Doran, John (1997) *Inferencing skills of deaf adolescent readers.* PhD thesis.

Inferencing Skills of Deaf Adolescent Readers

John Doran

February, 1997.

Thesis presented in fulfilment of the requirements for the award of the degree of Ph.D.

Department of Psychology,
Faculty of Social Sciences,
University of Glasgow.

Abstract

The great difficulty which deaf children have in learning to read is well documented. Previous studies have examined such aspects as problems with vocabulary and syntax but little work has been directed towards inferential and discourse skills. The present series of studies examines the inferencing skills of severely and profoundly deaf adolescents.

Different types of inference were examined using a variety of experimental techniques, ranging from on-line reading times, through memory probes after reading, to tracking the movements of subjects' eyes as they read. The deaf were found to be poorer at drawing inferences than hearing children matched on reading age, although they can recall as much detail from those texts in which they do infer correctly as the controls. The deaf were as successful as the reading age matched controls for material which required spatial inferences but not for more abstract temporal and causal inferences.

On-line studies suggested that the deaf, when drawing inferences, use a schemata, concept driven mechanism similar to hearing peers approximately matched for chronological age. A similar mechanism would seem to be operating when material is presented in the form of sign language. Thus many of the difficulties previously ascribed to deaf children's reading skills may in fact derive from more general language problems. It is suggested that these difficulties with inferencing are independent of modality of presentation and perhaps reflect a more impoverished experiential background for most deaf children. The comparative richness of scenarios for deaf and hearing children are then investigated.

A final study examined the eye movements of deaf and hearing children as they read. The findings confirmed that both groups do make inferences on-line but that these inferences take a small but finite time to generate. Their performance is compared to evidence from studies with adults and a scenario driven, situational model is proposed.

Acknowledgements

I would like to sincerely thank my supervisor, Dr Anne Anderson, Glasgow University, for all her support and guidance, and her unremitting patience. I would also like to thank Professor Simon Garrod, Glasgow University, for additional advice and suggestions. A special thank you goes to Mr Frank Shoveller, Audiovisual Services, Glasgow City Council Education Department, for writing the computer control programme for the signing study reported in chapter seven.

I must also thank my wife and son, without whose constant support and encouragement, none of the work reported in this thesis would be possible.

A final thanks goes to all the children who participated in the studies reported herein, as well as to their parents and teachers.

Contents

List of Tables

List of Figures

Chapter One

Reading and deaf children: A review of the literature

Reading is a basic skill central to any child's education. It is the gateway to all other forms and sources of knowledge. For the deaf child, it is a particularly difficult skill to acquire. Despite centuries of diligent effort by teachers, the sad fact remains that the typical deaf child cannot read as well as the typical hearing child.

1.1 Achievement Levels

Over the years many studies have shown the very poor achievement levels of severely and profoundly deaf students, eg DiFrancesca (1972) in the USA and Conrad (1977) in England and Wales. A typical 16 year old deaf student is found to be reading at the 8.5 to 9 year level using standard tests of reading. There also seems to be a levelling off in attainment between the ages of 11-16 years, a "plateau" effect (Brooks, 1978; Reich & Reich, 1974). Furth (1966) analysing data obtained by Wrightstone, Aronow and Moskowitz (1963) showed that for deaf children, their average reading age scores increased by less than one year between the ages of 10-11 and 15-16 years. These findings are supported in a later study by DiFrancesca (1972) who calculated an average growth in reading achievement age of 0.2 years per year of schooling. Hammermeister (1971) showed that 7 to 13 years after leaving school there was a significant improvement in word meaning (ie vocabulary) but not in paragraph (ie connected language) meaning for deaf subjects.

1.2 Vocabulary

Numerous authors, eg Fusaro & Slike (1979), Kyle (1980b), Hatcher & Robbins (1978), and Griswold & Cummings (1974), who have examined vocabulary development in the deaf, have confirmed that deaf students of all ages can comprehend from print substantially fewer words than hearing students and that the distribution of types of words, such as nouns and verbs, is different between the two groups. Paul (1984) examined the performance of deaf and hearing subjects on a picture vocabulary test with a multiple-choice response section. Some items required only one correct response while other multimeaning words required two responses. Paul found that the hearing subjects performed better than the deaf subjects on selecting two meanings for the words. In addition, they outperformed the deaf subjects on selecting at least one meaning for the same words. For both groups, however, knowing two meanings was more difficult than knowing only a single meaning. Also the subjects chose the primary, or more common meanings, more often than the secondary meanings. Paul concludes that both deaf and hearing subjects have difficulty with the notion that a word may have several meanings, and not being aware of the most common meanings of words appearing frequently in print may contribute to reading comprehension problems in the deaf. Paul (1984) reported that deaf children, as well as knowing only the most common meanings of multi-meaning words, cannot use the available context clues to help derive word meanings.

Some of these studies have attempted to isolate various factors that influence vocabulary development in deaf children. For example, Hatcher & Robbins (1978) found that deaf children they studied had somehow developed vocabulary and reading skills beyond the level expected from their knowledge of primary word analysis skills, if indeed phonic skills are essential prerequisites for reading. Kyle (1980) looked at the development of basic perceptual processes such as letter identification, discriminating and matching letter shapes, and the ability to associate words and pictures together. Over a 3 year period these skills showed no developmental difference in deaf and hearing children between about 7 and 11 years of age.

A more recent study by Laybaert & Alegria (1993) used the Stroop technique to study word processing in deaf children. This technique (Stroop, 1935) measures response times to words printed in different coloured ink. Children have to name the colour of the ink but access to the word meaning interferes with this process, causing an increase in·response time. Leybaert & Alegria, using both verbal and manual responses, found that access to orthographic, semantic and even phonological information can occur automatically when deaf children read.

In other words, simply examining the "bottom-up" processes such as ability to learn sight vocabulary shows little difference between young deaf and hearing children. The deaf child's reading problems begin as soon as an attempt is made to understand more complex text. Webster (1986) states that reading tests up to about the 8.5 year range tap such low-level skills as word recognition but beyond this point require more complex skills and this accounts for the plateau effect.

Dodd & Hermelin (1977) and Dodd (1980) have suggested that deaf children can use phonological information acquired through lipreading in spelling tasks (and possibly also in reading) although they tend to prefer to use visual information. Moores (1967) compared the performance using the cloze procedure of a group of deaf students with a group of hearing students, matched for reading age. Students had to replace words deleted from the text. Using methods constructed to indicate how well the subjects used vocabulary and syntax in this procedure, Moores found the deaf to be significantly deficient in their vocabulary and syntax compared with the hearing group. O'Neill (1973) confirmed Moores' conclusions in a study of deaf and hearing children's knowledge of phrase structure. She found deaf children to be significantly less able to judge correctly the grammaticality of pairs of grammatical and ungrammatical sentences, even although both groups were matched on reading achievement levels. Moores' and O'Neill's studies both indicate that standard reading tests may give spuriously high estimates of reading levels in the deaf population. Wood (1984b) noted that the deaf have difficulty with "diectic" functional words, for example, pronouns referring to elements mentioned earlier in the text. He quotes Clark

(1978) on how the use of such a class of words develops in the hearing child over time, and states that he believes that the problems experienced by the deaf are due to their relative lack of adequate experience in using such words in meaningful situations. Wood argues that comprehension relies upon the complex relationship between words and the systematic ways in which their meanings change, and that deaf children, because of their lack of good auditory memory for sequences of words, are denied access to or adequate experience of such situations. Unlike "concrete" words which can be exemplified by pointing, "function" words can only be mastered by appreciating the place they occupy in connected text.

In summary, when lower level skills are considered, there is no difference between deaf and hearing children in their perceptual discrimination skills, letter discrimination and ability to learn sight vocabulary. However, the deaf have narrower vocabularies; they understand fewer words, and for words with more than one meaning, they usually only know the more common meanings.

1.3 Syntax

Another critical component of effective readers is the ability to use syntax in written text. Several researchers have examined the syntactic structures used by deaf readers, their approaches usually being influenced by the prevalent linguistic theories of the time. The earliest attempts followed the lines of traditional grammar (counting the number of nouns/verbs etc used), followed by several studies which used Chomsky's (1957, 1965, 1976) transformational generative grammar and generative semantics (Chafe, 1970; Fillmore, 1968: McCawley, 1968) Probably the most comprehensive examination of how the deaf understand and use standard English syntax was carried out by Quigley and his associates eg Quigley, Wilbur & Montanelli (1974), Quigley, Power & Steinkamp (1977), and Quigley, Smith & Wilbur (1974). A summary of the differences between the deaf and hearing populations is tabulated in Table 1.1. It can be seen from this table that the average 8 year old hearing student scores higher on the various tasks than the average 18 year old deaf student. It can also be seen that various syntactic structures typically appear in a standard school reading series at an age level long before the deaf students can comprehend them. This would indicate a serious reading problem based on syntax alone. When the typical vocabulary, conceptual and experiential problems of deaf students are added, it can be easily seen how such commonly used reading materials might present serious difficulties for many deaf students.

TABLE 1.1

Performance of deaf and hearing subjects on various syntactic structures. Quigley et al, 1976. (Percentage of each group who have mastered each construction).

	Deaf Students			Hearing Students
Structure	Averaged over all ages	Age 10	Age 18	Averaged over all ages
Negation	76%	57%	83%	90%
Conjunction	73	57	86	92
Question formation	66	46	78	98
Pronouns	60	39	78	90
Verbs	58	53	71	79
Relativization	54	46	63	82

For example, Robbins & Hatcher (1981) found that controlling for word recognition and training in word meaning did not improve performance in reading single sentences containing various syntactic structures for deaf children aged 9-12 years. They found passive voice sentences most difficult to comprehend, followed by relative clauses, conjunctions, pronominalisation and indirect objects, a pattern similar to that in Table 1. Some studies (eg Gormley & Franzen, 1978) have shown that deaf students seem to comprehend syntactic structures more easily in connected discourse than in single sentences. This will be covered later when we discuss discourse more fully.

Wood (1984), Webster (1986), and King & Quigley (1985) among others have noted also the rather rigid and simplified structures which the deaf tend to produce in their writing and interpret/misinterpret in their reading, such as the Subject-Verb-Object and the Minimal Distance Principal, which reflect these difficulties in syntax. For example, when given the two sentences, "The boy kissed the girl" and "The boy ran away", most deaf students understand them, but when one sentence is embedded in another to form the relativised sentence, "The boy who kissed the girl ran away", most deaf students (even those aged 18) thought it was the girl and not the boy who ran away. [An example of the Minimum Distance Principal]. Similarly, "The boy was helped by the girl" is interpreted as "The boy helped the girl". [An example of Subject-Verb-Object]. Webster (1986) believes that the deaf do make slow but consistent progress over time, which runs contrary to the idea that deaf children's reading ability and general language development plateaus out.

In short, the deaf experience great difficulty with normal English syntax. Passive constructions, relative clauses, pronominalisation, and ellipsis cause the greatest difficulty. This is reflected in how the deaf (mis)interpret text and also in the rather rigid constructions used in their written output.

1.4 Reading tests

Given the differences noted above, the question arises as to the design and construction of the various tests which have been used to assess the reading standards of deaf children.

Webster (1986) questions the validity of using reading tests with the deaf which have been standardised on a hearing population. He argues that if one accepts Kyle's contention that there is no difference in intelligence between deaf and hearing populations, then a deaf 16 year old will be much more sophisticated, emotionally, physically, and conceptually than a 9 year old hearing child and therefore comparisons between these two populations on the grounds that they have similar reading ages on a standardised test is really meaningless. Webster et al (1981) replicated Conrad's (1979) work using the Widespan Reading Test (Brimer, 1972). They compared deaf and hearing children matched on their reading ages, but looked at errors made using Goodman's (1969) miscue analysis. In the Widespan test the child has to choose one word from the first of a pair of sentences to complete a gap in the second sentence. According to Webster, the test demands certain metalinguistic skills, "a reflective or explicit awareness, a kind of disembedded thinking". An example of the type of questions asked and one type of error are noted below:

"The salt of the earth are
the hewers of wood and the
drawers of water"

"Alongside the wardrobe
stood a tall chest of ...."

The answer expected is, of course "drawers" but the deaf tended to answer with "wood". Of the errors made by the hearing children, 81% were "linguistic", ie a word which is both semantically and grammatically appropriate, but incorrect (as above). In other words, the child is aware of some the linguistic constraints, although he gets the item wrong. Non-linguistic errors were those responses where no obvious connection could be found between the meaning or structure of the sentence and the word inserted into it. The hearing impaired children made twice as many errors as the controls and only 52% of these showed any linguistic strategy. Those children who achieved the highest reading ages in both groups also made relatively higher proportions of linguistic errors than non-linguistic errors. Also most of the hearing children stopped the test at their ceilings but the deaf continued well beyond theirs, very often right to the end of the test. In other words the deaf seem to have no self-monitoring skills, which according to Clay (1977) are essential in the development of reading skills. Webster interprets this as perhaps being a legacy of teaching practice itself, the often unwitting expectation of a passive, unquestioning style of learning from the deaf mentioned earlier. In any case the test is not measuring a delayed but a different process in the deaf.

In another experiment, Wood, Griffiths, & Webster (1981) tried to ascertain if the results of the above experiment were test specific. They duplicated Webster's procedures using the Southgate test (1962) which requires the child to choose one of several words given at the end to complete each sentence. Again they found that the deaf made more errors than the hearing children (47% versus 18%). They also found that the deaf do not self-monitor and continue well beyond their ceilings. The deaf did not seem to be distracted by similar looking alternatives in the choices. But most deaf children did agree on their choices of answer albeit incorrect. For example,

"Birds are covered with.....trees, skirts, sky, nests, feathers."
"Ducks can.....pond, swim, water, farm, sing."

In the first example, the correct answer is, of course, "feathers" but deaf children tend to choose "nests". For the second, they choose "pond" rather than "swim". Therefore, the deaf seem to be choosing on the basis of word association, a sense of which words tend to appear together in everyday usage and experience. In other words, the deaf are keywording.

In another study by Beggs & Breslaw (1982), deaf and hearing children were compared on the Picture Aided Reading Test (Hamp, 1975). Here the child is given a printed word and has to choose the correct picture from four possibles. The deaf identified correctly a quite different subset from the hearing group; they got some of the later, more difficult items correct and some of the earlier, easier items wrong. There was also agreement among the deaf in their (wrong) choices. By giving the series of pictures only to another group of hearing children, Beggs and Breslaw found remarkable agreement in their choices with those of the deaf. Therefore, the deaf seemed to be choosing on the basis of picture "saliency", ie the visual appeal of the picture.

The above findings show that the deaf use different strategies to hearing subjects when performing on standardised tests of reading. Even on tests which have been specifically designed for and standardised on a deaf population, deaf subjects use strategies other than those which might be expected to be taking place during the normal reading process. This, of course, poses the question of what exacting measured by such standardised tests?

1.5 Phonological Recoding: Hearing Subjects

An important aspect of any model of reading is how the visual images of the printed letters are converted into meaning. What cognitive mechanisms mediate this process?

It is noticeably to most fluent readers that as they scan over a text, they silently vocalise in their head what they are reading.

Several researchers have used electromyographic techniques (EMG) to measure the muscular activity of the speech organs (lips, tongue, larynx, throat etc) while subjects are reading. For example, Edfeldt (1960) showed that EMG activity increased for more difficult texts than for easier texts. McGuigan (1967) and McGuigan and Bailey

(1969) found that as readers became more skilled the level of EMG activity decreased.

However, whether this subvocal activity serves any useful purpose is less clear. Hardyck, Petrovich & Ellsworth (1966) and Aarons (1971) attempted to reduce EMG activity by using feedback methods. Hardyck & Petrovich (1970), using such techniques, found that when subvocalisation is reduced, comprehension suffers. However, Taylor & Taylor (1983) have suggested that the fact that subjects have to pay attention to eliminating muscle activity could in itself interfere with comprehension of the text.

Another approach was to have subjects engage in some concurrent articulatory activity such as reciting a well known nursery rhyme, counting or repeating a nonsense phrase while reading a piece of text. This minimises the demands on the subjects attention but still involves the vocal muscles in a task not directly linked to the reading task. Kleiman (1975), using such concurrent tasks, showed that the meanings of individual words can be accessed without reference to inner speech. He felt that inner speech was more important for accessing working memory. Others, eg Levy (1975, 1977); Slowiaczec and Clifton (1980), have presented evidence which suggests that the inner speech is important in these memory processes as it aids the comprehension of connected discourse.

There are possible methodological criticisms of these suppression techniques. Firstly, any decrements in reading might be due, not to the suppression of subvocalisation, but to an increase in cognitive load of having to carry out two tasks simultaneously. Alternatively, the suppression task may not adequately suppress inner speech. There is evidence (eg Posner & Boies, 1971) that in a dual task situation, the task which is perceived by subjects as being less important can suffer. In other words, the concurrent vocalisation task might suffer rather than the reading task. However, subvocalisation is not exactly the same as phonological coding. Subvocalisation refers to activity (muscular movement or articulatory processes) in the speech tract while phonological coding refers to the mental representations of sounds (Rayner & Pollatsek, 1989, pp 189, 201-202).

Alternative techniques have been used to assess whether inner speech is activated during reading. Several studies, eg Baron (1973), Doctor & Coltheart (1980), Banks, Oka & Shugarman (1981), Treiman, Freyd & Baron (1983), Treiman, Baron & Luk (1981), Treiman & Hirsh-Pasek (1983), used homophones (words sounding the same but with different spellings and meanings). Thus for a sentence containing words which sound correct but are misspelled, if a speech code is being used, then subjects should have no difficulty reading this sentence compared to a normally written sentence. A corollary of this, however, is that it should be more difficult to judge whether a sentence is meaningful. The findings indicate that there was no difference in the time taken to make semantic acceptability judgements between the control sentences and those containing the homophones but more errors are made with the homophone sentences. The results have been interpreted as showing that subjects can access the meaning of individual words directly by visual means but that speech coding does occur postlexically and is essential for interpretation at the level of the phrase or sentence.

Another approach favoured by some researchers utilised tongue-twister sentences which contain several words with the same initial consonant. Just as it is more difficult to say these sentences, so it has been found that it takes longer to read them silently and orally compared with straightforward control sentences (Haber & Haber, 1982; Ayers, 1984). Also semantic acceptability judgements took longer and were less accurate for tongue-twister sentences (McCutchen & Perfetti, 1982; McCutchen, Bell, France & Perfetti, 1991). McCutchen and Perfetti argue that specific phonemes are activated during reading and that this occurs postlexically. They and others (McCutchen et al, in press; Haber & Haber, 1982) see the interference as being due to the similarity of the phonetic representations generated. Other investigators (Baddeley & Lewis, 1981) have suggested that the tongue-twister effects mentioned could be due to grapheme rather than phoneme repetitions. However these difficulties were partially unconfounded by McCutchen & Perfetti (1982) by their use of tongue-twister sentences with mixed grapheme stimuli. Thus, instead of:

The talented teenager took the trophy in the tournament.

they used a sentence where the tongue-twister remains but where two different consonants alternate:

The taxis delivered the tourists directly to the tavern.

Similar tongue-twister effects were found in the mixed-grapheme sentence as in the same-grapheme sentence. Thus, although these mixed grapheme stimuli reduce the visual and phonetic confounding inherent in tongue-twister sentences, they do not eliminate it entirely.

An alternative strategy has been used by McCutchen, Bell, France, & Perfetti (1991). They asked subjects to judge the semantic acceptability of tongue-twister sentences but added a concurrent memory load. Two types of tongue-twister were used, either *alveolar fricative:*

eg *The sparrow snatched the spider swiftly off the ceiling.*

or *alveolar stop:*

eg *The taxis delivered the tourists directly to the tavern.*

Before seeing the sentences, five numbers were presented on the screen for subjects to memorise. The numbers either began with alveolar stops (eg 12, 2, 20, 25, 22) or with alveolar fricatives (eg 17, 6, 65, 16, 77). Subjects then saw the sentence, had to judge if it was semantically acceptable and then had to recall as many numbers as possible. McCutchen et al (1991) found that if the numbers started with a fricative then it took longer to respond to fricative sentences, if the numbers were alveolar stops then it took longer to respond to stop sentences. They also found the same interaction with respect to recall of numbers, with more errors when numbers and sentences shared initial phonemes. Since the memory load numbers were visually distinct from each other (7, 6, 16, 77) as well as from the sentence graphemes, it clearly demonstrates that the interference was not visual. Rather, it is specific interference between the phonetio content of the sentence and the memory load.

These studies show that phonological coding is important for hearing subjects during the reading process. Lexical access can occur directly from visual input and bypass phonological coding. However, phonological coding seems to be important as it allows retention in short-term memory of the surface structure of the text so that meaning can be derived across clauses and sentences (Baddeley & Hitch, 1974; Huey, 1908/1968; Perfetti, 1985; Shankweiler & Crain, 1986).

One other feature of the reading process where phonological coding may aid comprehension is in the use of prosodic structures (Slowiaczek & Clifton, 1980). Those facets of spoken language, such as rhythm, intonation, and stress are important sources of information in understanding speech. In contrast, written language by its very nature provides fewer of these cues. Slowiaczek and Clifton speculate that phonological codes may provide a platform for obtaining such detail. Kosslyn and Matt (1977) have offered some evidence to show that this might be possible. In their study they showed that the rate at which a passage is read silently by a subject can be influenced by the speaking rate of the person said to have written it. Prior to reading a passage, subjects heard the voice of the supposed author. Where this voice spoke quickly, the subjects read more quickly and vice versa.

1.6 Use of phonological coding by deaf

Since profoundly deaf readers lack direct (ie auditory) access to spoken English, it might be thought that they would have considerable difficulty in acquiring and using a phonological code.

One early study which investigated this was by Chen (1976), using an approach utilised by Corcoran (1966) with hearing subjects. Corcoran had his subjects read passages of text and cancel out all instances of a particular letter. He found that they failed to detect silent letters (such as the letter *e* in the word *tape*) more often than letters which are pronounced (such as *e* in the word *red*). Corcoran interpreted this as evidence that hearing subjects were using phonological coding as they read. Chen (1976) used this technique to compare congenitally deaf, partially deaf and hearing subjects. He found that the hearing and partially hearing subjects omitted more silent letters but found no difference between silent and pronounced letters with the congenitally deaf group, suggesting that they are using a visual rather than a phonetic code.

Locke (1978) conducted a similar study to that of Chen with deaf children, using three target letters: c, g, and h. While hearing subjects were almost three times as likely to miss a phonemically nonmodal use of a letter (eg *g* in *rage*) than they were to miss one which was phonemically modal (eg *g* in *rag*). However there were no differences for the deaf children suggesting that the deaf do not effectively mediate print with speech.

However, it is known that performance in proofreading tasks such as those of Corcoran, Chen, and Locke is influenced by factors such as the position of the target letter within a word, word frequency, and letter positional frequency (Frith, 1979; Smith & Groat, 1979). The stimuli of Chen (1976) and Locke (1978) were not controlled for these factors. However, in a study which controlled the position the target within a word, Dodd (1987) found a difference in target letter detection by deaf children as a function of pronunciation. Similar differences in letter detection by deaf children as a function of pronunciation were found by Quinn (1981) when word frequency was controlled. However no proofreading study with deaf subjects has controlled for letter positional frequency in words, yet this orthographic structure variable is known to influence the reading of deaf subjects (Hanson, 1986).

Other studies have examined whether the deaf use forms of coding other than speech. For example, Locke and Locke (1971) compared deaf children who had intelligible speech and deaf children who had unintelligible speech with hearing children. Subjects had to recall pairs of printed letters which were similar to each other phonetically (B-C), visually (P-F), or dactylically (K-P). The hearing children were found to be using a phonetic code, the intelligible deaf group using all three coding systems, and the unintelligible deaf group relying on a visual-dactylic system. Odom, Blanton & McIntyre (1970), Moulton & Beasley (1975) and Bellugi, Klima & Siple (1974) have shown that large numbers of deaf subjects encode printed words manually in terms of signs.

Jarvella (1971) and Sachs (1974) have shown that auditory short-term memory is necessary for the temporary storage of surface material during the reading process. Some studies, eg Hanson (1982), Engle, Cantor & Turner (1989), Hanson & Lichtenstein (1990) have used short-term memory studies to show that prelingually, profoundly deaf subjects can acquire and use a phonological code. For example, some deaf subjects are sensitive to rhyme and perform less well in the recall of printed lists of rhyming words than non-rhyming words. In addition, it has been shown that the use of phonological coding in such short-term memory tasks is characteristic of good deaf readers but not poor deaf readers (Conrad, 1979; Hanson, Liberman & Shankweiler, 1984; Hanson & Lichtenstein, 1990). In Conrad's study, he found that speech recoding was being used by a small percentage of profoundly deaf children who were good readers, even those whose speech intelligibility was very poor. However, although these studies indicate that phonological processes underlie skilled reading, even in deaf populations, they do not establish that phonological coding is actually used by these subjects during the act of reading. Leybaert & Alegria (1993), using the Stroop paradigm, found interference for vocal and manual responses to coloured letter strings presented on a VDU for deaf and for hearing children. They interpret their results as indicating that access to orthographic, semantic and even phonological information can occur automatically when deaf children are presented with written words.

