tape during long sessions and failure to activate recording). Also, there were unexpected interruptions during the testing procedure, but they were compensated for (see Appendix V).

The results of this study were able to provide clear answers to the research questions. The impact of approximations, not met in the literature, was presented.

5.3.2 Factors negatively and positively influencing results

There were some features of the research design which probably negatively influenced the results, thereby underestimating the potential benefit that word prediction might have in improving the spelling accuracy of children with spelling difficulties in typical usage of word prediction software.

Many of the accuracy and speed enhancing features of word prediction software were not utilised in this study. It is important to note that the task did not allow for access to some of the features of word prediction that the literature has shown could positively influence accuracy scores, such as speech output, frequency adjustments, text specific vocabulary, automatic learning of vocabulary and automatic space savers (see Table 2.1). For example, the automatic space facility, which would decrease by 27 keystrokes the minimum number of keystrokes used to type the 30 words in the wordlists with word prediction software, was not utilised in the research task. This negatively impacted the time scores achieved in the research task, thereby leading to underestimation of the time benefit word prediction could offer due to keystroke reduction.

The literature pointed out the positive influence of training in, and practice with, word prediction, increasing the influence that word prediction could have on improving spelling even further (see section 2.8). The subjects in this study had only rudimentary training and practice in word prediction use. The research project design did not include practice beyond a

10 minute training session. The modest improvements in word prediction usage in the research project could translate into significant improvements with more training and practice, with respect to both spelling accuracy and speed. There is sufficient evidence to hope that the frequent, guided use of word prediction will also serve to develop the writing capabilities of the user (Zordell, 1990).

Just as there were features in the research design which could have negatively influenced the results, so were there also features that could influenced the results positively. The single-word spelling task used in the research task was a feature of the research design which probably led to overestimation of the potential benefit that word prediction might have in improving the spelling accuracy of children with spelling difficulties. The words on the research task had a high probability of being present in the prediction list. A routine of typing letters and looking for the target word in the prediction list was established in the training. This does not reflect the functional writing situation, where words will often not appear in the prediction list at all and some words will be too short to bother to find them in the prediction list. The children may stop looking for the target words if they are often disappointed at not finding them there, or they may find the decision making required as to when to look overly demanding on their cognitive resources.

Assessing the factors that could have influenced the results in this study, it is considered that the effectiveness of word prediction to improve spelling accuracy and increase typing speed is underestimated in this study.

5.3.4. Limitations to the study

The single-word format of the research task was a limitation to the study. The accuracy improvements and speed enhancements with the use of word prediction seen in single word spelling in this study may not be realised in functional writing. Functional writing requires the ability to form words (i.e. to spell) as a foundational skill, but also requires the integration of a number of higher level skills not required in single-word spelling. However, the single-word format of this study effectively served to isolate the impact of spelling knowledge on word prediction use where there are spelling difficulties.

The Graded Word Spelling Test, used to determine the spelling knowledge and ability of the children, was limited in its effectiveness as a tool for assessment. It was not ideally suited to the South African situation as it was standardised to the British population and British educational system, and its age-norms were difficult to apply due to the large difference in school-starting age between Britain and South Africa. There was no test of its kind standardised for the South Africa population . The Graded Word Spelling Test appeared to be the most suitable alternative available to South Africans. The Graded Word Spelling Test was shown to be well suited to this study as there was a high correlation between the Graded Word Spelling Test and the Typing Only subtest (see section 4.2.3).

5.4 Recommendations for further research

Research that investigates the improvements in spelling accuracy with the use of word prediction across tasks other than single-word tasks would contribute to the understanding of the impact of spelling difficulties on the ability to use and benefit from word prediction. These would include tasks such as copying, dictation, formal sentence construction in class worksheets and free writing tasks. Each of the above tasks places cognitive demands on the user that are not present in single-word tasks, such as scanning from an external source, memory demands, grammar demands, decision making and internal message generation concurrent with external sentence construction. There would be benefit in designing studies that would attempt to understand how the performance of the subjects differs across a variety of different spelling tasks. This could be attempted by isolating the various skills and designing tasks that address each of them independently, or by designing an integrated task that combines many spelling tasks into a comprehensive functional writing task.

An investigation into improvements in spelling accuracy with the use of word prediction after more extensive training and practice would be valuable for practical application of word prediction as an educational tool to support writing where there are spelling difficulties.

Understanding the influence of cognition and chronological age on the benefit that word prediction may be able to offer children with spelling difficulties would add to the understanding of which children would benefit from word prediction.

As this study investigated the relationship between spelling skills and word prediction efficacy, a similar study could investigate the relationship between reading skills and word prediction efficacy.

An investigation into the impact of approximations in the use of word prediction may be useful in understanding how spelling knowledge relates to effective word selection from the prediction lists.

There was little research available to aid in the development of well-matched word pairs for the two wordlists. It would be useful to understand which factors make two words well matched in terms of spelling difficulty.

5.5 Summary

Word prediction has been shown to be a potentially useful tool in supporting children with spelling difficulties to spell more accurately, despite the costs associated with its use. However, the individual with spelling difficulties' response to word prediction usage varies and is difficult to predict. The results and limitations of this study have raised many questions about the effectiveness and use of word prediction, which can be addressed through further research.

References

Allred, R.A. (1990). Gender differences in spelling achievement in Grades 1 through 6. *Journal of Educational Research, 83(4)*, 187-193.

Anson, D. (1993). The effect of word prediction on typing speed. *The American Journal of Occupational Therapy, 47(11)*, 1039-1042.