A recent study (Hanson, Goodell & Perfetti, 1991) used tongue-twister sentences with a concurrent memory load, following the procedures of McCutchen et al (1991), to ascertain whether deaf college students use a phonological code during silent reading. Results show that both groups of subjects make more errors on acceptability judgments when reading tongue-twister sentences than when reading control sentences, confirming the effects found by Haber & Haber (1982), McCutchen & Perfetti (1982) and McCutchen et al (1991) for hearing subjects but extending them to deaf subjects. As stated in the previous section, criticism has been levelled at using tongue-twister effects alone, as there is an inherent confounding effect because of the similarity of visual and phonetic components in such sentences. However, the Hanson, Goodell and Perfetti (1991) study also found specific phonetic interference between the tounge-twister sentences and the concurrent memory load numbers. Both groups had difficulty making acceptability judgements about alveolar fricative sentences when the memory load consisted of fricative numbers, and had difficulty with alveolar stop sentences when the load consisted of stop numbers. Since the written forms of the sentences and numbers do not overlap, the results cannot be attributed to graphemic similarity but rather must be because of phonetic similarity. McCutchen et al (1991) found, with hearing subjects, that the interference effect could show itself in poorer performance in either the acceptability judgments of the tounge-twisters or in poorer recall of the numerical data. In the Hanson, Goodell & Perfetti (1991) study, the effects were only noticeable in the acceptability judgments. The results seem to provide strong evidence that deaf subjects do indeed use a phonetic code when reading silently. It should be noted that the deaf subjects in this study tended to have poorly rated speech production skills, as judged by listeners. They also had reading scores on standardised reading tests which would be considered exceptionally high for deaf readers generally. This supports the findings of Conrad (1979) and Hanson & Lichtenstein (1990) who found that phonological coding tended to be restricted to those deaf subjects who were good readers.

A parallel study to that of Hanson, Goodell & Perfetti (1991) was carried out by Treiman & Hirsh-Pasek (1983), who used "finger-fumbler" sentences. These were sentences whose words had manual signs which were formationally very similar on the hands. Treiman & Hirsh-Pasek found that deaf subjects who were less proficient readers had greater difficulty reading finger-fumbler sentences than control sentences, suggesting that they may have been using a sign code to mediate comprehension. An interesting finding is that the best readers in this study did not show finger-fumbler effects and are likely to have been using a phonetic code.

These studies show that deaf subjects can and do use a variety of coding systems when reading, ranging from phonological through signing, dactylic and visual. However, it seems likely that those who are the best readers use phonological coding even although their speech production may be poor, with poorer readers relying more on manual coding systems.

1.7 Metalinguistic skills - hearing children

As readers become more skilled, they move from purely mechanical, unconsciously applied processing to a largely conscious, more critical stance. Brown (1980) considers the distinction between these two stages as defining the essential difference between cognitive and metacognitive aspects of performance. Metalinguistic skills are, broadly speaking, those higher cognitive processes which consciously reflect on the linguistic process itself. They relate to the planning, monitoring, self-questioning, and summarising skills used in critical reading and studying, and assume an increasing importance in middle school and beyond (from about age 9). Here the student is moving away from *learning to read* to *reading to learn*. Skills vital to effective study are (a) the ability to delete unnecessary, unimportant or redundant information, (b) the ability to substitute a superordinate term for lists of items (eg furniture for chair, table etc), (c) the ability to substitute superordinate actions for a list of subcomponents of that action, and (d) the summarisation of the essence of a text by choosing a topic sentence from the text or (e) the invention of a topic sentence if one is not provided. Brown & Day (1980) studied how well these rules were applied in a large group of subjects ranging in age from 10 years through to college graduates. They found that even the 10 year olds could delete inappropriate material effectively. Students became increasingly skilled with age in applying the more complex rules. The most difficult skill, inventing a topic sentence, was hardly used at all by the younger children and was used most by the most skilled, older readers.

Some studies, eg Brown, Campione, & Day (1981) and Day (1980), have shown that these skills can be taught to students.

1.8 Metalinguistic skills - deaf children

The study of the metalinguistic skills of deaf children has been rather neglected. Research has really been limited to studies of judgements of grammaticality (Kretschmer, 1976; Quigley et al, 1976). Kretschmer (1982) suggests that future work could examine areas such as what deaf children think reading is, or look at their study skills. Webster (1986) quotes Donaldson's contention that very young hearing children can "decentre" (contrary to Piagetian theory) if the task is meaningful to them. Their metalinguistic skills are therefore well developed by the time they start school, whereas the deaf child has many hurdles to overcome, eg confronting a written code for a different language, a different syntax, and discourse features. The explicit awareness of language as a code in its own right, which is required for reading, the ability for detached, disembedded thinking about language as a tool to be manipulated independent of its context, will be less well developed in hearing impaired children. Therefore the deaf child must rely on other information sources, but Webster contends that they are encouraged subtly and covertly by teachers and other adults to take a passive and receptive role rather than actively interact in learning situations.

In a recent study, Strassman (1992) interviewed deaf students to assess their metacognitive knowledge about school related reading. She found that subjects' schemata for this were largely skill based and passive, lacking mature metacognitive knowledge that would enable them to gain the most from their reading. Indeed, Erickson (1987) has suggested that deaf subjects require a collaborative "externally guided thinking" approach to help them read at an evaluative level rather than just literally.

1.9 Figurative language - hearing children

As children become more proficient in their use of language, they tend to use constructions which are non-literal. This figurative language can take many forms but the most common are metaphor, simile, and idiom. Figurative language seems pervasive in English, even in elementary reading textbooks (Dixon, Pearson & Ortony, 1980). However, comprehension of figurative language and ability to explain what particular phrases mean, develops in an incremental manner. Children are able to recognise correct paraphrases of figurative language first (Pollio & Pollio, 1979), and are then able to paraphrase figurative language independently (Cometa & Eson, 1978). At an even later stage they are able to explicate their metaphors (Billow, 1975). Arlin (1977) and Billow (1975) have shown that different metaphors are understood at different times, with representational metaphors (eg "he is a rock") being comprehended first, similarity metaphors next (eg "night was a blanket"), and proportional metaphors (eg "time was a thief, robbing her of life") being understood last (Giorcelli, 1982). Increasing exposure to metaphorical language improves comprehension of such language. Cometa & Eson (1978) and Pollio & Pollio (1979) have shown that it is easier for children to understand commonly used metaphors (eg "time flies") than less commonly used phrases (eg "curtain of hate").

1.10 Figurative Language - deaf children

The difficulties that the deaf experience in learning English are compounded by the profusion of figurative expressions in the language in both its written and spoken form. Dixon, Pearson, & Ortony (1980) have shown that figurative language is used in the very earliest books of commonly used reading series.

Conely (1976) studied the comprehension of idiomatic expressions in the deaf compared with hearing subjects and found no difference in performance for students matched for reading age 7.0 to 7.9 years but a significant difference for reading levels of 8.0 years and above. Scores on the test of idiomatic expressions were significantly and positively related to reading for both groups of subjects. This experiment has been criticised though, for example, for using syntactic structures which were too complex for the deaf.

Iran-Nejad, Ortony & Rittenhouse (1981) used metaphorical expressions controlled for vocabulary and syntax along the lines suggested by Quigley et al (1976). They found that the deaf subjects at all age levels (9 to 17 years) scored highly on literal comprehension but also obtained unexpectedly high scores on the metaphorical tasks. In a related study they showed that deaf children's comprehension of metaphors improved with practice and concluded that the deaf had no cognitive deficiency which prevented their comprehension of metaphorical language and their natural tendency to interpret literally can be counteracted by practice. This is one of the few studies to show positive results for the deaf. There were design faults, eg how well can figurative language be presented when vocabulary and syntax are controlled at very low levels, but at least it demonstrates that the difficulties deaf children have with figurative language are not caused by a cognitive deficit.

A study by Orlando and Shulman (1989) examined deaf and hearing children's comprehension of similes, metaphors, idioms, and proverbs in single sentences controlled for vocabulary but not for syntax. They found deaf subjects were poorer at understanding these constructions than the hearing children but that the deaf groups responses became more abstract with increasing age and with improved reading ability.

Several other studies are quoted by King & Quigley (1985) to support the fact that the deaf can cope with figurative and idiomatic language when other factors such as syntax are controlled eg Wilbur, Fraser & Fruchter (1981), Page (1981), and Houck (1982). These authors have speculated that deaf children can understand idioms if there is sufficient contextual information in the written material or that at least some of the idioms might be memorised or learned as a whole so that vocabulary and syntax present less of a problem.

Giorcelli (1982) studied various aspects of figurative language in the deaf including analogical and syllogistical reasoning, associative fluency, linguistic problem solving, interpretation of anomaly, and discrimination between paraphrases of novel and idiomatic metaphors. Choices of idiomatic phrases, syntax, and vocabulary were carefully controlled. Isolated and short and long contextual conditions were used. He found that the deaf scored significantly lower than the hearing controls on most of the subtests but that the performance of the hearing impaired improved with addition of context although this was still well below the hearing children's levels. The 18 year old deaf subjects did not perform as well as the 9 year old hearing subjects.

Payne (1982) studied the understanding of verb-particle combinations in deaf and hearing children. This structure is one of the most common ways in which English is expanded. These combinations can have, in some cases, a literal meaning (eg "run up a hill") as well as an idiomatic meaning (eg "run up a bill"). Children were assessed at three levels (idiomatic, semi-idiomatic, and literal), in five syntactic surface structures, with vocabulary controlled at the 6 and 7 year levels. Payne found that the hearing subjects scored significantly higher than the deaf on all levels of semantic difficulty and for all syntactic structures.

This seemingly contrary evidence, and the question of whether the difficulty for the deaf is how concepts are expressed in English or the underlying concepts themselves, is partly clarified by the work done on discourse.

1.11 Discourse features - hearing children

The process of reading can be approached from two theoretical standpoints. The first, called "bottom-up" theories (Gough, 1972), see the reader as viewing the words on the page, analysing these phonetically, accessing word meaning, and by synthesising these into sentences using the rules of grammar/syntax, building up meaning in an incremental manner. The process is text-driven. The skills reviewed in the previous sections are concerned with these bottom-up processes.

An alternative view of the reading process are the so-called "top-down" theories (Goodman, 1970). These see the reader using his/her world knowledge to help interpret what is written on the page. Cognitive structures which organise this world knowledge, variously called schemata (Rumelhart, 1975; Kintsch & van Dijk, 1978), scripts (Schank & Abelson, 1977), scenarios (Sanford & Garrod, 1981), are used to help the reader predict what can be expected from the text. In reality, it is generally accepted that a combination of bottom-up and top-down strategies are used during reading.

One particular approach to studying top-down processes is discourse analysis and story grammar. Discourse analysis is concerned with how text is constructed to aid comprehension. Brewer (1980) classified written discourse into three basic types: (a) descriptive discourse, (b) narrative discourse, and (c) expository discourse. Each of these would represent a different genre, eg descriptive - travel guide, narrative - detective story, expository - political speech.

Mandler and Johnson (1977) felt that stories could be analysed to reveal the *deep structure* which represented the true or fixed order of story events, and the *surface* *structure* which, for certain purposes of the author, might vary from the order of events in the deep structure. The deep structure, termed *story grammar*, is linked to the surface structure by a series of transformations.

Stein (1978) suggested six sections of a basic story grammar. These were (a) the setting, (b) an initiating event, (c) an emotional or cognitive response of the protagonist, (d) an attempt by the protagonist to achieve his/her goal, (e) a consequence, eg attainment or non-attainment of the goal, and (f) a reaction from the protagonist relating to the consequence. Stein (1978) found that children as young as six years of age make few errors in recalling the correct temporal order of stories that correspond to the expected story grammar sequence. Also, when the story structure deviated from the expected sequence as noted above, children recalled the story in an order similar to the story grammar rather than in the order of presentation.

Hansche & Gordon (1983), examining children's own writing, found that more categories of the story grammar appeared as the children got older. Six year olds only included setting, initiating event and consequence. By age nine, they had added consequence, and by thirteen all six categories were included.

1.12 Discourse Features - deaf children

The study of schema development, discourse features and top-down features in the reading skills of the deaf is less well documented (King & Quigley, 1985). Prinz & Prinz (1985) describe discourse development in sign language of a group of deaf children. Their findings indicate that the children were acquiring appropriate discourse strategies comparable to those used by hearing children in spoken conversation and deaf adults who sign.

Kluwin, Getson & Kluwin (1980) showed that deaf adolescents interpret ambiguous passages differently from hearing subjects, confirming that they bring their own world knowledge (schema) to the reading process. Gormley & Franzen (1978) have shown that deaf students comprehend syntactic structures more easily in connected discourse than in single sentences. Nolen & Wilbur (1985) also found that the use of context facilitates comprehension of more complex sentences for deaf children. McGill-Franzen & Gormley (1980) studied deaf children's comprehension of truncated passive sentences, eg "The wolf was killed", presented in context and in isolation. Hearing impaired children could not understand such sentences in isolation but could in the context of a familiar fairy tale, and so the authors concluded that the hearing impaired could understand passives in context. But as Robbins & Hatcher (1981) point out, the deaf children could answer any questions without even reading the text if the story was already familiar to them.

In a latter study Iseralite (1981) found that use of context did not aid correct interpretation of passive sentences. Edwoldt (1981) had deaf children read stories and interpret them into sign language, and videotaped their performance. Using miscue analysis, cloze procedures and retelling of stories by the children, she concluded that deaf children can read and retell stories by translating them into their own sign language. She states that they used extensively the semantic and syntactic cueing systems but did not did not over-rely on graphic information, thus supporting a top-down theory. But it is claimed from mainstream reading research (for example, Perfetti, 1995) that poor or beginning readers do use a top-down approach because they have not yet acquired automatic/ unconscious bottom-up skills. King & Quigley (op cit) do not necessarily accept these findings as Edwoldt gives only two examples and used only four deaf readers. They believe other interpretations can be made of the results.

Gaines, Mandler & Bryant (1981) used stories previously used by Mandler (1978) to compare deaf and hearing children on immediate and delayed recall of three stories, one of standard prose, one containing non-phonetic mis-spellings (eg "throgh" for "through"), and one which contained confused anaphoric references. They found that the deaf were able to comprehend the overall meaning of the stories and recall an amount of story propositions similar to hearing children, but the deaf did have significantly more semantic distortions than hearing subjects. But this would typically be expected from hearing readers who were beginning or poor readers. Gaines et al, therefore, concluded that the deaf were using a "broad reconstructive top-down schematic approach" to reading. The authors feel that deaf children often get the general idea of a passage using top-down strategies but have a lack of understanding or even a misunderstanding of important details. However, it should be borne in mind that the deaf children used were highly selected, being high-flyers at a particularly good oral school.

Sarachan-Deily (1984) compared recall of stories read by deaf and hearing adolescents. She found that the hearing students recalled significantly more propositions than the deaf students, but both groups recalled a similar number of story inferences. Within the deaf group, the better readers recalled more explicit information than the poorer readers but there was no difference in recall of implicit information.

Coggiola (1983) has shown that deaf college students are less able than their hearing peers at identifying the level of importance of idea units (or propositions) in a piece of text. Also the level of importance of a proposition significantly affected how well it could be recalled but this was less so for the deaf than the hearing group. Coggiola concludes that this lower level of sophistication in identifying and using the structural importance of information in text may, in part, contribute to the deaf's difficulties in reading comprehension. However she speculates that since they show some sensitivity in this area, the deaf could be trained in strategies to identify text structure and rate its importance, and that this could lead to better comprehension and retrieval.

Banks, Fraser, Fyfe, Grant, Gray, MacAuley, & Williams (1989) felt that deaf children do not schemata when reading as they are over-absorbed in bottom-up processes. They found that actively teaching story structures to these children, and teaching the concepts necessary to understand each passage, improved comprehension. They interpret this as demonstrating that the pre-existing schema for each story is being activated by this cueing but it is possible that the teaching of necessary concepts for each passage is helping to generate a schema where none previously existed.

A recent study by Marscark et al (1994) compared the signed and written structures of stories by deaf 7 to 15 year olds with written stories of hearing controls. They found that both signed and written output for the deaf had similar discourse structures as indicated by patterns of causal goal-action-outcome episodes. However, the grammatical and lexical character of the deaf students' work lagged behind that of the hearing children. This might suggest that any problems which are experienced by the deaf may not be due entirely to difficulties with reading alone, but may point to aspects of general language processing.

There is no clear consensus from the studies quoted above about the exact nature of the discourse skills of deaf children. However, it is evident that they can and do use broad top-down strategies when reading, although the exact nature of these is yet to be determined.

1.13 Anaphoric relationships: hearing children.

As a reader moves through a text, the general purpose is the accumulation of information. Each sentence should follow logically from the previous one and should add some new piece of information. It should be linked to some part of the previous sentence. One common way of achieving this is by using anaphora. These can be nouns, noun phrases, or pronouns. For example, a sentence referring to John Major may be followed by a reference to "the premier", "the prime minister", or simply "he". In other words, anaphora are the links which give a narrative cohesion.

Bormuth (1970) identified four major types of anaphor, viz:

(1) PRO WORDS	eg	Joe picked up the box He is a good hitter.

(2) DELETED MODIFIERS	eg	The boys practised hard. This [practice] helped.
(3) ELLIPSIS	eg	The stands were green. The fence was [green] too.
(4) SEMANTIC SUBSTITUTE	eg	Jim gathered the bats and balls. He put this equipment away.

Bormuth (1970) tested 240 nine year olds on their understanding of several of these anaphors embedded in paragraphs of 4 to 5 sentences, using multiple-choice responses. He found pro-verbs to be easier than deleted nouns, which in turn were easier than personal pronouns which were the most difficult of the anaphors studied.

However, this study was criticised by Lesgold (1974) who felt that too much emphasis was paid to syntax and little or no attention given to the semantic aspects. Correcting for these problems, he obtained different results from Bormuth. Lesgold found personal pronouns to be easiest with pro-verbs much more difficult. The design of his study was different. Oral answers were obtained to the questions and the number of semantically plausible potential answers in each passage was controlled. No information on this last aspect is available in the Bormuth study.

Chai (1967) has shown that hearing children's ability to identify the referent of a pronoun is a developmental skill, improving with age. Richek (1976-77) found that, of three anaphoric forms, nouns were easier than pronouns, which in turn were easier than null forms (where the noun is deleted). Barnitz (1980) studying 7 to 11 year olds, found that pronouns and noun phrases are significantly easier than clause or sentence referents, while forward pronominal reference is easier than backward pronominal reference. It was also easier if the pronoun and referent were in the same sentence rather than separate sentences.

It can be seen from the above studies that how hearing children use anaphora is not entirely clear. It seems though that use of anaphora is a developmental skill which improves as children get older. It also seems that repeated nouns or pronouns are easier to comprehend than null or substituted forms.

1.14 Anaphoric relationships: deaf children.

Stoefen (1981) noted one linguistic rule used frequently by the deaf which deviates from standard English, namely the "object-subject deletion rule". For example, the two sentences:

"*The boy kicked the cat*" and "*The boy ran away*"

can be conjoined to form the single sentence.....

"*The boy kicked the cat and ran away*".

Since the subject of the second sentence is coreferential to the subject of the first (ie refers to the same boy), conjunction reduction can apply, in other words the subject of the second sentence can be deleted. However for the two sentences:

"*The boy kicked the cat*" and "*The cat ran away*"

conjunction reduction cannot be applied when the subject of the second sentence is coreferential with the object of the first. Many of the deaf, however, do apply this conjunction reduction, the so-called "object-subject deletion rule", in their written output. Stoefen considered how deaf children who use this rule compared with deaf children who did not, in their understanding of the anaphoric reference when conjunction reduction takes place in semantically acceptable and unacceptable conditions. For example, the rule could apply in sentence (a) below but not in sentence (b), as it would be impossible for a tree to cry:

(a) *The boy sat by the girl and cried.*

(b) *The boy sat by the tree and cried.*

These null forms of anaphor were compared with other types, viz repetition of the noun, a personal pronoun, and a pro-verb type,eg

"*Joe eats ice-cream by the gallon*" "*Jim does (so) too*"

Stoefen found that the null forms proved to be the most difficult for both groups of deaf children, the rule users and non-users, but only in the semantically acceptable condition. There were no significant differences found between the other forms of anaphor. She also found no differences between the different genders of personal pronouns.

In summary, the deaf can use anaphoric referents successfully. However, they do have difficulty with one particular type of construction, the so-called "object-subject deletion rule", when the sentence is not semantically unambiguous.

1.15 Summary

It is widely documented that deaf children have great difficulty learning to read. The typical deaf school leaver has a reading age of about 8.5 to 9 years. They know fewer words than their hearing counterparts and know fewer second meanings for multiple-meaning words. However there is no difference in their ability to perceptually analyse text.

Phonological coding is seen as important for hearing readers as it allows them to retain in short-term memory the surface structure of material they have read, so that meaning can be built up across clauses and sentences. Deaf children, however, use a variety of coding systems, for example, sign, finger-spelling, or speech, but those who use a phonological code are usually the most successful readers.

The deaf have great difficulty with syntax, finding passive voices of verbs, relative clauses, and pronominalisation particularly problematic. This leads them to use rigid, stylised structures in their written output.

Deaf readers have a poor understanding of figurative and idiomatic language. They are less able to stand back and dissociate themselves from the reading process, and as a result tend to be literal in their interpretation of the material read.

There is some evidence to suggest that deaf children bring world knowledge to their interpretation of meaning and they seem to use similar discourse structures to hearing children. However, they are less likely to be able to use the structural importance of information in the text to aid comprehension.

1.16 Beyond the literal

The meanings of words, sentences, paragraphs, or entire passages are affected by the contexts in which they occur and by the prior knowledge which the reader has about the topic of the passage. As developing readers progress beyond anything but the most elementary of texts, they have to draw inferences beyond the literal meanings of the words on the page.

Since inferencing, as it relates to deaf readers, is the main topic of this thesis, it will be dealt with more fully in the next chapter. The inferencing skills of hearing children and deaf children will be examined first and this will be followed by a survey of studies relating to adult inferencing.

Chapter Two

Inferencing skills in deaf and hearing children

Anything beyond the most elementary of texts requires skills additional to those necessary for the literal understanding of the words on the page. For example, consider the pair of sentences below, quoted from Sanford and Garrod (1981):

Jill came bouncing down the stairs.
Harry rushed off to get the doctor.

Most readers see these two sentences as being linked and interpret them in terms of Jill falling on the stairs, injuring herself, and as a result of this Harry calling the doctor. With a slightly different second sentence, as below, a totally different interpretation is given to the same first sentence.

Jill came bouncing down the stairs.
Harry rushed over to kiss her.

This illustrates that readers go beyond the literal meaning of the words on the page. The meaning which is conveyed is dependent on the reader bringing additional world knowledge to the process in an attempt to produce a coherent interpretation of the text as a whole.

2.1 Inferencing ability in hearing children

Piaget (1969) and Fraisse (1963) thought that children at the pre-operational stage (ie younger than 7 years) would be unable to carry out the underlying operations necessary for inferencing. They believed, for example, that a child of this age could not attend to order and must, therefore, be unable to co-ordinate separate items of information to draw a logical conclusion and make transitive inferences. However Bryant & Trabasso (1971) have shown that transitive inferences about quantity are within the capabilities of children as young as 4 years of age if one controls for the limitations of their memory. Similarly Brown & Murphy (1975) have shown that 4 year olds can reproduce picture sequences as logical, holistic units when the context is meaningful. Also, Shatz (1975) has shown that in situations with the context clearly comprehensible to the child, inferential behaviour in speech acts has been observed in children as young as 19 months. Brown (1975) looked at recognition, reconstruction, and recall of arbitrary and logical sequences by children aged 5-8 years. He found non-verbal reconstruction was similar for pre-operational and older children but in verbal recall the pre-operational children were significantly poorer. Thus constraints on output may mis-represent the child's ability. Several studies have used memory for picture sequences to test the principles of assimilation theory. Assimilation theory (Bransford & Johnson, 1972) proposes that the listener or reader actively constructs an internal representation which is elaborated with successive, related sentential information. In both verbal recall (eg Paris & Mahoney, 1974) and non-verbal recall (eg Baggett, 1975) construction of inferences has been shown to be an implicit aspect of both visual and verbal processing.

Bell & Torrance (1986) have shown that the ability of children, aged 4 to 9 years, to make and recognise inferences shows marked improvement as they get older. It is not certain whether this improvement is the result of increasing exposure to written language.

2.2 Inferencing in deaf children

The major study of inferencing abilities of deaf children is that by Wilson (1979). He compared deaf and hearing subjects matched for reading age on six types of inference, presented in short stories which were associated with particular syntactic structures. Each story was followed immediately by sentences requiring a true/false response. One of these verification sentences related to the literal content of the story, the other to information which had to be inferred. Half of the stories were read by the subjects, the other half being (a) read orally to the hearing subjects, (b) simultaneously signed and read orally to the deaf subjects. Response times were measured during the reading task.

The six types of inference examined were:

Locative: eg.

The shirt is dirty.

The shirt is under the bed.

The cat is on the shirt.

Therefore, the cat is under the bed.

Predicate nominal story type allowing a transitive relation to be inferred based on noun phrase equivalence:

eg

The boy liked the movie.

The movie was Star Wars.

Therefore, the boy liked Star Wars.

Passive story type allowing the agent of an action to be inferred eg

The boy was kissed.

Mary liked the boy.

Therefore, Mary kissed the boy.

The instrumental story type allows an instrument to be inferred as part of an action

eg

The girl threw the ball.

The ball broke the window.

Therefore, The girl broke the window with the ball.

The negation story type allows a negative relation to be inferred eg

The girl did not have lunch.

Father ate soup.

Therefore, the girl did not have soup.

The dative story type allows an extension of an action to be inferred

eg

The girl showed the boy the letter.

Therefore, the boy read the letter.

Wilson found that there were minimal differences between the deaf and hearing on measures of literal comprehension but that the deaf were significantly poorer at inferential comprehension. (Davey, LaSasso, & McReady (1983) found similar results. Their study is described below in more detail). There were also no differences for the hearing group in response times for inferred and literal tasks but the deaf were significantly slower in judgements involving inference. The order of difficulty for each inference type for each group is shown in Table 2.1. It should be noted that there are only slight variations in the order of difficulty for both groups. Wilson also found a significant developmental trend in the acquisition of linguistic inference for the deaf but not for the hearing subjects on the material used. It seems that the deaf are going through stages of inferential development already achieved by their hearing peers who have been matched on reading ability. Wilson found that while the inferencing ability of the deaf was poorer when they tackle texts with more complex syntactic structure, they were also delayed in acquiring inferential skills even when they had a complete understanding of the syntax of the underlying sentences. It was also noted that the hearing subjects performed significantly better with the spoken presentation than with reading while the deaf performed significantly better with the reading presentation than with the signed presentation.

TABLE 2.1 Percentage of correct responses for each inference type (Wilson, 1979)

Hearing Group		Deaf Group	
Predicate Nominal	(89%)	Instrumental	(76%)
Instrumental	(87%)	Predicate Nominal	(73%)
Dative	(76%)	Dative	(61%)
Negative	(71%)	Negative	(50%)
Passive	(67%)	Locative	(47%)
Locative	(61%)	Passive	(40%)

This last conclusion is supported by findings from other studies which have compared how much material is assimilated by deaf subjects when different modes of communication are used. For example, White & Stevenson (1975) found reading to be superior to the spoken word and total communication (signing and speech) as a method of presenting factual information to deaf children. Quigley & Frisina (1961) also reported that reading appeared to provide a more stable means of communication for deaf students than does speech or fingerspelling. While Stuckless & Pollard (1977) showed that deaf children who were raised using fingerspelling could process written material more readily than fingerspelling.

A study by Pinhas (1984) also examined inferencing in deaf children, specifically looking at three questions: (a) Is inferential processing of written text more difficult for deaf than for hearing children? (b) Is there any difference in deaf children's ability to draw inferences if the information is presented in signed form compared with written form? (c) Are some types of inference easier to comprehend than others, specifically lexical compared with contextual inferences? However, most of the hypotheses were not proved and results were contradictory, due mainly to design faults. For example, response times were recorded by the experimenter manually starting a timing device and stopping it when the subjects gave their response. Obviously the variation in the reaction time of the experimenter could lead to gross inaccuracies.

Several studies, eg Wilson, Karchman & Jensema (1978); LaSasso (1980); Wolk & Schildroth (1984); Davey, LaSasso & McReady (1983) have highlighted the fact that task demands and test-taking skills appear to obscure the nature of inferencing processes in the deaf. For example, Davey et al compared reading comprehension of deaf and hearing subjects by examining cloze and free response productions and modified cloze and multiple choice recognition under look-back and no-look-back conditions. They found that hearing subjects had better comprehension overall than the deaf on multiple choice and free response. Literal questions were easier for all subjects than inferential questions. For cloze and modified cloze the hearing subjects were superior to the deaf. Also modified cloze was easier for both groups of subjects than cloze procedure. Being able to look back and re-interpret the text resulted in improved performance for both groups but more so for the hearing subjects. Thus not only may inferential processing be more difficult for the deaf, but certain test conditions and procedures may measure these skills more accurately than others.

2.3 Summary

In summary, it can be seen that the inferencing skills of deaf readers have not been extensively studied. The few studies which do exist offer inconsistent findings and have designs and methodologies which are flawed.

It is clear that deaf readers have difficulty in drawing inferences from textual material. However, is this true for all types of inference? Is it a general linguistic problem which is present when an attempt is made to draw an inference from information which is presented in sign language? Also, although deaf subjects may be generally poorer at inferring, on those occasions when they can successfully understand implicit material, what is the mechanism by which this is achieved? The series of studies reported in this thesis attempts to address these questions.

Before this is considered, the work on inferencing in hearing adults will be examined. Much work has been carried out in this area over the past fifteen years or so, particularly with reference to *what* inferences take place during reading and *when* during the reading process they are generated. Since a wide variety of approaches have been used in these studies, the associated methodologies will then be studied in detail and their applicability to the deaf population considered.

Chapter Three

Inferencing Skills in Hearing Adults

As many researchers such as Sanford and Garrod (1981) have argued, inferencing is vital to the reading process. The meanings of words, sentences, paragraphs or entire passages are affected by the contexts in which they occur and by the prior knowledge of the reader about the subject being read. The details presented in the text are interpreted by the reader in terms of his or her prior knowledge, the schemata which have been activated by the subject of the text, and by what the reader predicts is ahead in the text. Inferencing is essential if comprehension beyond the most basic level of very literal material is to be achieved, in other words, just about all reading material with a reading difficulty level of about 8 years, and probably much that is below that level.

3.1 Taxonomy of Inferences

One of the first steps would seem to be to define what is meant by an inference. Several attempts have been made at classifying the process.

Anderson (1981) suggested a four level, functional classification.

(a) The first level is lexically based, where the inference follows from the meaning of words. This depends on the reader's knowledge of the language and is relatively independent of context.

(b) The second level sees inferences being driven by the need to connect textual propositions. This depends on readers using their world knowledge as well as information contained in the text to infer the semantic and logical connections between otherwise unconnected propositions or sentences. An example might be:

John was working at the top of the stairs.

Jill was in the kitchen. She heard John scream.

She phoned for an ambulance.

The reader uses his or her world knowledge to infer that John must have fallen down the stairs and injured himself (the scream) and therefore Jill summoned help by phoning for an ambulance.

(c) The third level is involved when the reader's schemata or knowledge structures have been activated by the text and have unfilled slots (due to missing or unsupplied information) and therefore make the schemata tentative.

(d) The forth level is one step up from the third and involves the constant and repeated interaction of text and schemata to select and refine the schemata, thus providing a constantly updated interpretive framework for text comprehension.

These inferencing processes operate (as do lower order skills such as decoding) at an automatic and probably unconscious level. In fact these skills become such an integral part of the reading process that skilled readers often cannot distinguish, at recall, what they read from a page and what they added by way of inference (Brewer, 1975; Spiro, 1977).

An alternative taxonomy was suggested by Trabasso, Nicholas, Omanson, & Johnson (1977).

(a) The first stage is <u>lexically based</u> and is similar to that already mentioned above.

(b) The second category is labelled inferences of space and time. Here episodes have to be linked by the reader in a sort of spatio-temporal framework for understanding of the narrative text.
For example:

Jim went to Dundee yesterday.

He met his old teacher.

It is assumed, therefore, that Jim met his teacher in Dundee.

(c) The third category is called extrapolative inferences and is similar to the second level of Anderson's classification mentioned above. The reader extrapolates beyond the information mentioned in the text and fills in the gaps in what can be regarded as an "elliptical" narrative, thus providing the links between the statements actually appearing in the text.

(d) This category is called evaluative inferences. The reader uses the context of the narrative to evaluate the significance and effect of a particular statement.
For example:

Harry could only find one pound in his pocket.

(1) If this comes in the context of Harry having just eaten a large and expensive meal in a restaurant, then he is in trouble and the police may be called.

(2) If he is on his way to the bank, he will know that he needs to withdraw some cash.

3.2 Backward vs Forward Inferences.

A very important area of study has been exactly when an inference is made during the reading process. If the inference is drawn in order to establish coherence between the present piece of text and the preceding text, then it is a backward or bridging inference. On the other hand, if the inference is not needed for coherence, but is simply drawn to embellish the textual information, then it is a forward or elaborative inference.

3.2.1 Bridging Inferences.

Clark (1975) suggested that to facilitate communication, there is an implicit contract between speaker and listener which utilises a partitioning of the information conveyed into "Given" and "New" (see Grice, 1975). The "Given" is the assumed common ground between speaker and listener, while the "New" is exactly that, new information which the listener then tries to link to the previously "Given". This strategy will allow the listener or reader to understand anaphoric references. An example is given below:

Herb unpacked the picnic things.

The beer was warm.

The reader knows that "the beer" refers to one of "the picnic things" although this is not stated explicitly. In other words, it is necessary to made a bridging inference.

Haviland and Clark (1974) showed that it takes significantly longer to read a sentence such as "The beer was warm" when it follows a sentence which refers to "picnic things" as above than when it follows a sentence which mentions beer explicitly, as below:

Herb unpacked the beer.

The beer was warm.

In other words, comprehension time is increased whenever a reader has to make a bridging inference. However, in the above examples it could be argued that the advantage in reading times when there is a direct antecedent present is due merely to lexical priming.

Haviland & Clark (1974) varied their material, an example of which is shown below:

1 *(Tom wanted an alligator for Christmas.*

1' *(The alligator was his favourite present.*

2 *(Tom got an alligator for Christmas.*

2' *(The alligator was his favourite present.*

Chafe (1972) has shown that verbs like "want" in the first pair of sentences have the unusual feature of not presupposing the existence of their objects. You can want something which does not exist. Therefore, the phrase "an alligator" in sentence 1 does not necessarily set up a direct antecedent for the subsequent anaphoric reference in sentence 1'. Therefore it is necessary to make a bridging inference. Since a direct antecedent is set up in sentence 2, no bridging inference is necessary in sentence 2'. However, in both cases, the first sentence contains the word "alligator" and so any lexical priming will operate equally over both pairs of sentences.

Haviland & Clark again found a comprehension time advantage for the condition with the direct antecedent. Thus bridging inferences take additional processing time, and this occurs at the time of encountering the critical sentence.

Other authors have shown that such bridging inferences are affected by the relationship between the anaphor and its antecedent. In particular, either the anaphor or the antecedent can be a generalised term or class and the other can be a specific example of that class. For example, the antecedent could be "vehicle" while the anaphor might be "bus" or "car". These could be reversed with the antecedent being the generalised term and the anaphor the specific exemplar.

A specific example of a class membership mentioned above could either be a common exemplar or might be less common. For example, a car would be seen as being more common then, say, a tandem, for the class of vehicle. Garrod & Sanford (1977) investigated whether there was any difference in processing times for bridging inferences for exemplars which were common (ie sharing a high conjoint frequency with the category) compared with those that were more uncommon (sharing a low conjoint frequency). An example is given below:

4 A goose would sometimes wander into the house.

4' A robin would sometimes wander into the house.

5 The bird was attracted by the larder.

They found a reading time advantage for the common exemplar (sentence 4') over the less common (sentence 4). In other words, common exemplars allow bridging inferences to be made more quickly.

3.2.2 Elaborative Inferences.

As stated above, an elaborative inference is one which is not essential for comprehension at the time it is drawn, but rather expands on or further defines explicitly stated information in the text.

The studies which have been carried out in this area give equivocal results. For example, some studies have found that readers do not construct instrumental inferences, eg Corbett & Dosher (1978); Dosher & Corbett (1982); Singer (1979). An example from Singer is shown below:

The boy cleared the snow with a shovel.	(STATED)
The boy cleared the snow from the stairs.	(UNSTATED)
The shovel was heavy.	(TEST)

He found that subjects took longer to read the test sentence when it followed a sentence where the instrument was unstated than when it was explicitly stated. In other words, elaborative inferences are not made, only bridging inferences.

Other studies find that subjects do construct instrumental inferences, eg Garrod & Sanford (1981). An example used by them is shown below:

Keith was going to give a lecture in London.	
He was taking his car there overnight.	(STATED)
He was driving there overnight.	(UNSTATED)
The car had recently been overhauled.	
Did Keith go to London by car?	(TEST)

They found no advantage for the stated condition over the unstated and interpret this as showing that forward inferences are being made. A fuller description of the Garrod & Sanford study is given in a later chapter in which a version of that study is replicated and expanded.

A study by Cotter (1984) suggested that the different results obtained by two of the studies mentioned above (viz Singer and Garrod & Sanford) were due to each study using different subsets of verbs, the examples used in the Garrod & Sanford study having stronger verb-instrument links than did the Singer study. These differences in strength of semantic relatedness could not be detected by the pretest questionnaires used but became obvious during the reading task because of the greater time pressures therein.

A series of experiments by O'Brien, Shank, Myers, & Rayner (1988) showed that subjects do generate and store elaborative inferences on-line but only under fairly limited conditions. Readers may draw elaborative inferences only when the context is sufficiently rich and therefore constraining. Without this constraint, the reader is likely to delay any inferential process until the text indicates unambiguously what the correct inference is or until it becomes necessary for comprehension. Once generated, elaborative inferences are stored as part of the long-term memory of a passage. An slightly simplified example from O'Brien et al (1988) illustrates this:

Sentence (1) = High Context, Sentence (2) = Low Context

Gus loved to play games that allowed him to spend time outside.

(1) *His dream of a hole-in-one made this particular sport his favourite.*

(2) *The leisurely pace made this particular sport his favourite.*

He used every chance he could to practise.

Playing golf was something he never got tired of doing.

In the high context condition, readers infer that the game is golf and take less time when fixating the word "*golf*" in the last sentence than they do in the low context condition.

3.3 Causal Structure of a Text.

Several researchers have thought that the chain of causal relatedness of a narrative is essential to the comprehension process, eg Graesser & Clark (1985) and Trabasso, van den Broek, & Suh (1989). The inferences drawn link together the various propositions stated in the text in a meaningful way. However, if the immediately preceding text fails to provide a cause for an action contained in the sentence just read (a "causal break"), the reader reviews earlier segments of the text in memory and, if necessary, uses world knowledge to generate a bridging inference between the current sentence and sentences read earlier. This idea is supported by several studies which show that reading time increases with decreased causal coherence, eg Haberlandt & Bingham (1978), Keenan, Baillet & Brown (1984), Myers, Shinjo & Duffy (1987), and O'Brien & Myers (1985). The variations in reading times are also accompanied by variation in memory for the text. When it is difficult to generate a causal inference memory is poor. However, when a causal inference is easily drawn in response to a coherence break in the text, text memory is actually better than in those cases where there is no coherence break (Keenan, Baillet, & Brown, 1984; Myers & Duffy, 1991; Myers, Shinjo, & Duffy, 1987; O'Brien and Myers, 1985).

Two models of comprehension based around the causal structure of a text are considered below. One views the reader as being relatively passive, with causal linkages being sought and constructed only in response to coherence breaks. The other considers the reader to be much more active, constructing a richly interconnected network of causal links between the various elements of the story. Once these models have been reviewed, a number of more recent studies are considered, each refining methodology and design in an attempt to build support for one or other of the theoretical stances.

3.3.1 Trabasso's Model of Causal Structure.

Trabasso's model (Trabasso & van den Broek, 1985; Trabasso & Sperry, 1985) assumes that a text's causal structure is parsed into statements of events which correspond roughly to clauses. All pairs of statements are considered and event A is said to cause another, event B, if in the context of the story B would not have occurred if A had not occurred. These events are seen as nodes which are connected to one another in a chain-like structure.

An example from Trabasso and Sperry (1985) which illustrates this is noted below:

> *A father and his son / were taking their donkey to town / to sell him / at the marketplace./ They had not gone a great distance,/ when they met a group of pretty maidens / who were returning from the town./ .... / The donkey, not liking to be tied / kicked so ferociously / that he broke the rope,/ tumbled off the pole into the water / and scrambled away into the thicket.*

Some of the causal connections are of a physical nature, eg "*kicked so ferociously*" being a cause of "*that he broke the rope.*" Others are less obvious. For example, "*to sell him*" is a cause of "*were taking the donkey to town*" as they have to go to town as that is were the market is. In addition, "*They had not gone a great distance*" can be seen as a cause of "*they met a group of pretty maidens*" since they presumably would not have met them if they had not left home.

Any such causal structure can be represented diagramatically. An example of how a text may be represented is shown in Figure 3.1 below:

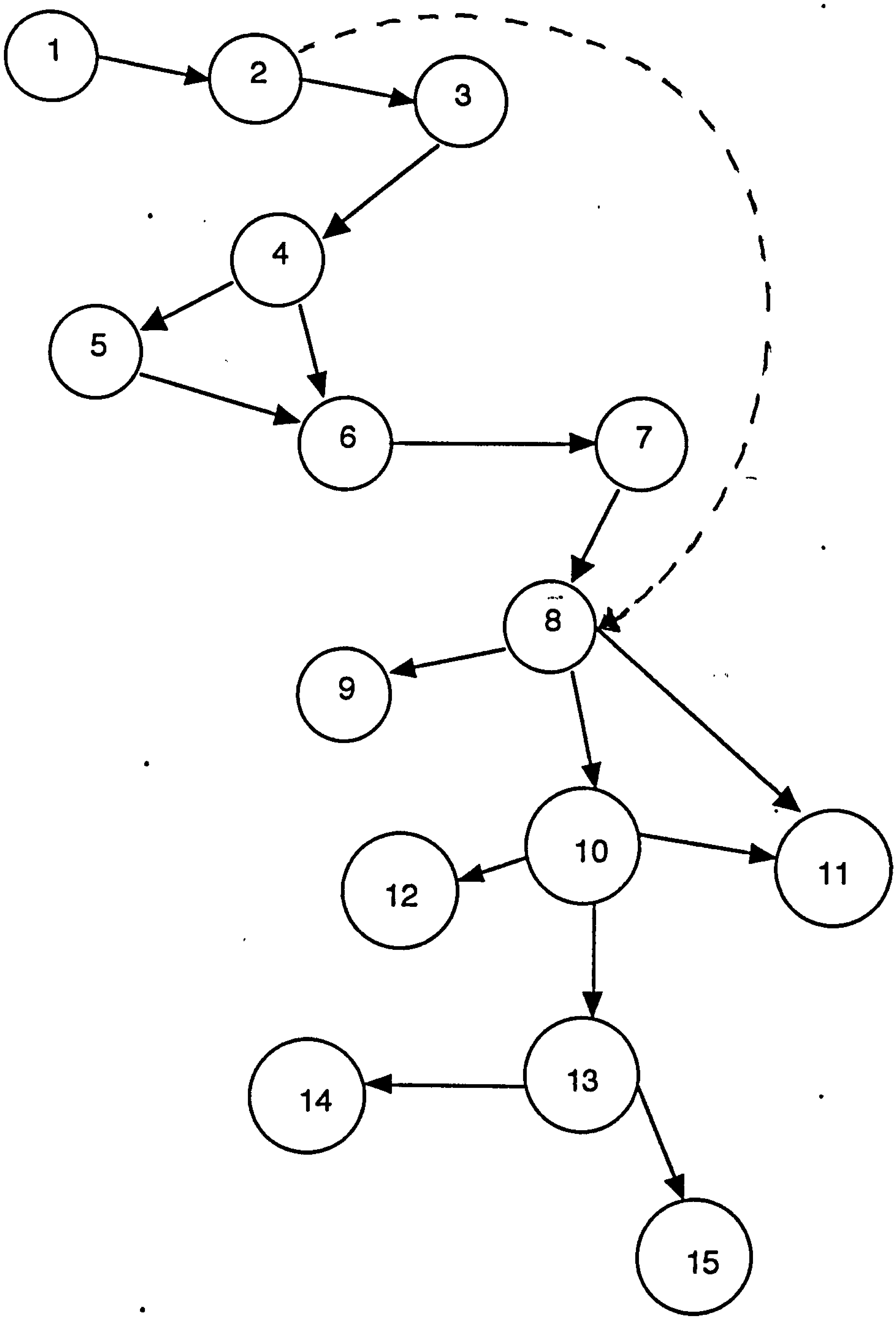


Figure 3.1: An illustration of Trabasso's causal structure for a text

Each statement is represented as a node which is linked to other nodes by a series of paths, the causal connections. Some nodes have more connections than others. For example, node 6 has three connections while node 8 has five.

Some statements can be on the main causal chain while others will be on a side chain, the main causal chain being the link of causal connections from the start of an episode (eg the introduction of the main characters) through to the attaining or otherwise of the goals of these characters. Nodes which are not on the main causal chain in the example above are 9, 11, 12, and 14.

It has been shown that recall of a statement increases with the number of connections to that node or statement (Trabasso & van den Broek, 1985). These authors also showed that nodes which are on the main causal chain linking the opening and closing statements of a narrative are better remembered than those which are on a side chain. However, it could be argued that these effects, rather than reflecting comprehension processes operating at the time of reading, actually are a function of the reconstructive processes operating at a later time when recall is required. This is discussed in more detail in a later section below.

O'Brien & Myers (1987) analysed the effects of causal structure upon reading times and recall times obtained immediately after reading. They found that times taken to read a statement containing an anaphoric reference to an antecedent in the text decreased as the number of causal connections to the statement containing the anaphor increased. In a subsequent study, they measured the time to produce an antecedent in answer to a question immediately after reading the text, eg "What was the stolen object?" (The anaphor was dropped from the text in this version).. They found that recall times were a function of the length of the shortest causal path from the setting information in the passage to the statement containing the targeted word.

Therefore, the patterns of causal connections clearly influence retrieval times. However the critical aspect of the causal structure depends on the nature of the retrieval task. It is likely that reinstatement by use of an anaphor and recall in response to a question start at different points. It can be assumed that reinstatement searches initiated by an anaphoric reference start from nodes corresponding to propositions in a buffer memory (Kintsch & van Dijk, 1978). Since these propositions come from different parts of the passage, then contact during such a search is more likely with a targeted proposition which lies on many different causal paths, that is one with many connections. On the other hand, recall in response to a question involves a search of long-term memory. This involves a spread of activation along causal links, beginning at the first proposition encountered. In such a search, the shortest path dominates and therefore path length is critical.

While these findings would seem to support the Trabasso model, Myers(1990) has reservations based on the experimental design. He feels that since many of the results rest upon regression analysis they are sensitive to, among others, (a) correlations among the predictors, (b) the sample sizes relative to the number of predictors. Another model discussed next, would also seem capable of predicting the results found above.

3.3.2 Fletcher and Blooms Current State Model.

Fletcher & Bloom (1988) proposed a different model which saw the reader as a less active processor. Each sentence is partitioned into statements using two rules. Firstly, a sentence boundary always terminates a statement. Secondly, clause boundaries terminate a statement only if it separates segments having different causal antecedents or consequences. The Fletcher & Bloom model assumes that fewer propositions are maintained in an active state and less frequent searches of long-term memory are needed, compared to the Trabasso model.

The Fletcher & Bloom model is very similar to that proposed by Kintsch & van Dijk (1978). The latter assumes that some propositions from one "cycle" are held over to the next cycle. New propositions are then connected to these older propositions and a series of links built up. Anything new which is not able to be connected remains in the buffer. If nothing can can be matched, a search can then be made of long-term memory. The longer propositions reside in short-term memory, the stronger the record made in long-term memory. Useful propositions held over are those which are richly connected and, therefore, likely to be of importance (ie the theme or gist). Others likely to be held over are those most recently mentioned, as these are more likely to connect or relate to to subsequently mentioned propositions than to distant ones.

The main difference in the Fletcher & Bloom model is the assumption that the most recent statement or proposition which is connected to an antecedent but not to a consequence is held in a short-term working buffer. This can be illustrated by an extract from a text used in a study by Bloom, Fletcher, van den Broek, Ritz, & Shapiro (1990):

(1) *Danny wanted to have the red bike /*

(2) *that he saw in the window of the neighbourhood bike shop./*

(3) *Danny knew that first he had to have $50 /*

(4) *to buy the bike. /* (5) .....

When processing the first sentence, a link is formed by the reader from Statement 2 to Statement 1; seeing the bike causes Danny to want it. Statement 1 is held in the buffer as the second sentence is processed and is causally linked to Statement 4; wanting to have the bike is causally related to wanting to buy it.

Fletcher & Bloom (1988) found that propositional recall correlated significantly with the number of cycles in the buffer and the number of causal connections derived from the current-state strategy. Their model accounted for more of the variance in propositional recall than either the Kintsch & van Dijk model or a model akin to that of Trabasso.