Berninger, V. W. (2000). Development of language by hand and its connections with language by ear, mouth, and eye. *Topics in Language Disorders, 20(4)*, 65-84.

Berninger, V.W., Abbott, R.D., Abbott, S.P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand language by eye. *Journal of Learning Disabilities, 35(1)*, 39-57.

DePoy, E., & Gitlin, L.N. (1994*). Introduction to research. Mulitple strategies for health and human services.* (pp. 109-110). St Louis : Mosby-Year Book, Inc.

Crick, J. (2000). *Clicker 4 User Guide*. Northampton : Crick Software Ltd.

Ganske, K. (1999). The Developmental Spelling Analysis: A measure of orthographic knowledge. *Educational Assessment, 6(1)*, 41-70.

Gillette, Y., & Hoffman, J.L. (1995). Getting to word prediction: Developmental literacy and AAC. *Proceedings of the Tenth Annual International Conference, Technology and Persons with Disabilities.*

Higginbotham, D.J. (1992). Evaluation of keystroke savings across five assistive communication technologies. *Augmentative and Alternative Communication, 8(4),* 258-272.

Horstmann, K.H. & Levine, S. (1990). Modeling AAC user performance: Response to Newell, Arnott, and Waller (1992). *Augmentative and Alternative Communication, 8(2)*, 92-97.

Hunnicutt, S., & Carthberger, J. (2001). Improving word prediction using Markov models and heuristic methods. *Augmentative and Alternative Communication, 17(4)*, 255-664.

Klund, J., & Novak, M. (1995). If word prediction can help, which program do you choose? *Closing the Gap Conference*.

Koester, H.H., & Levine, S.P. (1996). Effect of a word prediction feature on user performance. *Augmentative and Alternative Communication, 12(3),* 155-168.

Koester, H.H., & Levine, S.P. (1997). Keystroke-level models for user performance with word prediction. *Augmentative and Alternative Communication, 13(4),* 239-257.

Koester, H.H., & Levine, S.P. (1998). Model simulations of user performance. *Augmentative and Alternative Communication, 14(1),* 25-35.

Lesher, G.W., Moulton, B.J. & Higginbotham, D.J. (1998). Techniques for augmenting scanning communication. *Augmentative and Alternative Communication, 14(2),* 81-101.

MacArthur, C. A. (1999). Overcoming barriers to writing: Computer support for basic writing skills. *Reading and Writing Quarterly, 15(2),* 169-192.

MacMillan, J.H., & Schumacher, S. (2001). Descriptive statistics. *Research in Education : A Conceptual Introduction. 5th Ed.* (pp. 204-236). New York : Longman.

Newell, A. F., Arnott, J., & Waller, A. (1992). On the validity of user modelling in AAC: Comments on Horstman and Levine (1990). *Augmentative and Alternative Communication, 8(2),* 89-92.

Newell, A.F., Booth, L., Arnott, J., Beattie, W. (1992). Increasing literacy levels by the use of linguistic prediction. *Child Language Teaching and Therapy,* 138-156.

Rogers, S.L, Gordon, A. Y., Schanzenbacher, K.E. & Case-Smith, J. (2001). Common diagnosis in paediatric Occupational Therapy practice. *Occupational Therapy for Children*. In J. Case-Smith (Ed.). St Louis : Mosby.

Shanahan, T., & Lomax, R. G. (1986). An analysis and comparison of theoretical models of the reading-writing relationship. *Journal of Educational Psychology, 78(2),* 116-123.

South Africa (Republic). (1996). *National Education Policy Act no 27 of 1996*. Cape Town, South Africa: Office of the President.

South Africa (Republic). (1996). *South African Schools Act no 84 of 1996*. Cape Town, South Africa: Office of the President.

Spooner, R. (1999) *Penfriend for Windows*. (1999). Edinburgh : Penfriend Ltd.

Swiffin, A.L., Arnott, J. L., Pickering, J.A., & Newell, A. F. (1987). Adaptive and predictive techniques in a communication prosthesis. *Augmentative and Alternative Communication, 3*, 181-191.

Szeto, A.L., Allen, E.J., & Littrell, M. C. (1993). Comparison of speed and accuracy for selected electronic communication devices and input methods. *Alternative and Augmentative Communication, 9(4)*, 229-242.

Tam, C., Reid, D., Naumann, S., & O'Keefe, B. (2002). Effects of word prediction and location of word prediction list on text entry with children with spina bifida and hydrocephalus. *Augmentative and Alternative Communication,18(3)*, 147-162.

Van der Leij, A., & van Daal, V.H.P. (1999). Automatisation aspects of dyslexia: Speed limitations in word identification, sensitivity to increasing task demands, and orthographic compensation. *Journal of Learning Disabilities. 32(5)*, 417-428.

Venkatagiri, H. S. (1993). Efficiency of lexical prediction as a communication acceleration technique. *Augmentative and Alternative Communication, 9(3)*, 161-167.

Venkatagiri, H. S. (1994). Effect of window size on rate of communication in a lexical prediction AAC system. *Augmentative and Alternative Communication, 10(2)*, 105-112.

Vernon, P.E. (1998). *Graded Word Spelling Test*. (2nd ed). London : Hodder & Stoughton Educational.

Williams, S. C. (2002). How speech-feedback and word-prediction software can help students write. *Teaching Exceptional Children, 34(3)*, 72-78.

Zordell, J. (1990). The use of word prediction and spelling correction software with mildly handicapped students. *Closing The Gap, 9(1)*, 10-11.