Bloom, Fletcher, van den Broek, Reitz & Shapiro (1990) provided evidence from reading time experiments which seems to support the current-state model more than the Trabasso model. They predicted that: (a) time is required to connect statements; (b) causal connections between statements in the same sentence (sentence links) will be more rapidly made than connections between a new statement and one currently in working memory (working memory connections); and (c) when no causal antecedent is available, a search of long-term memory is required.

They also distinguished between allowable causal links (the essential links of the Fletcher & Bloom model) and potential links (the additional links in the Trabasso model). Normally a potential link will not be accompanied by increased reading time except (a) for coherence breaks, (b) when a statement satisfies a goal no longer in short-term memory, and (c) when a statement frustrates a goal no longer in short-term memory.

They found that potential links have no effect on reading times except for (a) and (b) above; (c) had no effect. They also found that both sentence links and working memory links do require time to construct.

In contrast with the Trabasso model, in which all potential links are encoded, the current-state model's assumptions about the processor seem more consistent with the response time data and suggest that subjects draw inferences only when required to maintain text coherence or when instructed to do so (Duffy, 1986; O'Brien, Shank, Myers & Rayner, 1988; Potts, Keenan & Golding, 1988; Singer & Ferreira, 1983).

There are some problems associated with the current-state model however. For example, in the Fletcher & Bloom model it is supposed that a statement is held in the buffer for another processing cycle only if it has an antecedent cause but no consequence in the previously read text. However a series of experiments by Duffy (1986) suggest that events which lack a causal antecedent in the preceding text are very likely to be held in working memory. Consider the example from Duffy below:

John was eating in the dining car of a train. The waiter brought him a bowl of soup. Suddenly the train screeched to a halt. The soup spilled in John's lap.

The sentence, *"Suddenly the train screeched to a halt"* , has no obvious causal antecedent and, as a result, tends to be "highlighted" by the readers important.

These occurrences can be viewed as script deviations, where something unexpected and unexplained happens in the narrative. The reader naturally focuses on it and searches the following text for a cause or consequence to which it can be connected. This would seem to be a necessary addendum to the current-state strategy.

Another problem relates to the assumption that reading a goal-satisfying statement initiates a search of long-term memory. An alternative explanation is possible. The goal is likely to be stated early on in the narrative, while the goal-satisfying statement is near the end. In other words, the two statements are likely to be separated by several intervening sentences and the initial goal is thus considerably backgrounded. The increase in reading time may merely reflect a search to find the antecedent to the anaphor.

In an effort to address some of these problems, Dopkins & Myers (1989) attempted to discover what exactly was held in working memory at each point in the text and what connections were established in long-term memory. They felt that goals are well defined causes of subsequent events and, as such, are central to the theories under consideration. From both Trabasso's and Fletcher & Bloom's models, it can be predicted that a concept will be more accessible in long-term memory when it is part of a statement which defines a goal than when it is not, since such a goal statement will have more connections and be on the main causal chain. An example of the type of passage used by Dopkins & Myers is shown below:

Goal Version

Al and Dave were in the process of breaking out of the state pen. With snarls on their faces, they raced down a corridor. They looked for a hostage to improve their bargaining position. When the warden appeared, Al crept up quickly behind him.

Control Version

Al and Dave were in the process of breaking out of the state pen. With snarls on their faces, they raced down the corridor. They grabbed a hostage to improve their bargaining position. When the warden appeared, Al told him to keep back.

The goal version contained a goal setting sentence (*They looked for a hostage...*) while in the control version, the goal has been met (*They grabbed a hostage...*). To test for the accessibility of the goal word (*hostage*), an item-recognition task was used. Since in the goal version, the goal setting statement should be linked to the goal satisfaction statement, then the goal word should be more accessible and the target word more quickly accessed. Target nouns were categories whose membership is not well defined (anyone can be a potential hostage). This helps reduce the lexical association between the goal word (*hostage*) and its subsequent instantiation (*warden*). Anaphoric reference to the target word in the last sentence is also avoided, and both versions have as similar wording as possible. Dopkins & Myers found that a word is more accessible in long-term memory when it is part of a goal setting statement than it otherwise would be (eg when the goal has been met early in the text).

It might be thought that this finding could be as a result of the reader paying more attention to the goal setting sentence. However, Dopkins & Myers proposed that this is not due to increased attention but requires the presence of a goal satisfaction sentence. They see a connection being established between a goal word and its later instantiation, but only in the condition in which a goal has been set.

One interpretation of the connection between the goal word and its instantiation might be that the reader attempts to actively anticipate subsequent text events, but this seems unlikely in the face of the many studies that have failed to demonstrate forward inferencing (eg Duffy, 1986; Potts, Keenan, & Golding, 1988).

A more likely mechanism is that proposed by Garrod & Sanford (1977) that reference to an object or individual is checked against locations in memory in order to store the new information. A variation on their original study by Dopkins & Myers (1989) supported this contention. The search for an anaphoric antecedent is a first response and if the resulting representation is coherent, then no further connections are made. The results seem to support the position that goal-satisfying statements are linked to statements initially setting the goal (eg Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990; Fletcher & Bloom, 1988; Trabasso & van den Broek, 1985).

However, Myers (1990) believes that it is not a case of the reader actively seeking causal relations between current input and previously read propositions. Nor does the connection between goal-setting statements and goal-satisfying statements seem to occur as generally as suggested by Bloom et al (1990). The process occurs if there is an anaphoric connection between the two statements or if there is no better place to locate concepts introduced in the goal-satisfaction statement. The reader's purpose may not be to establish a richly interconnected representation but to maintain what McKoon & Ratcliff (1986) have labelled minimal local coherence.

One problem with the studies cited above is that they say little about which links are constructed, and at what point during the reading process this occurs. The difficulty arises from the methods used. The majority of studies of causal comprehension have used indirect measures of the comprehension process - recall scores, importance ratings, frequency of inclusion in summaries. However, as will be discussed in more detail in the next chapter, these methods are susceptible to factors which occur after the text has been read. What is required is an approach which utilises on-line measures of comprehension. This is now considered in the work of Sanford and Garrod.

3.4 Sanford and Garrod's Model of Reading

Sanford and Garrod (1981) put forward a model of reading comprehension which emphasises top-down processing. They state that in order to understand a piece of discourse, it is necessary for the reader to relate it to some hypothetical or real state of the world, in other words, to determine to what it is that the discourse refers. They view this process in terms of the dynamics of memory access.

Memory is viewed as being divided into four partitions, each of which is independently addressable by the reader. These partitions limit the domain of any search which is needed to find the relevant information for understanding of what has been read. Two of these partitions are seen as being dynamic in nature since the contents which they hold change as the text unfolds. This Sanford and Garrod call current focus. That partition of focus where the memory representation derives purely from interpretation of the text, called explicit focus, is of limited capacity. The second partition, called implicit focus, comprises representations derived from scenario-based sources. Explicit and implicit focus provide a retrieval domain which incorporates the information most pertinent to understanding the text at any given time. In some sense, they embody the current topic of the text.

The two remaining partitions of memory are viewed as being more static and less susceptible to change over time. One of these is the long-term memory of the text, the other is long-term semantic memory, or the knowledge base of the reader.

Sanford and Garrod discuss their model mainly from the viewpoint of resolving references. After the initial parsing process, for any entity mentioned in the text, an attempt will be made to map it onto a representation of its meaning (a token) in explicit focus. At the start of a new passage, no tokens will exist in short-term memory as yet, so a new token is set up for the entity mentioned. Any subsequent reference to this entity, say by use of a pronoun, will lead to the pronoun being mapped directly onto the token in explicit focus. This mapping process is called *retrieval* by Sanford and Garrod.

A second process which is involved is *construction.* No match may exist in explicit focus and reference resolution may result from a search of implicit focus. There will be no token corresponding to the partial description (say a pronoun) but there may be a slot corresponding to it in the scenario. A token is then constructed in explicit focus which maps directly onto the scenario slot in implicit focus.

For example, consider the sentences below:

> *Keith drove to London.*
>
> *The car broke down halfway.*

Since "*Keith*" is a new entity, a token is created in explicit focus. When "*drove to London*" is encountered, a search of memory occurs for "*drove*". Since this too is a new entity, an appropriate scenario is identified in long-term memory and is moved into implicit focus. The subsequent referral to "*car*" can be mapped directly onto the appropriate slot in the scenario.

This search, mapping and constructing within explicit and implicit focus is described as primary processing. If this primary level processing fails to select a unique referent then secondary processing will take place. It may be that the entity referred to has not been mentioned recently and has therefore passed out of current focus into the long-term memory representation of the text. A search of this area of memory should find an appropriate match. Alternatively, if the entity has not been mentioned before and no slots can be found within the current scenario, then the scenario may be modified or a new, more appropriate scenario chosen. Any mapping with the new scenario will be at the primary level. Secondary processing is much slower than primary processing.

Characters or aspects of the scenario which are highlighted in explicit focus are said to be foregrounded. Qualifying an entity by additional information serves to bring it more into the foreground. Thus in a restaurant scenario, an "attractive waiter" will be more highly foregrounded than the simple mention of "a waiter". Highly predicted characters (eg waiter in restaurant scenario above) will be less foregrounded than poorly predicted characters (for example, a dishwasher), as they are more readily accommodated by the scenario. It is only the poorly predicted entities which are represented in explicit focus.

Not only are pronominal or noun phrase references resolved but also entire character-action complexes. If such a whole event cannot be assimilated into a scenario, then it should be foregrounded, that is highlighted, and receive a fairly strong representation in explicit focus. Thus, while explicit focus will primarily represent such foregrounded information not already represented in implicit focus, it will always attempt to map this into a scenario in implicit focus. It is when this cannot be done, that the scenario will be modified or changed.

It should be mentioned that the particular scenarios evoked will be dependent on the background, experience and inclination of the reader. Also the contents of similar scenarios are likely to vary between different readers. Garrod and Sanford (1980) have shown that if a title is used to introduce a piece of text, then this serves to bring in the appropriate scenario. This is illustrated by the sample text from their study noted below:

<u>Appropriate Scenario</u>

IN COURT (TITLE)

Fred was being questioned (by a lawyer).

He had been accused of murder.

<u>*The lawyer*</u> *was trying to prove his innocence.* (TARGET)

<u>Inappropriate Scenario</u>

TELLING A LIE (TITLE)

Fred was being questioned (by a lawyer).

He couldn't tell the truth.

<u>*The lawyer*</u> *was trying to prove his innocence.* (TARGET)

The text can have an appropriate or inappropriate title and the target entity (*lawyer*) can either be stated or unstated in the first sentence. Garrod and Sanford found no difference in reading times for the target sentences between stated and unstated conditions when an appropriate title is used. However, when the title is inappropriate, the target in the unstated condition takes significantly longer to read. Thus an informative title will invoke the appropriate scenario.

Anderson (1982) empirically investigated how time shifts signalled in the text cause changes in scenarios and the relative availability of both topic characters who are essential to the narrative and scenario dependent characters. If a text contains a time shift which is outwith the normally expected range of a scenario (eg "five hours later" when talking about a child's birthday party), then it is much more difficult to resolve a subsequent reference to a character who is bound in with that scenario (eg an entertainer for the party scenario above). There is no difference to the status of the key topic character. Also the actual sentence which signals the time shift is read more slowly, reflecting the effort needed to switch off the current scenario and either set up a new scenario or prepare for possible bridging operations.

Thus Sanford and Garrod's model attempts to describe a dynamic framework which utilises the background knowledge of the reader and how this links in with partitions of memory to help interpret narrative text.

3.5 McKoon and Ratcliff: The Minimalist Hypothesis

In contrast, McKoon & Ratcliff (1992) in a lengthy paper attempted to separate which inferences are automatically produced and which are the result of slower, goal-based strategic processes. From such a separation, they state their hope of laying the foundation for the development of a greater understanding of the database which is provided within the first few hundred milliseconds of processing. This knowledge could be used to determine what information strategic processes have to work on, which strategic inferences are likely to be easy and which difficult.

They do this in what they call the *minimalist hypothesis*. The minimalist standpoint assumes that:

(1) Readers do not automatically construct inferences to fully represent the situation described by a text.

(2) In the absence of specific, goal directed strategies, only two types of inference are constructed automatically:

(a) those necessary to establish local text coherence,

(b) those from information which is readily available from the text or from general knowledge.

Local coherence is defined as those propositions of a text that are in working memory at the same time, ie no further apart in the text than one or two sentences. McKoon & Ratcliff cite the Kintch & van Dijk (1978) model where propositions in short-term memory are connected together through overlap of their arguments and are ordered according to the most salient or topical proposition. One example of inference that is needed to establish local coherence is the connection between an anaphor and its referent.

Global inferences connect widely separated pieces of textual information into an overall causal chain or network. This information is not available in short-term memory and has to be retrieved from long-term memory.

McKoon & Ratcliff contrast the constructionist hypothesis with the minimalist hypothesis. They do this by considering previous studies on elaborative inferences such as semantic, instrumental and predictive inferences as these are not necessary for local coherence. They interpret the data for these studies (eg Singer, 1978, 1979; Corbett & Dosher, 1978; Dosher & Corbett, 1982) as contradicting the constructionist standpoint and suggest the minimalist hypothesis as an alternative explanation.

Another series of studies considered by McKoon & Ratcliff are "real life" situations, which are again interpreted as supporting the minimalist view. Several of these real life narratives (eg Morrow, Greenspan, & Bower, 1987; Morrow, Bower, & Greenspan, 1989) involve subjects being taught about a situation prior to reading the text. For example, subjects were taught with the aid of a map about the rooms in a laboratory and the layout of objects within those rooms. They then read a series of narratives describing a character moving through the rooms. Results showed that subjects were faster to answer questions about the location of objects when they were relevant to the character's current location or the character's goal location. One criticism of these studies is that the objects were made available not as a result of automatic (priming) processes but as a result of strategic retrieval processes to rehearse the position of objects within the rooms. This explanation is rejected by McKoon and Ratcliff as they believe it is not possible to predict what aspects of a situation are salient in any given circumstances and therefore no way of predicting which inferences should be included in the mental model. The results are considered by them, therefore, as being consistent with the minimalist standpoint.

They then discuss the results from five of their own studies which show that global inferences are not automatically encoded. For example, in one study they used textual materials which were locally coherent, consisting of two paragraphs, an introduction followed by a continuation. In the introduction, a general goal is set (eg *to kill the president*) and a goal which is subordinate to this goal is also described (eg *using a rifle*). For the continuation paragraph, there are three conditions: Control, Try Again, and Substitution. In the Control version, both goals are achieved (*the president is shot*). In the Try Again condition, a problem arises in attaining the subordinate goal (eg *the telescope falls off the rifle*) and the character tries the goal again by using the rifle in a different way (eg *sighting without the scope*). In the Substitution condition, a problem also arises with the subordinate goal, but instead of trying again, the character substitutes a new subordinate goal (eg *using hand grenades instead of the rifle*).

Immediately after reading the text, a recognition test word for the general goal (*kill*) is presented. According to the minimalist hypothesis, the general goal at this stage should no longer be used for comprehension under any of the three conditions and there should be no differences in response times to the test probe. In contrast, according to a constructionist approach, response times to the probe should be faster for the Try Again and Substitution conditions relative to the Control condition. This is because the character is still trying to achieve the general goal at the end for the first two conditions. In the Control condition, the main goal has been achieved and a new general goal takes over.

Indeed, the results support these predictions and McKoon and Ratcliff cite this as evidence endorsing the minimalist standpoint. However the design used relies on the null hypothesis, ie the finding of no differences to prove their thesis rather than finding positive evidence.

McKoon & Ratcliff feel that there is no evidence to directly support the constructionist viewpoint and in the absence of such evidence, there is no reason to reject a simpler hypothesis (the minimalist view) and this should be used until such evidence can be accrued.

3.6 Global Causal Inferences: Trabasso & Suh

Trabasso and Suh (1993) re-examined the question of whether global, causal inferences are made during comprehension. They used a three pronged approach, proposed by Magliano & Graesser (1991), which involved:

(a) identifying potential inferences a priori in the texts used
(b) using a "talk aloud" procedure to evaluate if these predicted inferences are actually made
(c) using converging methods which are less intrusive eg on-line.reading times

Trabasso & Suh consider goals to be very important in the study of causal inferences as they motivate other goals and actions, and organise complex structures into episodes. For any motivational inference the question has to be asked whether the goal is fulfilled or not. The interpretation of any action as an attempt to fulfil a goal depends on whether or not the goal is satisfied prior to the time of action. If it has, this would lead to the generation of a new goal/plan inference.

They used textual materials which include motivational inferences of a global nature which occur over large surface distances in the text, and where locally coherent inferences can be easily made for each sentence, for example because of anaphoric references. Contrary to the claims of McKoon & Ratcliff (1992) in their minimalist hypothesis, Trabasso & Suh found that motivational inferences of a global nature do occur. This demonstrates the advantage of the "talk aloud" procedure in that it is possible to monitor exactly what inference (if any) the subject is making at each stage.

The "talk aloud" procedure also revealed that, in addition to motivational processes, subjects also use operations to (a) maintain, (b) retrieve, (c) elaborate and (d) explain information in the focal sentence, with these mental operations occurring in working memory.

The evidence obtained from these "talk aloud" procedures identifies that those inferences predicted for the textual materials do, in fact, occur. This is supported by on-line measures such as reading times.

In order to integrate the current sentence information with that from prior text, the current sentence has to be *maintained* while the prior sentence is *retrieved* and used to *explain* the current sentence in working memory. These operations will facilitate a story's coherence and long-term memorability. However, if subjects tried to *elaborate* and add information that moved away from the sentences, they were less likely to integrate the sentence and forgot more information.

If a sentence can be readily integrated into what is know, then processing time should be reduced; if not search and construction is required and processing is slowed down. How connected a sentence is to other sentences in the discourse analysis and how connected it is in the talk aloud protocols are related conceptually and empirically. A sentence that is highly connected is shorter in its distance in the network to other sentences. The operations of retrieving, maintaining and explaining create connections and reduce distance. This aids assimilation which is reflected in a faster reading time. Elaborations of consequences or associations reduce intersentence connections and increase distance. This hinders easy assimilation with a correspondingly slower reading time.

Talk aloud protocols have been used in previous studies. Fletcher (1986) and Olson, Duffy, & Mack (1984) used such an approach to investigate what information was available to working memory during the comprehension of sentences. Subjects were given explicit instructions to comment on what connections there were with previous text, and to make predictions about what might occur. As this approach is open to the criticism that subjects are being directed towards cognitive strategies which might not occur otherwise, Trabasso & Suh generated a less directed protocol, with less specific instructions than those of Olson et al (1984) and Fletcher (1986). Subjects were given more freedom of expression and were only asked to comment on their understanding of each sentence and talk about whatever came into their minds.

The verbal protocol analysis method is criticised by Nisbett & Wilson (1977) and Trabasso & Suh admit that the thinking aloud that occurs in the verbal protocols, does not give direct access to the on-line thought processes that occur. However they believe they are correlated with the underlying thought processes and reflect the operations that have been performed during comprehension.

Trabasso & Suh (1993) feel that McKoon & Ratcliff (1992) failed to find evidence of global processing and motivational inferences because of procedural difficulties in their study. They believe that the McKoon & Ratcliff study may have used inferences which did not relate to the test sentences, since it was not possible to say which, if any, inferences the subjects actually made. Also they used only reaction time measures (the verbal protocol methods used by Trabasso & Suh at least confirmed which inferences subjects were making before on-line methods were used). McKoon & Ratcliff also relied on negative findings in their study to make their case by relying on proving the null-hypothesis; this, argue Trabasso & Suh, is not good scientific practice since there is a 95% chance of success. Trabasso & Suh rely on experimentally produced and statistically reliable differences to prove their claim that their findings support the constructionist processing model.

3.7 Superordinate Goal Inferences: Long & Golding

Graeser and his colleagues in a series of studies (Golding, Graeser & Millis, 1990; Graeser, 1981; Graeser & Clark, 1985; Graeser, Lang & Roberts, 1991; Graeser & Murachver, 1985; Graeser, Robertson & Anderson, 1981; Graesser, Robertson, Lovelace & Swineheart, 1980) found that when readers are asked why-questions about characters' intentional actions in stereotypical stories, they (the readers) reliably produce a small set of causal inferences. For a typical fairy tale about a dragon, for example, these include superordinate goals (eg *the dragon wanting to eat the King's daughters*), goal initiating states (eg *the dragon was hungry*), causal antecedent events (*the dragon saw the daughters*) and enabling conditions (*the dragon was nearby*). They found that superordinate goal inferences were produced in response to these why-questions more often than other types of causal inferences (subordinate goals).

Subjects also included superordinate goal inferences in recall and in summarization protocols in addition to generating them as answers to why-questions (Abbott & Black, 1986; Black & Bower, 1980; Graeser, 1981; Graeser & Clark, 1985; Graeser, Robertson, & Anderson, 1981; Schank & Abelson, 1977; van den Broek & Trabasso, 1986; Wilensky, 1983).

However these protocols (recall, summarization, and question-answering tasks) do not discriminate between inferences drawn during comprehension and those drawn at the time of the test probe.

The minimalist hypothesis states that superordinate goal inferences should only be drawn if (a) they are necessary for referential coherence or (b) they are supported by readily available world knowledge. The global coherence model sees text coherence being established when the reader is able to construct a network of causally interrelated actions and events. Inferences may be required in order to understand a statement even though the inference is not required for referential coherence.

Long & Golding (1993) consider the on-line status of superordinate inferences as an interesting test of the minimalist/global coherence models.

In previous studies, Long, Golding, Graesser, & Clark (1990) and Long, Golding, & Graesser (1992) reported evidence that readers do spontaneously generate superordinate goal inferences as they read statements in stories when they are given sufficient time to do so (ie when there is a long delay between the presentation of the inference-eliciting sentence and the test probe). Long and her colleagues compared the activation level of concepts relating to superordinate goals and concepts relating to subordinate goals using a lexical decision task. Subjects read stories presented one sentence at a time. Each sentence was followed by a delay and then a test word chosen from superordinate goal inferences, from subordinate goal inferences, and from non-words. Subjects then had to respond and say whether the test probe was a word or not. If an inference has been drawn, then the level of activation of that concept should be elevated and the subject should be able to respond more quickly.

Superordinate goal inferences were compared with subordinate goal inferences rather than unrelated control words for two reasons. Firstly, since both superordinate and subordinate goal inferences tend to be generated from the same knowledge structures, once these knowledge structures are accessed, all concepts within that structure will receive some activation. Therefore, both inference types should show activation. Secondly, within the tests used, each pair of superordinate and subordinate goal inferences was produced in response to questions about the same action statement. Any differences in activation levels should not be due to differences in their compatibility with the preceding context: thus both sets of words should be equally susceptible to the influence of context checking at the time testing.

Long and her colleagues found evidence to support the global coherence model. Decision latencies to words from superordinate goal inferences were faster than latencies to words from subordinate goal inferences.

In the 1990 and 1992 studies, Long had used a self-paced presentation of the text with a relatively short SOA (stimulus onset asynchrony), ie the time between the presentation of the action statement and the test probe. One criticism of the findings is that the subjects may perhaps have spontaneously generated the superordinate goal inferences because they had sufficient time to elaborate the text representation. Till, Mross & Kintch (1988) found that subjects spontaneously generate topical inferences after a SOA of 1000 milliseconds but not for SOA's which were shorter (eg 250-500 milliseconds). The SOA used in the 1990 and 1992 studies was 100 milliseconds after subjects had read the sentence plus the actual time to read the sentence, a total in the order of 2000-2500 milliseconds.

Therefore in the Long & Golding (1993) study, an attempt was made to assess the activation of superordinate goal inferences under relatively demanding time constraints, again using a lexical decision task. Instead of a self-paced presentation of text as before, the material was presented one word at a time (a rapid serial visual presentation- RSVP), with lexical decision targets presented after a short SOA. This meant that subjects were forced to read each sentence at the rate dictated by the presentation rather than at a more leisurely rate which was possible in self-paced protocols. After all the texts had been read, a comprehension test was administered. Thus longer term retention is being tested than in some previous studies of inference generation, eg Till, Mross, & Kintsch (1988). Long and Golding found that:

(1) Subjects' performance on the comprehension test varied considerably. The authors felt that this might be due to the fact that an RSVP procedure was used instead of a self-paced method. As a result, some readers may have failed to form a discourse representation and thus performed poorly as regards comprehension.

(2) Superordinate goal inferences are more likely to be generated during comprehension than were subordinate goals but only by those subjects who performed well on the comprehension test.

These findings support the global coherence model and replicate the pattern of findings from Long et al (1990) and Long et al (1992).

However, it might be argued that the faster decision latencies for superordinate items may be due to the influence of context on the postaccess process. For example, Balota & Chumbley (1984), Chumbley & Balota (1984), Forster (1981), and Potts, Keenan & Golding (1988) have found that lexical decision time can be facilitated by context checking at the time of testing. Therefore subjects do not generate superordinate goals when they comprehend action statements but rather a backward association is elicited between the test word and the preceding action statement. This may operate for superordinate but not subordinate goal inferences. However, Long et al (1992) investigated a context-checking explanation of their results by using a word naming task, where subjects have to say the probe word out loud, and which is less sensitive to context checking. They again found superordinate goal inference words facilitated relative to subordinate goal inference words, and concluded that the evidence was against a context checking explanation.

Long et al (1992) also found that superordinate and subordinate goal inference words are both facilitated by semantic associations to lexical items in the preceding action statements but to the same extent.

None of the subjects who did poorly on the comprehension test reported any difficulty in understanding the texts with the presentation rates used, but some were surprised that they could not answer the comprehension questions. Some readers commented that they could remember what the stories were generally about but could not recall the details. One possible explanation is that perhaps they had time to complete word and sentence level processes (eg word recognition, case-role assignment) but did not have time to integrate information across sentences (Perfetti, 1985: Perfetti & Roth, 1981). Without the information being sufficiently integrated, no lasting memory representation is created.

It appears that superordinate goal inferences are facilitated relative to subordinate goal inferences within about 250 milliseconds. These findings are incompatible with the minimalist hypothesis which predicts that neither superordinate nor subordinate goal information would be generated as neither is necessary for local text cohesion. The findings do support the global coherence model. Superordinate goal inferences provide a causal network which facilitates overall textual coherence and offers support for the global coherence model.

3.8 The Time Course for Generating Causal Antecedent and Causal Consequence Inferences: Magliano et al

As well as considering which types of inference readers may or may not make, the question could be asked as to when these inferences are generated. Are they generated at the time of reading (ie on-line) or later when a probe question is answered (ie off-line)? Even when an inference is generated on-line, are all inferences generated with the same speed?

Some researchers who have studied the time course in which inferences are generated claim the time course of inference activation varies with different inference categories. For example, McKoon & Ratcliff (1986, 1989a, 1992) distinguish between on-line inferences which are either automatically or strategically generated.

Automatic inferences are generated without awareness and relatively quickly (within 650 milliseconds after comprehending a sentence). They are stored in highly automized knowledge structures and consist primarily of those inference categories needed to establish local coherence, such as causal antecedents. For example, if the sentence "*John left the bar and got into an accident*" is read, a causal antecedent is likely to be "*John had been drinking*".

Strategic inferences depend on reader goals and the text genre. They take more time to generate than automatic inferences and are not mandatory for comprehension. For the same sentence in the paragraph above, a strategic inference could be the causal consequence that "*John was hurt*".

Magliano & Graesser (1991) believe that, while the above distinction is useful, some inferences are inevitably generated on-line but take more time to construct than automatic inferences. Inevitable inferences (to use Magliano & Graesser's term) are mandatory for comprehension but are not generated automatically.

Till, Mross, & Kintsch (1988) investigated the time course of thematic inferences, inferences which capture the main point of a sentence. They used an RSVP (rapid serial visual presentation) procedure with a varying SOA (stimulus onset asynchrony) intervals of 200, 300, 400, 500, 1000 and 1500 milliseconds, with a lexical decision task to measure activation of inferences. In other words, subjects read a short text presented one word at a time and displayed for a short interval (RSVP). After the last word is presented, there is a short delay (the SOA interval) before the test word is presented. These test words were either words related to the inference or unrelated, or non-words. If an inference has occurred during reading then the concept should be activated and response to the test word should be faster than when the concept is not activated. They found that activation increases monotonically as the SOA increases, with a critical interval of 1000 milliseconds. In other words, the inference is stronger as the interval increases between the presentation of the last word of the text and the presentation of the test probe.

However there are some methodological problems with the Till et al (1988) study. Firstly, they do not distinguish among inferences categories. For example, some texts include superordinate goal inferences while others include inferences for unintentional events. It is possible that some inference categories were generated at SOA intervals shorter than 1000 milliseconds but their effect was cancelled out by other inference categories in the statistical analysis. Sharkey & Sharkey (1992) found facilitation for script-based inferences at an SOA interval of 200 milliseconds. A second criticism is that the RSVP rate of 333 milliseconds may not allow enough time between each word presentation for subjects to comprehend critical words. It could also lead to a high amount of buffering (ie the holding of many words in short-term memory) and thus place high demands on working memory at the point when the test probe is received. This could have an impact on what inferences are generated as well as the time course in which readers generate them.

Magliano, Baggett, Johnson & Graesser (1993) focused on (a) causal antecedents, which are directed backwards in a causal network, and (b) causal consequences, which are directed forward in such a network.

Causal antecedent inferences establish text coherence by connecting a sentence with prior text and are generated on-line in narrative texts (Keenan, Baillet, & Brown, 1984; Myers, Shinjo & Duffy, 1987; Singer & Ferreira, 1983; Singer, Halldorson, Lear, & Andrusiak, 1992; van den Broek, 1990b). An example used by Magliano et al is the couple of sentences below:

A thirsty ant went to to a river.

It became carried away by the rush of the stream.

The second sentence can be causally connected to the first by the inference "*the ant fell into the stream*". This causal antecedent establishes text coherence.

Causal consequence inferences predict or forecast subsequent episodes in the plot and are not generated during comprehension, ie on-line (Graesser, Haberlandt & Koizumi, 1987; McKoon & Ratcliff, 1989a; Millis, 1989; Potts, Keenan, & Golding, 1988) or at best are only minimally inferred (McKoon & Ratcliff, 1986). For the example above, the reader might infer that "*the ant drowned*".

Magliano et al (1993) used a three pronged approach advocated by Magliano & Graesser (1991), and similar to that used by Trabasso & Suh (1993). This is discussed below:

(a) They used a question-answering methodology to expose what inferences are made. Subjects were asked "why" and "what-happens-next" questions as they read the narrative statement by statement.

(b) Three theories of discourse processing were used to predict what inferences are generated on-line. These were:

(1) Textbase model: This model assumes that the text is adequately represented by a list of explicit propositions that contain its meaning or gist (Kintsch, 1974; Kintsch & van Dijk, 1978). Text cohesion is established through argument overlap. The only class of inference that is generated on-line is anaphoric reference. This model predicts that neither causal antecedents nor causal consequences are generated on-line.

(2) Bridging model: This includes models and hypothesis by several researchers eg Graesser & Clark, 1985; Kintsch, 1988; McKoon & Ratcliff, 1986, 1989a, 1992; Singer, Halldorson, Lear, & Andrusiak, 1992; van den Broek, 1990a; van Dijk & Kintsch, 1983. This model assumes that the text representation contains the explicitly mentioned text (where propositions are related by argument overlap) plus coherence-based inferences, such as causal antecedents which connect the current sentence with previous text. Causal consequences are seen as elaborative inferences which are not normally generated on-line.

(3) Prediction-substantiation model: This model has its origins in both artificial intelligence (Dyer, 1983; Schank & Abelson, 1977) and cognitive psychology (eg Sharky, 1986). It assumes that knowledge structures (eg frames, scripts, schemas) are used by readers to generate expectations about subsequent events and actions in

narratives. It predicts the generation on-line of both causal antecedents and causal consequence inferences.

(c) Behavioural methods were used to provide a rigorous empirical test of whether a particular inference category was generated on-line. In this particular case a lexical decision task was used in conjunction with a rapid serial visual presentation (RSVP) of the text.

Magliano et al (1993) list some advantages of using the three-pronged method. Firstly, in other studies (eg Kintsch & Mross, 1985) experimenter generated scripts are used which intuitively identify actions in several script-based activities (eg eating in a restaurant). The question-answering methodology reveals what (if any) inferences subjects actually make, not what the experimenter thinks they will make.

Another linked advantage is that such a methodology allows the study of inferences in the context of naturalistic texts. Passages had hitherto used only one or two sentences (eg McKoon & Ratcliff, 1986; Potts, Keenan, & Golding, 1988). These short texts allowed the researchers greater experimental control over the materials. However such short passages seldom have a rich enough context to support the normal volume of inferences that are presumably generated when reading naturalistic text.

The design used was similar to that used by Till, Mross, & Kintsch (1988) but because of the difficulties noted above, Magliano et al (a) manipulated inference category (antecedent versus consequence) and (b) varied the RSVP rate (250 versus 400 milliseconds), as well as the SOA interval (only the SOA interval was varied by Till et al).

Subjects read short passages which were presented one word at a time for either 250 milliseconds or 400 milliseconds. After the final word, a target word was presented after intervals of 250, 400, 600, or 1200 milliseconds. Subjects had to say if the target was a word or not: again inference concepts, non-inference concepts, and non-words were used.

Magliano and his colleagues found evidence which supported the bridging model, ie they found that causal antecedents were generated on-line but causal consequences were not.

Magliano and his team point out certain potential difficulties with their study. Firstly, causal consequences have been shown by other researchers to be generated on-line when the context is highly constrained and allows for only one or two possible outcomes rather than several alternatives (eg Murray, Klin & Myers, 1991; van den Broek, 1990b). The Magliano passages possibly did not provide a sufficient amount of contextual constraints to elicit consequence inferences.

Secondly, the consequence inferences revealed by the question-answering protocols extended far into the future compared with the antecedents which tended to be more locally bound, ie the consequences are a greater distance in the causal network than the antecedents and, as a result, are less likely to be generated. Future investigations in this area would need to equate the distance of causal antecedents and consequences within the causal network.

With regard to the time course of inference generation, Magliano et al found that causal antecedents were generated on-line if readers were given a sufficient amount of time to interpret the text and generate inferences. Specifically, the presentation rate of text (the RSVP) had to have an interval between words of 400 milliseconds or greater. In addition, the interval between the presentation of the last word of the text and the presentation of the test probe (the SOA) had also to be 400 milliseconds or greater.

This having been said, Magliano and his colleagues admit that it is not possible to say exactly when within the combined SOA interval and the decision latency period the inference is actually generated.

The implications of this study are that adult readers tend to generate at least causal antecedents on-line. They may generate causal consequences but the Magliano et al

(1993) study has certain implicit problems with design which failed to stimulate these inferences. At the very least, they have shown that the generation of an inference requires a minimum amount of time of around 750 milliseconds. However, it is possible that other categories of inference may require more or less time than this.

3.9 Network Representations of Causal Relations: van den Broek & Lorch

The identification of causal dependencies in a text can involve establishing connections between a focal sentence and immediately preceding text or textual information that is further removed in the surface structure of the text.

Two theoretical models have been put forward to describe how readers may connect this information. *Linear models* see each text unit connected to the immediately preceding text unit. Connections are not established between non-adjacent units so long as connections can be made between adjacent units (eg Fletcher & Bloom, 1988 (above); Kintsch, 1988). *Network models* propose that multiple connections are possible to a single text unit. This can include those connections suggested by the linear model as well as others that relate units which are quite far apart in the surface structure of the text (eg Trabasso & van den Broek, 1985 (above); Trabasso & Sperry, 1985).

van den Broek & Lorch (1993) in a series of studies tried to determine which of these models readers actually use, ie do they use a network of multiple relations or a linear chain. The narrative texts which were used in these studies could be represented either by a linear chain or by a network. An example is shown below:

There once was a boy named Bob. One day, Bob saw his friend's new 10-speed bike. Bob wanted to get a 10-speed. He looked through the yellow pages. He called several stores. He asked them about prices for bikes. He went to the bike store. He asked the salesperson about several models. The salesperson recommended a touring bike. Bob looked at the selection of touring bikes. He located some that were his size. He found a bike that was metallic blue. Bob bought the beautiful bike.

A speeded recognition procedure was used to test subjects' memories for story events. After reading the text, a pair of sentences were presented as the recognition task. The first, the priming sentence, reminded them of either the story (general prime) or a specific event in the story (specific prime). The second sentence was a target sentence. The specific primes were either causally related or unrelated to the target event, and were either adjacent to or distant from the target in the surface structure of the text. The general prime for the story above was " *Remember the story about Bob*". A specific prime was "*Remember Bob looked through the yellow pages*". One target sentence paraphrased a subsidiary action in the story ("*Bob called several stores*") while the other target paraphrased the outcome of the story ("*Bob bought a bike*").

van den Broek & Lorch found faster response times when the targets followed a specific prime that was causally related than when it followed either a specific but causally unrelated prime or general prime. Most importantly, this effect was found both when the specific prime and target were adjacent and when they were distant in the surface structure of the text. The textual materials used in these studies always had a causal relation to the immediately preceding event. Therefore from a linear chain perspective, there was no need for subjects to reactivate the distant information to establish coherence. The fact that priming is observed for both indicates that multiple relations are represented.

These results converge with evidence from Trabasso & Suh (1993). However the present results additionally show that the relations between distant statements in a text are not only detected but are also incorporated in the reader's mental representation of the text. They support a network model of the representation of causal relations in narratives; they are inconsistent with a linear chain model.

3.10 Situation Model Construction: Fincher-Kiefer

It is generally accepted, for example by Johnson-Laird (1983), Just & Carpenter (1987), Mani & Johnson-Laird (1982), Perfetti (1989), Perrig & Kintsch (1985), Sanford & Garrod (1981), and van Dijk & Kintsch (1983), that when readers comprehend a text, they construct a multilevel representation of the text. Initially a surface structure representation is generated; this includes word identification and syntactic parsing. This process is seen as being obligatory and impervious to knowledge effects (Perfetti, Beverly, Bell, Rodgers & Faux, 1987; Rayner, Carlson & Frazier, 1983; Kintsch & Mross, 1985). A semantic analysis of the text results in a propositional textbase which provides the reader with meaning, albeit in a restricted form (Perfetti, 1989). Typically a more enriched level of representation is constructed wherein the reader supplements the propositional textbase with prior knowledge. This has been referred to variously as a situational model (van Dijk & Kintsch, 1983) or a mental model (Johnson-Laird., 1983). This leads to much richer inferences being possible (Perfetti, 1989).

McKoon & Ratcliff (1986, 1989b, 1989c) have argued that inferential processing should be considered within a framework which allows for variability; in a restricted context inferences may be encoded to a "high degree" but others, such as those that concern predictable events, may be only partially encoded.

As little previous research has been directed towards inferencing in light of this multilevel representation model of comprehension, Fincher-Kiefer (1993) proposes and tests another framework which is consistent with that of McKoon & Ratcliff, 1986, 1989b, 1989c). She assumes that the generation of the surface structure code involves no inferential processing, while construction of the propositional textbase entails limited inferential processing which is concerned with maintaining local coherence. The development of the situational model involves elaborative inferential processing eg predictive inferences.

Fincher-Kiefer assumes that different tasks tap different levels of mental representations (Lucas, Tanenhaus & Carlson, 1990). The first experiment tested the hypothesis that predictive inferential processes are not employed in the generation of a propositional textbase. Subjects were interrupted during their reading of predictive texts and asked to make word recognition decisions for various target types. For the second experiment, the nature of the decisions subjects were being asked to make were changed to reflect more closely the processes used in constructing a situation model. A word prediction task was used in which subjects had to decide if a word might appear in future text instead of having to decide if a word had appeared in previous text. The third experiment was similar to the second but used a lexical decision task. Since this does not require the reader to address a particular level of text representation, it should be sensitive to processes that occur in multiple levels of representation (Lucas et al, 1990; Masson & Freeman, 1990).

An example of the type of text used is shown below:

Henry was going to the dentist. This time he especially dreaded the trip because he knew he had several cavities. Sure enough, the dentist located the cavities and asked Henry to open his mouth wide. Henry knew he should have been more careful about flossing. Every time he went to the dentist this procedure was mentioned to him, but he just hated to do it. The dentist warned him that gum disease was worse than having cavities .

The predictable event which is elicited by the first three sentences is that Henry had his teeth drilled. The last three sentences neither confirm nor refute this prediction. Thus the initial inference can generate expectation without there being a later need for referential coherence. The target decision tasks were only presented to subjects after the first three sentences (and the predictable event had been read.

In her first experiment, Fincher-Kiefer compared word recognition response times for (a) a word associated with the predicted event (eg *drill* in the example above) with (b) a semantic associate, ie a word which was appropriate in terms of general knowledge for the situation described in the story (eg *teeth*), and (c) neutral control words. No difference was found between predicted targets and semantic associates. This is interpreted as supporting the idea that anticipatory or predictive inferences are not encoded into a reader's propositional textbase, thus confirming previous findings eg McKoon & Ratcliff (1986).

The second experiment suggested that predictions are considered during reading. When readers are encouraged to attend to the meaning of a text instead of what is explicitly stated (the word prediction task), the predicted targets show greater facilitation effects than schematic associate targets. The third experiment, using a lexical decision task, provided converging evidence for the results of the second experiment. The facilitated response times for the predicted targets show that the construction of a situational model involves more than the temporary activation of contextually appropriate knowledge; it involves the generation of inferences concerning highly predictable actions and events.

Fincher-Kiefer feels that word-naming is insensitive processes
occurring at levels of text analysis higher than the word access level. The lexical decision task is more sensitive to inferential processes that are known to occur during encoding.

The results of these three experiments support a model in which multiple levels of representation are constructed during reading. The first level appears to provide only restricted meaning and approximates to the explicit text presented. If any question arises concerning the details of the explicit text, the textbase can be consulted. The situation model, however, provides enriched meaning and interpretation.

Both levels of representation (the textbase and the situation model) occur in parallel in that they are initiated simultaneously. However, while the propositional representation may be generated, the situational model can still be incomplete, or at least tentative. The situation model construction lags slightly behind the textbase construction because the reader must have basic propositions to work with before a text model can be built.

The results support the minimalist view of inference generation. However the data also suggest that the situation model is an active, flexible and continuous level of representation, whose generation includes many types of extratextual influences, including predictive inferences.

3.11 Spatial Representations in Naturalistic Text: Zwaan & van Oostendorp

Although many studies, as described above have focussed on how readers make causal inferences whilst reading, other types of inferencing, such as spatial relations, have also been studied.

Some studies have shown that skilled adult readers are able to construct rather detailed spatial representations and use them in further processing of the text (Morrow, Bower & Greenspan, 1989; Morrow, Greenspan & Bower, 1987). However, subjects in these studies first memorised a layout of a building and proceeded to read a relatively simple story detailing actions in that building. Zwaan & van Oostendorp (1993) feel that all that these studies reveal is the competence of readers to construct spatial models but they may not reflect their actual performance during text comprehension. If the spatial information is presented with little or no context (eg in an isolated sentence) then the only dimension for subjects to apprehend is spatial information, thus facilitating the construction of a spatial model. More complex stretches of text may yield other results.

In naturalistic text, various properties of the text structure have been shown to facilitate the construction of spatial representations. These are (a) determinancy, (b) continuity and (c) condensation.

Determinancy refers to how restricted a text reference is and evidence has been reported by Mani & Johnson-Laird (1982) for its effect on the formation of spatial representations. If a text supports only one representation (or better, class of representations) readers tend to form that representation. However if a description is indeterminate (ie can give rise to more than one class of representations), readers only form a surface and propositional representation rather than a situational one. For example, if a text refers to "*a house*", this is indeterminant. However, mention of "*a bungalow*" can lead to a particular spatial representation, ie one would expect all the accommodation to be on one level.

Continuity describes how well connected the text is. Ehrlich & Johnson-Laird (1982) have shown that for continuous descriptions, where each sentence after the first refers back to an object mentioned in the immediately preceding sentence, readers could construct spatial representations. If this chaining principle is violated and the description is discontinuous, then readers have difficulty integrating the spatial information into a single model. Parallel research into spoken language shows that speakers attempt to make their spatial descriptions continuous (Levelt, 1989; Linde & Labov, 1975).

Even when a description is continuous in the sense that each object and its location is mentioned after an adjacent object, the surface structure of the text may not be continuous. The spatial information may be interspersed with nonspatial information, eg the description of the location of a building could be followed by a description of its inhabitants or of the history of the building. Zwaan (1993) has shown that the amount of nonspatial information between spatial information influences the quality of the spatial representation constructed. If the spatial description is distributed over a large surface text area, even although continuous in terms of the spatial model,then the poorer will be the spatial representation which is constructed by the reader. The more *condensed* the description, the better the spatial representation. This could be due to limitations of short-term memory, and therefore is not unique to spatial processing (Sanford & Garrod, 1981).

It seems that these three properties are only partly present in most naturalistic, fictional narratives; therefore, perhaps the intention of authors is to convey a general impression rather than a distinct spatial representation.

If readers adopt a minimal coherence strategy (McKoon & Ratcliff, 1992, 1990), then causal information is likely to be more important than spatial information in constructing coherence. Spatial situation models are only likely to be considered more important if subjects read with a special purpose or with particular instructions, perhaps to pay attention to spatial layout.

Zwaan & van Oostendorp (1993) examined the effects of two different reading instructions. Subjects were either told to read the text as they normally would (the normal condition) or to focus on spatial information and keep track of the location of objects mentioned in the text (the spatial condition). They measured reading times on a self-paced reading task and reading times and accuracy of response on an inference judgement task. They proposed that the more accurate the inferences are drawn, then the stronger the spatial model subjects had constructed.

They had four hypotheses:

(1) Subjects in the spatial condition should allocate more encoding time for spatial information than subjects in the normal condition; therefore spatial information would be encoded more slowly during the reading task.

(2) Subjects in the spatial condition would have a stronger spatial representation than the controls, so that the spatial inference judgement task should yield faster and more accurate responses.

(3) Subjects in both conditions should have relatively good representation of non-spatial information, such as information about actions, traits, and motives of the characters. This test acts as a control. If subjects in both groups perform equally well on non-spatial inferences, any differences on spatial inferences would suggest that the different reading instructions produce qualitatively different modes of reading and different cognitive representations; the difference could not be explained by the spatial condition subjects just being better overall at processing textual information.

(4) Wender (1989) proposes that a well established spatial representation has beneficial effects on the processing of subsequent information which elaborates or transforms the spatial layout. Zwaan & van Oostendorp added such new spatial information and measured reading times for this information (which should be faster for the spatial condition) and judgements about spatial inferences pertaining to these modifications of the spatial model (for the spatial condition, these should be more accurate).

The textual material was naturalistic and selected to meet the conditions of determinacy, continuity and condensation mentioned earlier. In addition, a detective story was chosen as spatial representations are quite important in this genre since they frequently contain clues about the crime. Experienced readers should have internalised this conventional wisdom and so spatial information would not be completely ignored. Individual differences in reading rate and spatial ability were partialled out using covariance.

Zwaan & van Oostendorp found support for all four hypotheses. Reading times for the spatial condition were indeed slower than than for the normal condition. Subjects in the spatial condition made more accurate spatial inference judgements although there was no difference in response times. Subjects in both the spatial and normal conditions had equally strong representations of non-spatial information. In the spatial condition, readers checked this non-spatial text for potentially relevant information, hence the longer reading times, but this added time did not lead to better performance. Subsequent spatial information was processed more quickly and accurately by the spatial subjects.

Therefore, construction of spatial representations does not seem to be a crucial aspect of naturalistic story comprehension. Under normal conditions, readers construct relatively weak spatial representations. However, they can construct relatively strong spatial representations, but only if specifically instructed to do so.

Zwaan and van Oostendorp mooted, but did not investigate in their study, that the relevance of an object to the causal chain of a story might influence the reader's knowledge of the location of such an object.

3.12 Overview

If one wishes to explore the mechanisms that constitute competent reading, it seems sensible to examine this process in fluent, adult readers. Since inferencing has been seen as central to the comprehension process (Schank, 1976), this has been an area to which extensive research has been directed. However, despite almost twenty years of such investigation, and continual refinements in methodology and experimental design, it has proved very difficult to construct a unified theory of inference processing.

The studies discussed above focus on issues of current concern. In particular, what inferences are generated (eg inferences concerned with local cohesion in the text or more global inferences), as well as when these inferences are generated: are they generated on-line (ie during the comprehension process) or off-line (ie during a subsequent retrieval task)? However, the variety of studies quoted have tended to focus on different types of inference. Some categories of inference have been extensively studied while others have almost been neglected (Graesser & Kreuz, 1993). These are listed below, starting with the most studied (references) through to those least studied (readers' emotions):

referential inferences > causal antecedent > causal consequence > instrumental > superordinate goal > subordinate goal > .... > thematic > emotion of reader.

The methodologies used have varied also, from memory probes, through latencies on lexical decision, to on-line measures of reading times. Each seems to have advantages as well as disadvantages.

The majority of the studies mentioned in this chapter have concerned themselves with establishing support for either the constructionist/ global inference theory (Trabasso, Trabasso & Suh, Long & Golding, van den Broek & Lorch), or the minimalist/local coherence theory (Fletcher & Bloom, McKoon & Ratcliff, Magliano et al). The constructionist standpoint sees the reader as making elaborative inferences which embellish the text representation while the minimalists see the reader as making bridging inferences for local text cohesion. Fincher-Kiefer proposes a multilevel processing model in which the reader tends to adopt a minimalist approach but also constructs a situational model, although this is likely to be tentative and incomplete.

Although these studies fall into two theoretical camps, the types of inference used in each cut across the divide. Trabasso & Suh and van den Broek & Lorch consider bridging inferences while Fincher-Kiefer, Long & Golding, and Magliano et al consider elaborative inferences. (In the Magliano et al study antecedent inferences are used. Although these are backward directed, they are not bridging inferences since the antecedent goal does not appear in the text. They are considered to be elaborations).

There is remarkable consistency in the findings of the Trabasso & Suh and the van den Broek & Lorch studies. Texts were locally coherent and attempts were made to control for context checking of the test probe. van den Broek and Lorch also tried to address the issue of lexical priming. Trabasso & Suh's approach had the advantage of converging results from three sources: the inferences predicted from their theoretical model matched the inferences which readers reported they had made, with the additional evidence of reading times to confirm that these inferences were made on-line. These two studies, along with that of Long & Golding, provide support for the global/constructionist view.

Although the second group of studies examined elaborative inferences, each looked at a different subset of causal elaboration. Long & Golding's included superordinate and subordinate goals: Magliano et al included superordinate goals with causal consequences: Fincher-Kiefer included causal consequences. Despite some discrepancies among these studies, they do provide some evidence of elevated activation of causal elaborations during (compared to other inference categories).

The efficacy of a single word probe (as in naming or lexical decision tasks) has been questioned by Magliano & Graesser (1991). Can such a probe access an adequate level of representation of a complex inference? A sentence or phrase would seem to be better, for example, "*Did the dragon eat the princesses?*" would seem to be better than "*eat*". An obvious disadvantage, though, is that use of a longer phrase does permit context checking to occur more easily. Trabasso & Suh and van den Broek & Lorch presented probes in this more complete way.

It is obvious, when investigating knowledge-based rather than text-based inferences, that the readers should have a sufficient knowledge base, in order to generate such inferences. This is an evident advantage of the "talk-aloud" procedures used by Trabasso & Suh and the question answering protocols used by Long & Golding and Magliano et al in that they allow the experimenter to check what inferences the readers are actually making, and not rely on speculation about which they might be making. While it cannot be claimed absolutely that what readers report is exactly what they infer, the reported items can at least be viewed as very likely candidates for on-line detection. Conversely, an item which does not appear in the protocols can be regarded as being very unlikely to accompany comprehension. However, one major disadvantage of talk-aloud protocols is that such procedures might orient the reader in a particular way, such that inferences could be drawn that would otherwise not be made. At the very least, it could be argued that the whole process bears little resemblance to normal reading.

Another point which is important but often overlooked is the fact that the type of material being read and the readers purpose in reading it can have an effect on which inferences are likely to be made. Zwaan & van Oostendorp have shown that readers' goals determine how much spatial information is constructed when reading a mystery novel. Thus text genre may determine the particular goals that readers adopt during comprehension. For example, keeping track of the spatial layout of a house might be useful when reading a thriller about a jewel thief but not when reading a romance novel.

Although it is difficult to incorporate all the findings of the above studies into one overall model, some general precepts can be drawn. Inferences which aid local coherence of a text probably occur automatically. More global inferences occurring over a greater surface distance of the text are likely to occur if there is a break in cohesion. However, the motivation of the reader is also important, and such global inferences may occur if the reader is actively seeking causal relationships within the text, for example, if it is a spy thriller. More elaborative types of inference also seem possible but perhaps only in circumstances where the context is highly constrained.

Given the range of experimental techniques which have been used in the various studies considered so far, the next chapter examines the advantages and drawbacks of each technique more closely.

Chapter Four

Methodological Problems when Studying Inferencing.

As mentioned in previous sections, the results of various studies on inferencing often produce contradictory results. Keenan, Potts, Golding & Jennings (1990) cite several examples of studies which give differing results from different studies, not only between different types of elaborative inferencing but also different results from various studies of the same type of inference; for example, Dosher & Corbett (1982); Lucas et al (1987); McKoon & Ratcliff (1981); Paris & Lindauer (1976); Singer (1979) - INSTRUMENTAL INFERENCES; Anderson & Ortony (1975); Garrod & Sanford (1977); McKoon (1988); O'Brien et al (1988); Whitney (1986) - INSTANTIATIONS OF GENERAL TERMS; Duffy (1986); McKoon & Ratcliff (1986); Potts et al (1988); Singer & Ferreira (1983) - LIKELY CONSEQUENCES OF EVENTS.

These studies have concerned themselves primarily with what type of inference (if any) was drawn. They used a variety of paradigms and materials varied greatly, from singles sentences in some cases to paragraphs in others. In some paragraphs previous mention is made of the inference concept, in others the context is highly constrained. Two different methods are used to detect whether an inference has been drawn, usually memory measures or activation measures. Additionally, the time when the test probe is initiated can vary from during the reading process itself, to just after the text is read, to a slightly greater delay after reading. Each of these methodological aspects will be considered below. Inherent difficulties will be discussed and possible solutions suggested.

Activation measures, which will be described in more detail below, detect inferences by seeing whether an inference concept is primed, in other words, whether it is more activated after reading an implicit version of the text compared to a control version.

Activation measures include naming, lexical decision and the modified Stroop task. An inference concept can be activated by (a) intralexical associations or word-based priming, and (b) the reader drawing the inference from his or her knowledge of the situation described in the text. Most researchers would not consider word-based priming to be true inferencing and would try to control for it in their experimental designs (eg Keenan, Golding, Potts, Jennings & Aman, (1990); McKoon & Ratcliffe (1986).

Memory measures, which again will be described later, require the reader to access his or her representation of the text to see if the inferential information is part of that representation. They include cued recall, sentence verification, question answering, and recognition tasks.

Which type of measure is used (activation or memory) often depends on what the researcher considers constitutes an inference. Simple activation of a concept can be considered as an inference and so an activation method seems appropriate. Others might consider activation of a set of concepts through to activation of a higher order structure such as a schema to be necessary before considering an inference to have been drawn. They would, therefore, reject activation of a single concept as sufficient. Their counterparts could in turn argue that activation of a whole set of concepts can be detected by measuring the activation of a single concept from that set.

Another factor which is often considered is the level to which an inference is processed. This can range from simple activation, through incorporation in working memory, to inclusion in the long-term representation of the text (Kintsch, 1988). Those who believe that a true inference is only that which is included in the final representation of the text will, therefore, prefer memory measures over activation.

4.1 Cued Recall

Cued recall was one of the earliest techniques used to detect elaborative inferences. An example from Corbett & Dosher (1978)
illustrates its use:

The athlete cut out an article with scissors for his friend.	EXPLICIT
The athlete cut out an article for his friend.	IMPLICIT

After reading a sentence such as the above, the information which is either stated explicitly or implied (in this case the scissors) is used as a cue and the subject is asked to recall the sentence. Paris & Lindauer (1976) found that an instrument was just as effective as a cue to recall in the implicit sentence as well as the explicit. They concluded, therefore, that the instrument had been inferred during encoding.

Singer (1978) challenged this conclusion and suggested that the cued recall reflects a reconstruction process which occurs during retrieval rather than the inference occurring at encoding. Since the cued recall occurred some time after encoding, the sentence is not readily available and must be retrieved from memory. This involves processing the sentence again which could induce construction of information (the inference) which was not part of the original encoding. In other words, the test demands themselves could have caused the inference to be made.

Singer verified this suggestion by using the cued recall method for sentences which contained actions such as "stirred the soup". Such an action can imply two possible instruments, spoon and ladle. However, these instruments do not relate to the verb "stir" in a symmetrical way. There is a strong forward but weak backward association between "stir the soup" and spoon, whereas there is a strong backward and weak forward association between the action phrase and ladle. If the instrument is inferred during reading, then the subject should infer "spoon" more often than "ladle" because it is more strongly associated in that direction. However, if the inference is made at retrieval, then "ladle" should be more likely than "spoon" due the stronger backward association of that word. This is indeed what Singer found, confirming that inferencing is occurring as a result of reconstruction.

4.2 Sentence Verification

This approach involves the subject reading either an implicit or explicit piece of text and then responding "true" or "false" to a sentence. An example from Keenan & Kintsch (1974), which involves backward or bridging inferences, is shown below, although the technique has been used to detect forward inferences also (eg implied instruments, Singer, 1979; inferences concerning the consequences of events, Singer & Ferreira, 1983):

Gas leaked from a butane tank.)
The explosion levelled a service station and a new home.)
IMPLICIT

Gas leaked from a butane tank and caused an explosion.)
The explosion levelled a service station and a new home.)
EXPLICIT

The subject is then asked to respond by answering true or false to the statement: *A gas leak caused the explosion*. If subjects are just as fast to verify the information after the implicit condition as they are after the explicit version, then it was assumed that they drew the inference while reading.

However, there is a problem with the above logic as the paradigm involves accepting the null hypothesis. That is, evidence for the inference having been made depends on finding no significant difference between the two conditions. Of course, a lack of significant difference could be caused by other reasons, for example lack of power in the statistical analysis being used.

Another problem with this approach has been pointed out by Keenan & Kintsch (1974) and by McKoon & Keenan (1974). The probe statement to be verified will have been presented explicitly in the control version of the text but not in the implicit version. It is possible, therefore, that the subject's memory for the surface form of the statement can facilitate or prime the processing of the test statement following the control condition. Consequently, it could be that subjects consistently draw inferences in the implicit condition but that this is masked by the decrease in verification times under the control version because of priming.

A further problem can arise when sentence verification is used to detect forward inferences. It has already been noted that elaborative inferences are likely, but not necessarily true, elaborations of the text. As a result, when the subject is faced with a verification statement, he or she knows that the statement is likely to be true, but since it was not stated explicitly in the text, there is always the possibility that it may not be true. A example from Singer & Ferreira's (1983) study was:

Bob threw the report into the fire.

It is likely that the report was burned but it is also possible that Bob had a poor aim and missed the fire altogether. Certainty can only be assumed in the explicit version or in a version where a backward inference can be drawn, eg *The report's ashes went up the flue.*

Therefore, even when subjects do make forward inferences, they may hesitate to endorse it as readily as they would under explicit or bridging inference conditions. This could slow down the verification times, suggesting that no inference was drawn when in fact it had been.

4.3 Sentence Reading Times

Because of the difficulties mentioned when using the two methods considered above, it was realised that to assess inferences occurring at the time of encoding, it would be necessary to test during encoding, that is, on-line.

Use of sentence reading time measures has already been mentioned in section 3.2.1 above. Haviland & Clark (1974) showed that for bridging inferences, an implicit version of a text takes significantly longer to process. This is so even when word-based priming effects are controlled.

This use of the reading time paradigm for bridging inferences works well because evidence for the inference is an increase in reading time. However, its use in detecting elaborative inferences is less straightforward. To illustrate, the example from Haviland & Clark (1974) quoted earlier can be used:

> *Herb unpacked the picnic things.* IMPLICIT VERSION
> *The beer was warm.*

If the subject makes an elaborative inference that includes beer among the picnic things when reading the first sentence (because the subject knows, perhaps, that Herb would never go on a picnic without taking along some beer), then the time to read the second sentence would be the same under implicit and explicit conditions. The difficulty then arises that to accept that an inference has been drawn requires one to accept the null hypothesis. Therefore, the same results would be obtained as when no inference is drawn under either condition (ie no significant difference is obtained).

One possible solution to this difficulty is to design the materials so that the target sentence does not confirm the expected elaborative inference, but rather contradicts it.

This should result in the reader taking longer to read the target sentence if the inference is drawn. An example from Sanford & Garrod (1981) illustrates this:

(John/The teacher) was on his way to school.
The bus trundled slowly along the road.
Last week he had trouble controlling the class.

If the reader makes the elaborative inference that John is a school boy, then it will take longer to read the final sentence than if the reader had not made that inference, as in the teacher version. However, in the example above, there is still the problem of word-based, associative priming noted earlier. The explicit version, which contains "teacher" is more likely to prime words in the target sentence, such as "class" and "control", than would the implicit version. A possible solution to this problem is suggested by Keenan, Golding, Potts, Jennings, & Aman (1990), They suggest incorporating "teacher" into the implicit version as well:

As John was on his way to school, he thought about what an effective teacher would be like.	IMPLICIT
As he was on his way to school, the teacher thought about what it would be like to be effective.	EXPLICIT

One disadvantage of the sentence reading time paradigm is that it does not allow one to know exactly when an inference is drawn. A more exact process is to use word-by-word reading times by using eye movement measures (eg Rayner & Pollatsek, 1989). This procedure is discussed later in this chapter.

Another possible disadvantage of reading time measures is that increased reading time on the target sentence could simply reflect a difference in the ease with which the target sentence can be integrated with the first sentence because of the different syntactic structures of the first sentences under the explicit and implicit conditions.

In addition to the above, the reading-time paradigm does not reveal exactly what inference is drawn. In the example from Sanford & Garrod above, the "John" version may show that the reader does not infer John to be a teacher, but it does not indicate if he is inferred to be a schoolboy, a bus driver or someone else. The measures considered next address this problem of what the content of the inference is.

4.4 On-line Question Answering

This methodology involves interspersing various questions throughout the text to tap into the reader's developing representation of text meaning, checking unspecified information about what has and what is going to happen (Graesser & Clark, 1985; Olson, Duffy & Mack, 1984). Although this method can be used at any point in the text and can reveal the content of inferences, it is still the case that those inferences might not be drawn in the absence of the questions. The measure itself is invasive, perhaps causing inferences to be drawn that would not normally be made during reading. It is none-the-less useful for revealing potential inferences that a subject may make when reading.

4.5 Recognition

Like the last method above, recognition requires the subject to access his or her representation of the text but it does not ask the subject to make the inference. After reading an implicit or a control version of a text, the subject has to determine whether a test word representing the inference actually occurred in the text. If an inference has been drawn in the implicit version, then the concept will have been activated and it will be more difficult to state if the test word actually occurred in the text or not, leading to longer response times.

Again the drawback with this method is that it is impossible to know whether the inference was drawn while the text was being read or drawn during test.

Keennan et al (1990) point out another potential difficulty with this method. If the subject is led to draw the inference while checking the test probe, this would slow the response time. (One may not have thought of an instrument used to stir the soup, but when presented with "spoon", infer that it would be an appropriate instrument). This is exactly the same result one would expect if the inference was made while reading the text (the concept of "spoon" is activated when the inference is made and this delays confirmation of the test probe).

McKoon & Ratcliff (1986) recommend the use of a deadline procedure with the recognition task in order to eliminate the possibility of subjects drawing inferences at the time of testing. With a strict deadline, it was felt that subjects would not have time for the strategic comparisons of the probe to the text that would allow an inference to be drawn at retrieval. However, Keenan et al (1990) argue that this restriction does not completely eliminate the comparison process and therefore does not remove the problems mentioned.

4.6 Activation Measures

In order to guard against inferences occurring at the time of testing, it is necessary to use a test that does not require subjects to evaluate the probe against the text. Activation measures are designed to do just that. If an inference has been drawn, then the activation level of the inference concept is elevated, thus facilitating lexical access. Thus tests involving lexical access can be used to assess whether an inference has been drawn (eg reading, naming, lexical decision, and modified Stroop).

Unfortunately, as Keenan, Potts, Golding, & Jennings (1990) point out, these tasks involve not just lexical access, but other processes as well. For example, Schustack, Ehrlich & Rayner (1987) have shown that single word reading time or gaze duration involves the time to carry out lexical access plus time to integrate the word's meaning with the preceding text. Therefore, if an inference has been drawn in an implicit version of a text, then gaze duration on the inference word will be shorter than in the control version. However, it is not possible to say which process caused this facilitation, lexical access priming or the post-access process of integrating the word with the text (which might be easier in the implicit version). In a sense, it does not matter as either process constitutes evidence for an inference having been drawn. If lexical access is being primed, then it is a forward inference; if post-access integration is facilitated, then it is a backward inference.

4.7 Lexical Decision

In some ways this is similar to recognition. The subject reads a passage in its implicit or explicit control version and then has to say whether a string of letters is a word or not. This involves two components: lexical access and the decision process. Since they have only to say whether the letter string is a word or not, subjects have no reason to compare the letter string to the text. This should, therefore, render any relatedness between the probe and text irrelevant. Since this could otherwise affect the decision process, then lexical decision would appear to be a good measure of priming and thus of inferencing.

Unfortunately, a number of studies have suggested that such a comparison nevertheless occurs (eg Balota & Chumbley, 1984; Chumbley & Balota, 1984; Forster, 1981; Neely, Keefe & Ross, 1989; Seidenberg, Waters, Sanders & Langer, 1984; West & Stanovich, 1982). This context checking occurs after lexical access in order to help the decision process; if the target is related to the text, this promotes the decision to say it is a word.

Studies by Keenan, Potts, Jennings & Golding (1988) and Potts, Keenan, & Golding (1988) suggest that the lexical decision process may overestimate the occurrence of inferences. They suggest that it may be possible to eliminate this context checking by forcing subjects to respond quickly through the use of a deadline procedure.

4.8 Naming

After reading the text, subjects are presented with the inference word and have to say it out loud. The naming process thus reflects the time for lexical access and the time for articulation. As there appears to be no connection between the post-access process (ie articulation) and the relatedness of the probe to the text, this method is sometimes viewed as a purer measure of lexical access. It has been shown that it is not affected by factors such as whether the target and prime are related or the presence of backward association between them (Seidenberg, Waters, Sanders, & Langer, 1984). Balota, Boland & Shields (1989) have shown semantic effects on articulation but Keenan, Potts, Golding, & Jennings (1990) feel these are spillover effects from lexical access.

One possible disadvantage of this method is that it may have a reduced sensitivity to detecting inferences compared with other methods (Potts, Keenan, & Golding, 1988).

Another potential disadvantage is that subjects may learn to anticipate what the probe is going to be, a type of expectancy priming (Neely & Keefe, 1989). However, Keenan, Potts, Golding, & Jennings (1990) asked subjects to try to guess what the probe might be, and under these optimum guessing conditions found that correct guesses were made only about 15% of the time. Nevertheless this may still affect the significance of the findings of any study which uses this approach..

4.9 Modified Stroop Task

The original Stroop test (Stroop, 1935) involved subjects naming the colour of ink in which a colour name such as red was printed. The modified Stroop task involves subjects reading a text and then naming the ink colour of a non-colour test probe. If the test probe has been activated or primed by the text, then it should take longer to name the ink colour used for that word (Conrad, 1974). For example, it should take longer to name the colour of ink for "doctor" following a text about hospitals than one about restaurants.

The modified Stroop task has been used to study forward inferences for the instruments of certain actions (Dosher & Corbett, 1982), instantiation of general terms (Whitney, 1986; Whitney & Kellas, 1984) and the likely consequences of events (Keenan et al, 1988).

As well as the priming of lexical access, the time to articulate the ink colour is also included in latency measures. Like naming, it is difficult to see how this can be affected by the relatedness of the probe to the text.

Since latency times tend to be longer than for naming (Keenan et al, 1988), the modified Stroop is more likely to avoid potential floor effects. Also, since the subject is focusing on ink colour, this should discourage expectancy priming.

A major advantage of this method is that it allows distinction between priming at the conceptual level and priming at the lexical level. In the example already given above, the latency response time is increased for a word like "doctor" following a text about hospitals (inference at conceptual level). However, if the probe "doctor" has been preceded by the explicit generation of the lexical form "doctor", then the opposite effect occurs, ie the ink colour of the probe is identified faster than if it had been preceded by a control version of the text (Dosher & Corbett, 1982; Whitney, 1986).

4.10 Movements of the Eyes

Perhaps one of the most powerful techniques to emerge in recent years has been the study of eye movements. Indeed, Rayner & Pollatsek (1989) claim that "eye movements are by far the best tool to understand the process of normal silent reading" (p115). This has been due to the refinement of the technology used to measure such movements and also to the accumulating bank of knowledge of what our eyes actually do when we perceive words printed on a page, the perceptual processes involved and what these perceptions tell the brain.

From a large pool of experimental data, it is now known that the eyes of a reader do not sweep continuously across a page of text (as one may intuitively think). Rather the eyes move in synchrony with one another (Rayner, 1978a) but move in a series of jumps or "saccades". The eyes come to rest on a word (called a "fixation") for a period of about 200 to 250 milliseconds, although this can vary anywhere from between 150 and 500 msecs). The eyes then move, usually to the right in English text, the duration of the saccade typically taking about 20 to 35 msec. It is known from several studies, eg Matin, 1974; Campbell & Wurtz, 1978; Wolverton & Zola, 1983) that input of visual information to the brain during a saccade is suppressed, or if any information does get in, it is of little practical importance. With each saccade the eyes generally move forward by about 7 to 9 characters.

While most of the saccades made are forward (ie towards the right of the page in English text), a small number, about 10 to 15 percent for fluent, adult readers, move backward in the text; these are called "regressions". Readers are usually unaware of these regressions. When the eyes near the end of a line of text, they move rapidly to near the beginning of the next line; this is called a return sweep. The return sweep usually starts about 5 to 7 character spaces from the end of the line and generally goes to somewhere between the third and seventh character space of the next line. This may land on the second word of the new line and often there is an additional short right-to-left saccade. These small saccades are probably corrections for errors in aiming the eyes. There are also slight movements of the eyes called microsaccades or drifts, small and rather slow movements caused by the less than perfect control of the oculomotor system by the nervous system. When this happens, there is often a small and rapid microsaccade to bring the eye back to where it was. Such drifts are usually considered to be noise and are ignored when eye movement data is analysed.

Various measures have been used when studying eye movements. Those generally agreed to be the most useful are:

(a) *first fixation duration* . This is the length of the first fixation, when the eye first lands on a word.

(b) *gaze duration.* The eyes may fixate a word and then move to another character within the same word. Gaze duration is the sum of these fixation times before the eyes move off to another word.

(c) *total viewing time.* This measure includes fixation times for first pass fixations as well as for regressive fixations.

(d) *saccade length.* This is the average number of character spaces within a saccade and gives an estimate of how far the eyes travel during each sweep across the text.

(e) *percentage of regressions.* This indicates how often the eyes look back from the current position to an earlier place in the text.

Although it can be said that most words are fixated during reading, some types of words will be fixated more often than other types. A distinction is often made between function words and content words. Content words include nouns, verbs, and adjectives; function words are prepositions, conjunctions, articles and pronouns. Adverbs can be classed either way depending on context and use. Fluent adult readers on average fixate about 90% of all content words but only about 30% of function words (Rayner, 1995).

Various factors may influence eye movements. For example, there is an optimal length of line which seems to facilitate the reading process (Tinker, 1963, 1965; Morrison and Inhoff, 1981). This appears to be about 52 characters. Since readers extract information from more than one word on a line during a fixation (McConkie & Rayner, 1975), this optimal length may be explained by a trade-off between extracting this parafoveal information and the difficulties experienced by making an accurate return sweep if the line is too long.

Orthographic factors also appear to influence how the eyes move. It has been found that for those languages which use different characters or symbols for the written forms of the language, the length of saccades varies greatly. For example, while readers of English typically have saccades of about 7 to 9 character spaces in length, readers of Hebrew have saccades of about 5.5 characters (Pollatsek, Bolozky, Well, & Rayner, 1981), readers of Japanese about 3.6 characters (Ikeda and Saida, 1978), and readers of Chinese about 2 characters (Stern, 1978). The important factor seems to be density of information contained within the text; as this increases so does the length of saccade decrease. There is a corresponding variation in average fixation times with longer fixations for the more densely packed text.

This pattern is also noticed within English when one examines different types of text (Rayner & Pollatsek, 1989), with more difficult text (eg a physics textbook) requiring longer fixations, shorter saccades, and more regressions than an easier piece of text (eg light fiction).

It is generally accepted that first pass fixations reflect primary processes such as lexical access. Lengthy gaze durations, say at the end of a sentence, are likely to reflect processes such as sentence wrap-up or places where the reader pauses slightly while making a bridging inference. Regressions and second pass fixations are thought to indicate places where confusion or ambiguities have arisen and the reader re-reads the text seeking clarification.

4.11 Summary:

The range of methodologies discussed above illustrates that, no matter which is adopted, there will be advantages as well as disadvantages to its use. The debate begins when attempts are made at defining exactly what constitutes an inference. Some researchers accept that an inference has been processed if the concept has simply been activated while others insist that the inference should be incorporated into the reader's long-term representation of the text (see Kintsch, 1988). The former would accept activation measures such as lexical decision or modified Stroop as being sensitive enough to detect if an inference had been processed while the latter would insist on using a memory probe.

It can be argued that any probe which is used after the text has been read, such as cued recall or sentence verification, could itself cause the inference to be made, rather than the inference being made on-line.

Even measures which are taken on-line, such as sentence reading times, rely for their efficacy on establishing the null-hypothesis. It is possible to set the target so that it contradicts the expected elaborative inference and so leads to a measurable increase in reading time. However, difficulties then arises with associative priming.

On-line question-answering seems attractive in that it allows the reader's developing representation of the text to be mapped. But again this method is intrusive and may cause an inference to be made which would not otherwise have taken place.

Activation measures theoretically do not require the reader to evaluate the probe against the text but rather rely on measuring ease of lexical access. However, even with lexical decision, there can still be a degree of context checking. This is reduced with naming but this procedure can give rise to expectancy priming. This in turn can be decreased by using the modified Stroop procedure.

However, perhaps the most effective protocol is that of eyetracking. The exact movements of the reader's eyes can be studied, how long particular words are fixated, which words are referred back to in earlier text, can all be precisely measured and their significance assessed. The process is also close to "normal" reading.

The current consensus is that there are trade-offs associated with each methodology, that no single protocol is adequate by itself, and that the wisest approach is to use multiple methodologies (Keenan, Golding, Potts, Jennings, & Aman, 1990).

4.12 Overview of the literature:

From the literature reviewed in the preceding three chapters, it can be seen that there are several facets of the reading process which cause difficulties for the deaf. While it seems that the deaf are equally as good as their hearing counterparts in visual analysis of text and other basic, bottom-up skills, they nonetheless have a poorer vocabulary. Not only do they know fewer words but they are also less likely to know alternative meanings for multi-meaning words.

Studies have shown that deaf subjects can and do use a variety of systems for recoding the words that they read on a page. The majority use visual, sign or fingerspelling modes, but a small number of the profoundly deaf use phonetic recoding and these seem to be the best readers.

The syntax of the English language is another major stumbling block for deaf readers. They do not hear language and find it difficult to assimilate the language structure. They are likely to use a manual form of communication such as British Sign Language, which has a totally different structure. For similar reasons, the deaf are likely to find figurative or idiomatic language rather difficult to grasp.

As a result of these difficulties, deaf students are much less likely to be able to stand back and view language as a tool, to experiment with language or to play language games. This is reflected in their rather impoverished written output.

Even when factors such as vocabulary and syntax are controlled for, the deaf still have more difficulty in comprehending text than hearing subjects. They seem to level off at a point where textual material becomes more implicit in content. The few studies which have examined the inferential skills of the deaf show that in this area they also experience difficulty. Not only do deaf children infer less well than hearing children, they do not seem to use their prior knowledge to construct appropriate schemata to help interpret the text.

In contrast, much work has been directed towards the inferencing process in hearing adults. Recent research has examined what type of inferences take place, when they take place, the role of the causal chain of the narrative, the effect of local text coherence, etc. Although many slightly different models have been proposed, and different methodologies have been used, there now seems to be a consensus about how and when hearing, fluent readers draw inferences from textual material. It seems likely that elaborative inferences are not drawn on-line unless the context is constraining enough to make an inference very likely. Otherwise readers will not make elaborative inferences as the consequences of doing so in terms of computational load would be too great and could lead to comprehension difficulties. Unless the context makes certain inferences more or less a certainty, they will not usually be made. The reader tends to adopt a "wait and see" approach.

The present set of studies will attempt to investigate the inferencing processes in a sample of severe to profoundly deaf children as they read textual materials, by adopting the general Sanford and Garrod model, which views the reading process from the stance of the subject employing background knowledge in the form of scenarios. A variety of approaches will be used in an attempt to avoid some of the methodological difficulties mentioned above.

Chapter Five:

Can deaf children draw inferences from simple textual material ?

5.1 Introduction

From earlier sections, it can be seen that there is a paucity of studies relating to inferencing skills in deaf children. Those which have been carried out offer at best inconclusive, and sometimes conflicting results. Given that hearing impaired children face so many additional difficulties when they come to read, can we be certain that it is the inferencing process which is failing them and not, for example, their limited vocabulary or the complexities of English syntax which is having the greater effect?

Many questions remain unanswered. For example, if deaf children can draw inferences on occasion, is the quality of these inferences commensurate with their level of competence in reading? Are the inference concepts activated briefly or are they incorporated into the reader's longer-term representation of the text?

Do the hearing impaired have difficulty with all types of inference or are some easier or more difficult than others, for example, considering the work of Zwaan and van Oostendorp (1993), can spatial inferences be made more easily than say causal?

Another important question which needs to be addressed is when any such inference might take place. Does it occur as the text is read or later during some retrieval process, ie does it occur on-line or off-line? If it occurs on-line, is the inference elaborative or bridging?

Since it would not be possible to address all of these issues in a single study, it was decided to initially investigate the inferencing abilities of deaf adolescents in relation to a group of hearing children matched for reading age and a group of hearing children who were nearer in chronological age. This allows inferencing skills to be examined in those deaf children who have acquired some basic literacy skills but who have reached a plateau in their progress, as described by Brooks, 1978; Reich & Reich, 1974; DiFrancesca, 1972. Using a group of younger children who approximate in reading ability to those deaf children being studied, allows direct comparison of inferencing skills across both groups. In other words, can the deaf children infer as well as hearing children with the same level of reading skills? The older group of hearing children are nearer in chronological age and should be more closely matched in terms of emotional maturity and experience to the hearing impaired children.

This initial study proposed to look at a large group of severely and profoundly hearing impaired children from a range of schools. It would include children who used different methods of communication (oral methods and total communication), children who were pre-lingually deafened as well as those who were deafened after the acquisition of speech, children whose parents might be deaf as well as those whose parents are hearing. The only provisos were that the children had to fall within the average range of ability, not suffer from any additional handicap and have an average hearing loss in the better ear of at least 70dB.

The study used simple material similar in style to that used by Wilson (1979). However, a different system of classification of inference types was used, spatial, causal, and temporal, this last being considered by several experienced teachers of the deaf to cause hearing impaired children particular difficulty. The material was particularly simple, being controlled in terms of vacabulary and syntax. Any inference which might be drawn by the subjects was in each case fairly straightforward.

Unlike Wilson's study, the present project included deaf children from a wide range of backgrounds, those who were prelingually deaf as well as those who lost their hearing after acquiring speech, those from an ethnic minority background, and not excluding any child whose parents were also deaf. Thus the present study gives a survey of a wider, less select population of deaf children.

5.2 METHOD

5.2.1 Subjects

One experimental group and two comparison groups were used. The experimental group comprised 39 deaf children (25 boys, 14 girls) with hearing losses of 70dB or greater in the better ear averaged over the four frequencies within the range 500-4KHz. Losses ranged from 75 dB to 116 dB, with a mean loss of 99 dB (standard deviation 11 dB). All were within the normal range of intellectual ability. Their most recently recorded IQ had to fall within the average range. Ages ranged from 12 years 2 months to 18 years 6 months with a mean age of 14 years and 5 months (standard deviation 16 months). A few of the children were from ethnic minority backgrounds. Both prelingually and postlingually deafened children were included in the group. All the children attended special schools for the deaf or units for the hearing impaired attached to mainstream secondaries in Strathclyde Region.

The first comparison group comprised 39 normally hearing children (18 boys, 21 girls) at the Primary 4 stage in a local primary school, aged from 8 years 0 months to 9 years 0 months, with a mean age of 8 years 6 months (standard deviation 4 months). The second comparison group contained 39 normally hearing children (15 boys, 24 girls) in the first year of a local secondary school, aged from 12 years 0 months to 13 years 1 month, with a mean age of 12 years 6 months (standard deviation 4 months). Since deaf children's reading skills tend to plateau at about the 8.5 year level, the younger control group could be expected to match the deaf children on reading achievement level, while the older control group, although slightly younger, would be approaching the deaf children's level of emotional and conceptual maturity.

5.2.2 Materials

The test materials consisted of four booklets. the first booklet contained 9 short pieces of text each of which was followed by a question designed to assess inferencing ability. These were preceded by four practice items. The second booklet contained cued recall items designed to measure the number of propositions correctly remembered from the text in the first booklet (see Appendix A1). The third bookletpresented a further 9 pieces of text with inference questions, while the fourth booklet contained cued recall material for these items. Three types of inference were examined, viz. spatial (could subjects infer the relative positions of objects), temporal (could they infer the time sequence of events), and causal (could they infer cause and effect). Examples of each type of inference are given below. The full list of text items is given in Appendix A1.

Temporal

The car crashed into the wall.
The car was badly smashed.
Later the policeman came.
Did the car crash before the policeman came?

Causal

The children were playing with a ball.
John threw the ball to Susan.
The ball broke a window.
Did John break the window?

Spatial

The bread was in the cupboard.
The butter was beside the bread.
The cheese was on the table.
Was the butter on the table?

The vocabulary was carefully controlled, having been selected from lists published by Bench and Bamford (1979), of words which should be known by deaf children similar to those in the experimental group. Bench and Bamford had recorded groups of deaf children orally telling stories based on a set of pictures. Thus if the deaf children spontaneously used a word, it could be taken that this word was indeed familiar to them. The groups of words were then categorised according to the children's degree of hearing loss and level of intelligence. Only those words used by children of average or below intelligence, and with an average hearing loss of 70 dB or greater were included in the texts of the present study. The syntax of the text was also carefully controlled bearing in mind the findings of Quigley, Wilbur, Power, Montanelli, & Steinkamp (1976), for example, nouns andnoun phrases were used in preference to pronouns wherever possible. The passages were scored for level of reading difficulty using the Fry graph (1977) which gave a difficulty level of 6 years, and the Spache formula (1953) which gave a difficulty level of 6.9 years, well within the capabilities of all three groups of subjects. The question types were randomly distributed throughout the booklets. Response requirements, ie affirmative or negative response, were also randomly distributed.

The aim was to make the reading of the texts within the capabilities of the deaf subjects so that any difficulties which they do face with comprehension can be ascribed to failures in drawing inferences and not to problems with other textual features such as syntax or vocabulary.

5.2.3 Design of Study

A 3x3 factorial repeated measures design was used. The first factor, between subjects or treatments, was subject group and had three levels - deaf, primary aged, and secondary aged. The second factor, a within subjects factor, was inference type and had three levels - spatial, temporal, and causal.

5.2.4 Procedure

The four booklets were administered to the subjects by their class teachers following a standardised protocol (see Appendix A2). For the first booklet the subjects were given two practice items with the answers given to illustrate what was required and a further two practice items which they had to attempt by themselves. They were also given two items to illustrate what was required with the recall test in the second booklet. Teachers were allowed to help with any difficulties which arose at this stage. Apart from this, no help or coaching was allowed. The booklets were presented in order, one booklet being removed before the next was presented. This ensured that on the cued recall items, subjects could not check back on the original texts. Subjects were allowed a maximum of one minute for each piece of text or recall item. For the questions in booklets one and three, subjects were expected to respond with a yes/no answer to the inference question but also to expand upon this response. In the recall items, booklets two and four, they were given the first sentence of each passage as a cue and had to complete the remainder of the story. The children were forewarned that they would be asked to recall what they had read after a few passages, but care was taken not to suggest either rote memory or gist memory. There were six passages for each inference type, making eighteen passages in total.

5.3 RESULTS

5.3.1 Inference

The accuracy of the inferences (ie the number of questions answered correctly) were recorded by totalling the number of correct responses for each inference type for each subject. These were then averaged for each group of subjects. The mean scores for the two comparison groups and the experimental group for each of the inferencing types are shown in Table 5.1 below. (Scores out of a possible maximum of 6):

TABLE 5.1: Inference Scores

	Spatial	Temporal	Causal
Deaf	4.5385 (76%)	3.6667 (61%)	4.4359 (74%)
Primary	5.5385 (92%)	5.4615 (91%)	5.4872 (91%)
Secondary	5.8462 (97%)	5.8205 (97%)	5.9487 (99%)

Thus while the secondary group are achieving a success rate of 98% over all inference types and the primary group 92%, the deaf group only make correct inferences in 70% of the texts.

5.3.1.1 First Analysis

These results were analysed using a 2 way, repeated measures analysis of variance, collapsing across questions and treating subjects as random variables, with inference type as within subject repeated measure and deaf, primary, and secondary group as between subject grouping factor. This revealed a significant main effect between the groups ($F_{2,114}$=55.390, p<0.001). Post-hoc Tukey tests showed that the deaf group was significantly poorer overall at inferencing than the primary or secondary groups and that there was no difference between the comparison groups (deaf= 4.2137, primary= 5.4957, secondary= 5.8718). There was also a significant interaction effect ($F_{4,228}$= 4.054, p<0.005) between group and inference type. A Tukey test found that all the deaf results were significantly poorer than all the comparison groups' results (p<0.01) and that there were no differences within or between the comparison groups. Within the deaf group, temporal inferencing was significantly poorer than the other two types (p<0.01) - see Table 5.1. The analysis of variance also revealed a significant main effect between inference types ($F_{2,228}$= 3.909, p<0.01), with post-hoc Tukey tests indicating that temporal items were poorer than either causal or spatial (p<0.01) - temporal= 4.9829, spatial= 5.3077 causal= 5.2906.

Clark (1973) discussed the methodological errors often made by researchers in psycholinguistics, paying particular attention to the "language-as-a-fixed-effect fallacy". Clark points out that in order to generalise one's findings beyond the specific example of language used in the experiment, one has to apply statistical analysis to the language materials themselves. In other words, the textual material itself cannot be regarded as a fixed entity. It is only one sample of many other possible samples of language and must, therefore, be regarded as a random effect. Clark recommends ideally the calculation of a so-called quasi F'-ratio, after Winer (1971), which takes account of the random effects of subjects and textual materials simultaneously. However, in practice, it is easier to calculate min F', the minimum possible value of the range of quasi-F'. If min F' is significant then so too must quasi-F' be. Thus the treatment effect can be confidently expected to generalise beyond the specific textual material used in any one experiment.

An analysis of variance was carried out, therefore, collapsing across subjects and treating texts as random variables. For ease of calculation, the raw scores were converted to percentages. This analysis confirmed the main effect between the groups ($F_{2,30}$= 85.034, p<0.01). A Tukey test confirmed as before that the deaf group was significantly poorer at inferencing than the primary and the secondary groups ($p<0.01$), but in addition revealed the primary group to be significantly poorer than the secondary group ($p<0.05$). The by-materials analysis of variance also confirmed the interaction effect between group and inference type ($F_{4,30}= 2.606$, $p<0.06$). Although this result just fails to reach statistical significance, comparisons of differences between means on a Tukey test revealed a similar pattern of results to that of the by-subjects analysis thus confirming those results, viz that all the deaf results were significantly poorer than all the comparison groups' results ($p<0.01$) and that there were no differences within or between the comparison groups. Also within the deaf group, temporal inferencing was significantly poorer than the other two types ($p<0.05$). The by-materials ANOVA did not support the main effect found for inference type in the by-subjects ANOVA ($F_{2,15}= 1.008$, $p>0.05$). This may have been caused by some factor within the materials themselves. For example, upon visual inspection of all the raw data, two questions in particular, one temporal and one spatial, did seem to produce lower than average results for all three groups of subjects.

Since variance measures the sum of the squares of deviations of scores from the mean of the distribution, it is particularly susceptible to outliers, that is scores which are judged to be inconsistent with the rest of the sample on some objective grounds (Collett & Lewis, 1976). Such spurious scores can distort the means of samples significantly. Other approaches use more robust measures of dispersion. One such method for detecting outliers is to regard any score which is more than one and a half times the interquartile range above the third quartile or below the first quartile as suspect (Lovie, 1986).

5.3.1.2 Second Analysis

Notched boxplots were drawn for the by-materials distribution of scores in each of the cells used in the ANOVA, eg Deaf-Spatial, etc. Using the criterion mentioned above, only one true outlier was detected, namely the third text in the spatial set. This was clearly an outlier for both the primary and secondary groups and on inspection, was the poorest text for the deaf group with only fifty percent correct responses on average, ie no better than chance. (The boxplot for the secondary group's causal items indicated a possible outlier, text 5, but on inspection this proved to be false. Since all scores in this section but one were 100%, the median, first centile, and third centile were represented on the boxplot by a single line. Therefore the score for text 5, a respectable 95%, or indeed any different score from 100% would fit the criterion defined above for detecting outliers. Also text 5 did not seem to cause any particular difficulties for the other two groups. Therefore this was not counted as an outlier.) Text 3 in the spatial section was removed and the remaining scores, collapsing across texts, were prorated for each subjects scores in all three groups. The distribution of subjects scores was then examined for outliers. Four scores in the deaf-spatial cell and two in the primary-temporal cell were replaced by the respective new cell averages.

These new sets of by-subjects and by-materials raw data were analysed as before, by ANOVA.

A significant main effect was found between groups in both the by-subjects analysis ($F_{2,114}$= 64.316, $p<0.001$) and the by-materials analysis ($F_{2,30}$=90.386, $p<0.001$), with a significant min $F'_{2,114}$= 37.577, $p<0.001$. Analysis by post-hoc Tukey tests revealed that the deaf group was significantly poorer than the two comparison groups ($p<0.01$) for both by-subjects and by-materials analyses. The post-hoc tests also revealed a borderline effect ($p<0.08$) in the by-subjects analysis and a significant difference ($p<0.01$) in the by-materials analysis, with the primary group performing more poorly than the secondary group. A significant main effect was discovered between types of inference in the by-subjects analysis ($F_{2,228}$= 19.55, $p<0.001$) and a borderline effect in the by-materials analysis ($F_{2,15}$= 3.615, $p<0.06$), with a borderline min $F'_{2,21}$= 3.051, $p<0.07$. Post-hoc Tukey tests failed to indicate a similar pattern of significance for mean differences in the by-subjects and by-material analyses. In the by-subject approach, spatial texts were significantly better than causal and temporal ($p<0.01$) and causal were significantly better than temporal ($p<0.05$).

However, in the by-materials approach, only spatial texts were significantly better than temporal ($p<0.05$). This apparent mis-match can be explained by the fact that the by-materials ANOVA revealed a difference for inference type of borderline significance, with a consequent borderline value for min F'. Therefore, according to Clark's (1973) arguments, the only result within the inference type area which could be reliably reproduced with different subject and language samples is that spatial inferences are significantly easier than temporal ones.

If the textual materials are examined closely, it is obvious that certain passages within the temporal grouping are causing particular difficulty, especially for the deaf children.

The third temporal passage appears to be more difficult than the others for the deaf and primary groups: Deaf 49% correct (average for all temporal items is 61%), primary 74% correct (average 91%). The passage is reproduced below:

The car crashed into the wall.
The car was badly smashed.
Later the policeman came.
Did the car crash before the policeman came?

Four of the five remaining passages conform to a similar pattern, viz. "Event A.... then Event B." The passage above breaks this pattern by using "later" to signal a time shift and thus a sequential order. It would seem that the deaf and primary children are susceptible to this change in form.

Additionally the deaf group found two further temporal texts particularly difficult, ie the second and sixth passages. These are noted below:

The milkman dropped the bottle.
The milkman got a brush and bucket.
Then the milkman cleaned up the mess.
Did the milknan clean up the mess before he dropped the bottle?

One possible explanation might be that nowadays milk is sold mostly in cartons, not bottles. The deaf children, with their limited life experience generally, may not realise that the bottle which was dropped contained milk which would make a mess. The mess, in their minds, could already exist and would thus be unconnected to the event of the bottle being dropped. Also more bottles are being made from plastic rather than glass nowadays that this may be an additional source of confusion for them, as such an item would not shatter when dropped.

Ann was walking along the street.
Ann slipped on some ice.
Ann fell down.
Did Ann fall down before she slipped on the ice?

Just over half the deaf children who took part in this study were educated in schools or units which used total communication (56%). As this system uses signing, it is possible that these children, at least in part, code the read text in signs (Odom, Blanton & McIntyre, 1970). Unfortunately the signs in BSL or Signed English for "slip" and "fall down" are rather similar in form and would tend to be merged into one movement thus veiling the distinct elements of the action. Emmorey & Lillo-Martin (1995, footnote 3, p 636) illustrate how more than one English word is sometimes needed to translate a single sign in American Sign Language.

The ANOVAs also revealed a significant interaction effect between subject group and inference type in both the by-subjects ($F_{2,228}$= 13.107, p<0.001) and the by-materials analyses ($F_{2,30}$= 4.033, p<0.05) with a significant min $F'_{4,51}$= 3.084, p<0.05. However, while there was overall agreement between the by-subjects and by-materials post-hoc comparisons for the majority of the means, one significant item in the by-subjects comparisons (deaf spatial significantly poorer than primary spatial) was not supported in the by-materials, and one significant item in the by-materials (deaf spatial significantly better than deaf causal) was not supported in the by-subjects comparisons. Out of 36 possible comparisons, it can be seen that these mis-matches represent only a small percentage of the total. For the reasons given above (Clark, op cit), these two comparisons could not be considered to offer reliable evidence of true differences. However, from the other comparisons, deaf temporal and causal inferences were significantly poorer than all types of inference for the two control groups. Deaf spatial inferences were significantly poorer than all types of inference for the secondary aged control group but there was no difference compared to all types of inference made by the primary aged group. Finally, within the deaf group itself, temporal inferences proved to be significantly poorer than spatial or causal, with no significant difference between the last two.

The new mean scores after removing outliers for each group on each inference type are shown below in Table 5.2:

TABLE 5.2: Inference Scores

	Spatial	Temporal	Causal
Deaf	5.2564 (88%)	3.6667 (61%)	4.4359 (74%)
Primary	5.7436 (96%)	5.6667 (94%)	5.4872 (91%)
Secondary	5.9744 (100%)	5.8205 (97%)	5.9487 (99%)

In summary, the second analysis clarifies uncertainties in the first analysis. The by-subjects ANOVA in the first analysis was not fully matched in the by-materials ANOVA on the two main effects, ie between group differences and differences for inference type. In the second analysis these two by-subject main effects are fully matched in the by-materials ANOVA. It is now clear that the deaf group performs significantly more poorly than both comparison groups and that the primary group does more poorly than the secondary. Also temporal items are more difficult overall than the other two types of inference.

The by-subjects interaction effect in the original analysis was matched on the by-materials and showed that the deaf group were poorer at inferencing than both hearing groups but also that the deaf found temporal inferencing more difficult than the other two types. The second analysis again revealed a matched by-subjects/by-materials interaction effect with almost the same results. Again the deaf group do more poorly over all inference types than the secondary group. The deaf also find temporal inferencing more difficult than causal or spatial. However, while the deaf do more poorly than the primary subjects on temporal and causal inferencing, there is no significant difference between the deaf's spatial inferencing and all types of inferencing by the primary group.

5.3.2 Recall

Booklets 2 and 4 were marked for the number of propositions correctly recalled by each subject. All marking was by the author and was blind, ie the scorer did not know which group the subject belonged to, nor if the subject was male or female.A proposition was defined as a basic idea within the text which was distinct and separate and added some new piece of information. This is illustrated below for the examples already given in the Materials section. Since the first sentence is given in each case as a cue, it is not included in the scoring:

Temporal

The car crashed into the wall.
The car / (was) badly / smashed.
Later / the policeman / came.

Causal

The children were playing with a ball.
John / threw / the ball / to / Susan.
The ball / broke / a window.

Spatial

The bread was in the cupboard.
The butter / was beside / the bread.
The cheese / was on / the table.

To be scored as correct, recalled material had to include the basic content of each proposition. For example, in the temporal item given above, incorrect tense of the auxiliary verb would not be counted as wrong; thus "was smashed" and "is smashed" would both be considered correct. The fact that the auxiliary verb is separated from the verb proper did not mean that it was scored as a separately. However, a synonym would be counted as incorrect eg "mangled" for "smashed". Prepositions had to be correct. For example, in the causal item illustrated above, "John threw the ball to Susan" was acceptable, but "John threw the ball at Susan" was not acceptable. Propositions which were in the original text but which were recalled out of sequence or in isolation were also scored as being correct. A few examples of how recalled material was scored is shown below:

EXAMPLE 1: The car / smashed / very / bad.
(PRIMARY) Then / the police man / came.

The word "bad" rather than "badly" is scored as correct. "Then" rather than "Later" is counted as an error. The subject is not penalised for splitting "policeman" into "police man". The score given was 5 correct propositions and 2 errors.

EXAMPLE 2: The car / was badly / damaged.
(PRIMARY) Later / on / the policeman / came.

"Damaged" is rated as an error as is the addition of "on". The score given was 5 propositions and 2 errors.

EXAMPLE 3: The cup / was on / the table.
(PRIMARY) The butter / was beside / the bread.

"The cup" is scored as an error. The fact that the sentences are in the incorrect order is ignored. The score given was 5 propositions and 1 error.

EXAMPLE.4: John / threw / at / Susan.
(DEAF) The window / was / bread.

The substitution of "at" for "to" is counted as an error. The fact that "window" is now the subject rather than the object of the sentence is not regarded as a mistake. "Was" and "bread" are obviously wrong. Thescore given was 4 propositions and 3 errors.

EXAMPLE 5: The cheek / on / the table.
(DEAF) The bread / was on / the top of / cupboard.

The mistaken use of "cheek" for "cheese" is rated an error. The second use of "(was) on" when "on" has already been used in the previous sentence is also an error. (Only one "on" appears in the original text.) The score given was 2 correct propositions and 5 errors.

Since the total number of propositions contained in the texts was different for each inference type, the number recalled by each subject in each category was converted to a percentage so that direct comparison could be made for each subject group across inference type. The mean scores for the percentage of propositions correctly recalled for each group is shown below in Table 5.3:

TABLE 5.3: Percentage of propositions correctly recalled.

	Spatial	Temporal	Causal
Deaf	50.68	43.85	49.56
Primary	53.74	45.95	49.90
Secondary	79.23	75.85	74.49

5.3.3 Inter-rater reliability

Since the partitioning of the material recalled by subjects into propositional units calls for some degree of subjective judgement, a second scorer was used to rate a sample of the responses (all the responses of 10 deaf subjects and 5 subjects from each of the control groups chosen at random). This second scorer was given general guidelines as to how the propositions were defined in the original scoring (similar to that noted in the early paragraphs of the Recall section above) but was left to define the exact details of her own scoring criteria for each passage. For example, it was noted that auxiliary verbs were not counted separately from the main verb; thus "were playing" is counted as one propositional unit. Also a preposition and the verb to be were counted as one unit in phrases such as "was on". Definite and indefinite articles were not counted as separate units. Adverbial prepositions which could take alternative forms on recall were counted as separate units, for example, "John / threw / the ball / to / Susan" not " at ". The second scorer's marking criteria corresponded closely with those of the original scorer. The second scorer's results were then compared with the original results for these selected subjects (Pearson's product moment correlation): $r=0.9896$, $df=18$, $p<0.01$. Because of the spread of scores, especially for the deaf group, it was necessary to compare the second scorer's results with the original scorer's results for exactly those same subjects. Therefore, the original scoring would appear to be reliable.

5.3.4.1 First Analysis (of recalled material)

These results were analysed using a 2 way, repeated measures analysis of variance, collapsing across texts and treating subjects as random variables, with passage type as within subject repeated measure and subject group as a between subject factor. This revealed a significant main effect between the groups ($F_{2,11}$= 45.489, $p<0.001$). Post-hoc Tukey tests revealed that the secondary group recalled significantly more propositions (76.52%) than either the deaf group (48.03%) or the primary group (49.86%) at the $p<0.01$ level. There was no significant difference between the primary subjects and the deaf subjects. The analysis of variance also revealed a significant main effect between inference types ($F_{2,228}$= 13.356, $p<0.001$) with Tukey tests indicating that temporal items were significantly poorer than causal ($p<0.05$) and spatial ($p<0.01$), and causal significantly poorer than spatial ($p<0.05$) [temporal: 55.21%, causal: 57.98%, spatial: 61.21%]. The interaction of group membership by inference type failed to reach significance ($F_{4,228}$= 1.854, $p>0.05$). In other words, all groups showed the same pattern of recall results by inference type although the secondary pupils recall more.

An analysis of variance, collapsing across subjects and treating texts as random variables, confirmed the main effect between the groups ($F_{2,30}$= 77.403, $p<0.001$) and yields a significant min $F'_{2,124}$= 28.651, $p<0.001$. Tukey tests on the by-materials data disclosed a similar pattern to the by-subject analysis, thus confirming that the secondary group recalled significantly more text propositions than the other two groups ($p<0.01$) with no difference between the primary and deaf children [secondary: 77.22%, primary: 51.33%, deaf: 49.72%]. The significant main effect on passage type in the by-subject analysis was not supported in the by-materials analysis ($F_{2,15}$= 0.960, $p>0.05$).

5.3.4.2 Second Analysis (of recalled material)

Since the by-materials analysis did not support the by-subjects finding of a significant difference on inference type, the procedure used in the previous section to remove outliers was utilised. From inspection, only the temporal section contained any outliers, text 5 for the primary group and text 2 for the secondary group. Since text 3 in the spatial section was considered to be inappropriate when analysing the accuracy scores, it was also omitted for all three groups in the present analysis. The empty values created by omitting these outliers were filled with the new average for each cell (eg primary temporal). With the outlier texts removed, the subjects' scores were re-averaged, collapsing across the remaining texts and the distributions of subject scores examined for outliers, using the same method as above. Out of a total of 351 subject scores, only 10 were omitted as outliers, viz. one in the spatial and one in the temporal section for the primary group, and two in the spatial, three in the temporal and three in the causal sections for the secondary group.

The new percentage of propositions correctly recalled for each group after the omission of outliers is shown below in Table 5.4:

TABLE 5.4: New percentage of propositions correctly recalled.

	Spatial	Temporal	Causal
Deaf	49.28	43.85	49.56
Primary	54.74	42.51	49.90
Secondary	76.21	79.72	76.90

These new arrays of raw data were again analysed using ANOVA. This revealed a significant main effect between the groups in the by-subjects analysis ($F_{2,114}$= 63.114, p<0.001) and in the by-materials analysis ($F_{2,30}$= 85.157, p<0.001), with a significant min $F'_{2,112}$= 36.249 p<0.001. As before, post-hoc Tukey tests on both the by-subjects and the by-materials data showed that the deaf and primary groups were significantly poorer than the secondary group (p<0.01) and that there was no difference between the deaf and the primary groups. A significant main effect between inference types was again found in the by-subjects analysis ($F_{2,228}$= 8.360, p<0.001) but this was not reproduced in the by-materials analysis ($F_{2,15}$= 1.590, p>0.05; min $F'_{2,21}$= 1.336, p>0.05).

A new significant interaction between group and inference type was found on the by-subjects analysis ($F_{4,228}$= 7.956, p<0.001) but this was not reproduced in the by-materials analysis ($F_{4,30}$= 2.168, P>0.10; min$F'_{2,48}$= 1.704, p>0.05).

Thus the removal of outliers from the arrays of raw data has not clarified the analysis. A visual inspection of the test material was made to see if any particular text may have caused the pattern of results obtained, eg were some temporal texts especially difficult, were some spatial texts easier than the others etc. Unfortunately, no obvious pattern emerged; the temporal text which was the most difficult for the deaf group was different from the one which was most difficult for the controls. Similarly, the easiest spatial passage varied across the groups. This variation would have been allowed for to some extent in any case by the removal of outliers.

In summary, it can be stated that the older hearing subjects recall significantly more propositions than the younger hearing subjects or the deaf children, with no difference being found between the last two.

One possible explanation for the results obtained in these analyses is that what is being measured is, in fact, the total amount of material recalled, regardless of whether it comes from texts whose questions have been answered correctly or not. If some relationship exists between accuracy of answering text questions (ie comprehension of the text) and correct recall of the material of the text, then the results could be re-analysed in the light of this finding.

5.3.5 Accuracy of inferencing and amount recalled

To examine whether any relationship existed between the number of propositions recalled and how well inferences were drawn from the text, a Pearson product moment correlation was carried out. The number of correct inferences made, arranged over all inference types, correlated positively with the number of propositions in the text which were correctly recalled for both the deaf and the primary group. However, the secondary subjects did not show this correlation, perhaps because they are skilled at inferencing and have reached a ceiling with the relatively simple materials used in this study. This is obvious when one considers that for the secondary group, the percentage of material recalled over all three types of inference ranged from 45 to 92 with a mean of 77 and standard deviation of 10. The other two groups by comparison have lower amounts recalled and the distribution of scores is much more scattered: deaf scores range from 10% to 79% (mean of 48, standard deviation of 19), primary group scores from 12% to 83% (mean of 50, standard deviation of 14). The correlations are shown below in Table 5.5

TABLE 5.5: Correlation between inference and recall.

Deaf:	$r=0.6875$,	$p<0.01$,	$df=37$	$r^2=0.473$
Primary:	$r=0.4726$,	$p<0.01$,	$df=37$	$r^2=0.223$
Secondary:	$r=0.1238$,	$p=ns$,	$df=37$	$r^2=0.015$

This pattern is also found when each type of inference is examined separately, with the exception that for the primary group there is no significant correlation for temporal items. The data are shown in Table 5.6 below:

TABLE 5.6: Correlation matrix by inference type.

	Spatial	Temporal	Causal	
Deaf Group	r=0.6161	r=0.3791	r=0.6925	df=37
	p<0.01	p<0.05	p<0.01	for all
Prim Group	r=0.4219	r=0.2141	r=0.3403	correlation
	p<0.01	p=ns	p<0.05	
Sec. Group	r=0.2290	r=0.1359	r=0.0444	
	p=ns	p=ns	p=ns	

5.3.6 Re-analysis of Recalled Material

Since it has been shown, therefore, that a direct relationship exists between the amount of material recalled and the number of correct inferences made, at least for the deaf and primary groups, it would seem wise to re-analyse the data for percentages recalled over inference type for the deaf and primary groups, but only for those questions which have been answered correctly.

The mean scores for the percentage of propositions correctly recalled for those questions which have been correctly answered are shown below in Table 5.7:

TABLE 5.7: Percentage of propositions correctly recalled for those questions correctly answered.

	Spatial	Temporal	Causal
Deaf	51.97	48.72	51.97
Primary	53.59	48.05	48.72

These results were analysed as before using a 2 way, repeated measures analysis of variance, collapsing across texts and treating subjects as random variables, with passage type as within subject repeated measure and subject group as a between subject factor. This failed to show any significant main effects or interactions. The F-ratios and relevant means are shown below:

Deaf: 50.89	Primary: 50.12		$F_{1,76}$= 0.038, $p>0.05$
Spatial: 52.78	Temporal: 48.39	Causal: 50.35	$F_{2,152}$= 2.645, $p>0.05$
Group x inference type interaction:			$F_{2,152}$= 0.810, $p>0.05$

An analysis of variance, collapsing across subjects and treating texts as random variables, also failed to show any significant main effects or interaction (means and F-ratios below).

Deaf: 54.17	Primary: 51.22		$F_{1,15}$= 0.950, $p>0.05$
Spatial: 55.25	Temporal: 50.08	Causal: 52.75	$F_{2,15}$=0.468, $p>0.05$
Group x inference type interaction:			$F_{2,15}$= 0.204, $p>0.05$

In other words, the deaf children do not appear to be performing any differently from the primary children. For those passages where a question was answered correctly, both groups subsequently recall similar amounts of information. Also, for both groups, no type of inference is any better or any worse than any other.

The deaf and younger hearing group recall a roughly similar number of propositions (approximately 48% and 50% respectively). However, the primary subjects infer correctly in about 94% of the texts whereas the deaf only infer correctly about 74% of the time. Therefore, since the deaf appear to recall a disproportionate number of propositions overall in relation to their level of comprehension, it would seem that they must be recalling information correctly in some cases where comprehension has failed.

It could be argued that recall of propositional content or surface structure is a poor measure of degree of comprehension (Johnson-Laird & Stevenson, 1970). It is also true that the recall task in the present study occurred after subjects had read several texts. However, subjects were asked to elaborate on their answer and not to simply respond with a "Yes" or "No" immediately after they read each text. Examination of these responses for the deaf subjects showed that when they responded correctly to the question, in most cases they also appeared to genuinely understand what they had read. This applied over all inference types.

5.3.7 Accuracy of inferencing and nature of recalled material

Several studies, eg Anderson, Pichert, Goetz, Schollert, Stevens, & Trollip (1976), Garnham (1979), Brewer (1975), and Spiro (1977), have all shown that fluent non-deaf readers do not recall accurately exactly what they have read. They tend to recall the completed inference and not the implied version presented to them originally.

With these findings in mind, the recalled material for each subject in which the inference question was answered correctly was classified as either containing the completed inference or not. Paraphrasing was allowed so long as the inference was completed. For example: "Mary could not find Tom" was just as acceptable as "Mary could not see Tom" in one of the Spatial texts used (see Appendix A1).

The percentages of cases were the completed inference was made and those were it was not made are shown averaged over all inference types for each of the three groups in Table 5.8 below:

TABLE 5.8: Completed and uncompleted inferences in recalled material.

	Inference completed	Inference not completed
Deaf Group	13.6%	86.4%
Primary Group	17.0%	83.0%
Secondary Group	5.7%	94.3%

Analysis of these results by t-tests reveals that significantly more texts are remembered in the non-completed inference form than in the completed form by all groups (Table 5.9):

TABLE 5.9: Completed Vs Incompleted Inferences.

Deaf Group	$t=14.2876$	$df=38$	$p<0.001$
Primary Group	$t=18.5627$	$df=38$	$p<0.001$
Secondary Group	$t=34.9858$	$df=38$	$p<0.001$

Similar results were found when each inference type is analysed separately. These are shown in Table 5.10 below:

TABLE 5.10: t-test results for each inference type.

Deaf Group	Spatial	t=8.6600	df=38	p<0.001
	Temporal	t=11.1050	df=38	p<0.001
	Causal	t=14.2876	df=38	p<0.001
Primary Group	Spatial	t=7.1241	df=38	p<0.001
	Temporal	t=25.0299	df=38	p<0.001
	Causal	t=9.2516	df=38	p≤0.001
Secondary Group	Spatial	t=21.0748	df=38	p<0.001
	Temporal	t=74.3506	df=38	p<0.001
	Causal	t=23.8777	df=38	p<0.001

If the percentage of those questions which were answered correctly and recalled in the completed inference form are compared for each group, no significant difference is found (Chi-squared=5.555, df=2, p>0.05). Thus the deaf are recalling the textual material in a similar manner to the hearing children.

5.4 DISCUSSION

From the accuracy results, it can be seen first and foremost that the deaf are significantly less accurate at drawing inferences than either hearing children matched on reading achievement age or hearing children nearer in chronological age (70% accuracy compared with 92% and 98% for the primary and secondary groups respectively). There was a greater spread of scores for the deaf group, ranging from 6 to 18 correct responses (standard deviation of 3.32), compared with a range of 9 to 18 correct for the primary group (standard deviation of 1.76), and 16 to 18 for the secondary group (standard deviation of 0.54). It should be noted that some deaf subjects scored as well as the best of both the comparison groups and indeed better than the lowest scorers of these groupls. In general, however, these results confirm the findings of Wilson (1979) and Davey, LaSasso & McReady (1983) that deaf children find difficulty in drawing inferences from textual material, even when the material has been designed for maximum ease of reading.

The results also indicated a trend where the younger primary aged children performed more poorly than the older secondary aged children. This would confirm the findings of many studies that inferencing ability is a developmental skill that improves with age, eg Paris & Mahoney (1974), Stein & Glenn (1979), Paris & Carter (1973), and Paris & Upton (1976).

It was also found that, as well as being poorer generally at inferencing, the deaf found some kinds of temporal inferencing significantly more difficult than either causal or spatial inferencing. There may be good reasons why this is the case and these issues have already been discussed in the Results section. It certainly shows that particular care must be exercised when constructing materials for the deaf to read as several extraneous factors may have a bearing on the results obtained. However, the present results would seem to confirm the findings of Ormanson, Warren & Trabasso (1978) that inferencing ability increases with age since the primary group perform less well than the older children in the secondary group. However, Ormanson et al (op cit) found that no particular form of inferencing causes any special problems whereas in the present study it is found that spatial items are found to be easier than other forms of inferencing by all groups of children. Wilson (1979) did find that some types of inference were more or less difficult than other types for both deaf and hearing subjects. In Wilson's case, however, different types of inference were examined than in the present study.

In the present study, spatial texts were answered as accurately by deaf subjects as by hearing subjectss matched on reading age, but the deaf found temporal items particularly difficult.

One reason why deaf readers find spatial inferences easiest might be that they develop an expectancy that certain passages will contain a spatially formulated storyline, and therefore, a spatial inference will be required. If subjects can anticipate in this way, then they could orientate themselves and, according to Zwaan and van Oostendorp (1993), create a spatial representation, so helping them with the inference. A quick inspection of the texts used, however, shows that there is no such obvious pattern discernable from the first sentences of the stories.

It is possible that temporal inferencing causes particular difficulty for the deaf simply because the questions for these texts generally contain more propositions than either causal or spatial types and are therefore more difficult to retain in short term memory or require more reasoning to figure out the relationship between two propositions than to comprehend a single proposition. This is unavoidable when asking such questions as the two incidents have to be juxtaposed; did event A occur before event B ? The question of load imposed on short-term memory by temporal items is considered later when the evidence from the recall data is examined.

Alternatively it might be that the deaf find it more difficult to cope with the temporal nature of the information, because they have been deprived of auditory input which in itself is necessarily in a temporal form compared with visual input which would generally be considered to be both temporal and simultaneous in nature. In other words, the deaf might have had less practice handling temporally sequenced information than hearing children. Blair (1957) found that deaf children performed as well or better than hearing children on tests of spatial memory but poorer on tests of temporal-sequential memory. The finding from the present study that deaf children's spatial inferencing is no different from any type of inferencing by the primary group would appear to lend strength to this argument, as the deaf obviously operate mostly through the visual modality, and thus having such experience can compete effectively with their hearing counterparts.

A third possible explanation arises later when consideration is given to the form in which the material is recalled. It should be noted that both these findings are observed even with very simple textual material, controlled for both vocabulary and syntax. It is conceivable, however, that despite this attempt at control, the syntactic structure of the temporal texts was beyond the ability of many of the deaf subjects.

The deaf and the primary aged subjects show no significant difference in the percentage of correct propositions recalled but both groups remember significantly fewer propositions than the secondary group. This applies over all inference types. This could be because the secondary children have reached a ceiling and presumably the reading and inferencing are automatic for them, therefore they can concentrate more on remembering what they read, whereas for the primary and the deaf groups, the reading and inferencing tasks make greater cognitive demands and therefore less is remembered.

For the deaf and the primary groups the number of propositions that can be recalled about a text is related to whether or not the inference has been understood, and since the correlation co-efficient is larger for the deaf group than for the primary, this connection would appear to be operating more strongly for the deaf. Therefore when the recall data is re-analysed only for those items which have been inferred correctly (ie comprehended), we find that there is no difference in the amount recalled by the deaf or primary groups. In other words, both groups recall the same amount of information proportionately when they comprehend the textual material. The comprehension question responses may on some occasions underestimate the deaf readers' ability. They got something from the passage even when they got the question wrong.

Following on from this, it would seem unlikely that the first explanation as to why deaf children find temporal inferences most difficult, is valid. Since no difference is found between the deaf and primary groups or between any inference type on the amount of information recalled when a text is correctly comprehended, memory overload caused by lengthy temporal texts cannot be offered by way of explanation.

A third explanation to possibly explain why temporal inferencing appears to present so much difficulty for deaf subjects is now discussed. First it is necessary to consider the manner in which textual information is recalled and how this relates to what is known from other studies on this topic. Are the deaf children and the younger comparison group assimilating the information contained in the texts into their pre-existing cognitive schemata when they comprehend successfully (as judged by whether a question is answered correctly or not)? If this were so, then there might be more inaccuracies on recall as the textual material becomes absorbed into a pre-existing schema, with a resulting altering of both schema and textual material. Such a view would be consistent with the findings of, for example, Anderson et al (1976), Garnham (1979), Brewer (1975), and Spiro (1977) who found that for fluent, non-deaf readers it is the completed inference which is recalled and not the original implicit version. However, from the present study it is obvious that none of the groups is following this pattern. For each group and for all inference types, what is recalled is not the completed inference but the non-inferred version. With the secondary group, recall was generally fairly accurate, either verbatim or a close paraphrase on the original. It is possible that because these subjects manage accurate recall, they have no need to infer. The deaf also recalled in this fashion but to a much lesser extent, more frequently introducing a few inaccuracies, or in some cases failing to recall any information for some texts. The primary children could also recall verbatim or in paraphrase, but not nearly as well as the older hearing children. It was this younger group who tended to invent plausible but totally inaccurate stories when their memory for the original text failed, along the lines proposed by Bartlett (1932).

Spiro (1975) has argued that the demand characteristics of an experimental situation, the nature of the materials used, and the relatively isolated context in which discourses are presented all increase the likelihood that subjects will tend to isolate the read material from pre-existing cognitive structures and treat them purely as an exercise somewhat removed from a real life situation. Obviously some interaction with cognitive structures must go on, but Spiro contends that this will be minimal. This may be what is happening in the present experiment. The materials used were rather short and simple in construction (necessarily so to reduce syntax and vocabulary difficulties for the deaf). From the way the temporal items were recalled, it was noted that the secondary hearing children tended to recall almost verbatim, eg "James finished his apple. Then he put his book down" The deaf group also recalled temporal items in a similar manner. The primary hearing children too recalled in this form, but occasionally might recall as follows: "When he had eaten the apple, James put his book down". This is close to what would be expected for recall in the inferred form: "After he had eaten the apple, James put his book down". Therefore it is possible that the temporal items have been presented in the most acceptable form within the text, ie event A then event B. So that the form used in the questions, the completed inference form, ie event A before/after event B, is less acceptable to the subjects of all groups. Brewer (1975), in a study of synonym substitution in recalled material, found that substitution tended to take place in one direction only. For example, the sentence: "The nightgown was too *little*" was always recalled as "The nightgown was too *small*" and never in the reverse direction. These favoured words were rated by subjects as being more "natural". Therefore, in the present experiment it is possible that something similar is happening and that the temporal items are already in the most "natural" form for the subjects as they are presented in the body of the text, with the completed, inferred form used in the question being unnatural.

In summary, the present study has confirmed the findings of previous authors, namely that the deaf find it difficult to draw inferences from textual material, even compared with younger children matched on reading age. The wide variation in reading skill noted within the deaf group is also predictable from the literature.

The relative efficiency or otherwise of processing different types of inferred material by the deaf is less clear. The disparate types of inference which have been used throughout the literature have failed, unsurprisingly, to show any homogeneous pattern. In the present study, it is found that one type, spatial, is processed by the deaf as efficiently as all inference types by the younger comparison group. This could be because of the experience of the deaf in using a limited range of modalities to communicate hampering temporal and possibly causal inferencing, but their "normal" spatial experience enhancing their performance with spatial items.

Some types of temporal items are particularly poor for the deaf. This would seem to be because the deaf are especially sensitive to the form in which such texts are presented.

For those texts which are comprehend correctly, the deaf can recall a similar amount of information as the primary group. They also recall in a similar style, without completing the inference. This is probably due to the simplistic nature of the material used.

One advantage of the present design is that the process undertaken by subjects is akin to "normal" reading, in that they are presented with the complete text on the page, can read at their own pace, and can look back to earlier parts of the text if necessary to clarify any ambiguities.

While this study offers some reassuring findings, it also presents some unanswered questions. For example, do the deaf and younger hearing children infer in similar ways; do they assimilate textual information into their cognitive schemata? Because a memory probe is used, it is not possible to say whether inferences are occuring on-line or off. Neither is it possible to distinguish between bridging and elaboration.

Further studies are required using textual materials which are more challenging, while bearing in mind the limiting aspects of vocabulary and syntax on deaf reading skills. Temporal items in particular would need to be presented in an alternative form which would eliminate some of the complicating factors discussed above. In the present study, the probe question is presented after the text has been read. As mentioned in the chapter on methodological difficulties, it is possible that the probe itself could cause the inference to be drawn. To examine these issues further, it will be necessary to use on-line procedures in the next study.

Chapter Six:

Do deaf children draw inferences at time of reading?

6.1 Introduction

From the results of the first study, it can be seen that deaf children have more difficulty in comprehending inferred material than their hearing counterparts. In order to examine whether or not deaf children infer in a similar manner to hearing children, on-line methods are needed. In other words, a method is needed which will determine whether deaf children draw inferences as they read as opposed making the inference at a later stage, for example, during a subsequent retrieval task. It was decided that methods similar to those used by Sanford & Garrod (1981) should be employed. In other words, by using a reaction time method it would be possible to estimate how easy or difficult any particular sentence is to comprehend by calculating the time subjects take to read it. The actual process is described more fully in the Method section below.

If we consider the theory outlined by Sanford & Garrod (1981), and the empirical evidence which they use to substantiate their speculations, then it is possible to devise a procedure which will confirm or refute the hypothesis of the present study that deaf children use different cognitive processes when attempting to comprehend a piece of discourse than do hearing children. It should be noted that Sanford & Garrod's empirical investigations and the studies of most of the other researchers in this area have concentrated on adult subjects who are fluent readers. It would remain to be seen, therefore, whether the younger subjects of the control group did indeed process text in a similar manner to the adult subjects used in these other studies.

In their theory of how discourse is processed, Sanford & Garrod (1981) describe how readers use foregrounding and topicalisation as well as syntactic information to resolve pronominal reference to entities mentioned explicitly in the text. They also propose a mechanism, founded on much empirical evidence, of how readers use what they describe as "scenarios" to bring background knowledge to help interpret material which required inferences to be drawn for full comprehension. As the text is read an appropriate scenario is invoked by the reader and this provides an extended domain of reference which can be used for direct mapping of subsequently mentioned entities in the text onto slots in that scenario. These slots provide roles and the necessary connecting relationships between them, giving a programmatic component to the scenario. The role slots also bring a certain predictability of behaviour for the characters they represent. The main characters are so closely bound in with the scenario that if the scenario fades from current "focus", they too fade. Secondary characters which are less strongly bound in with the scenario can be more highly foregrounded because of their unexpectedness in that particular scenario, (Anderson, Garrod & Sanford 1983). Unusual or unexpected adjectival attributes also serve to foreground characters. Scenarios do change with appropriate cues in the text, the old scenario fading from current focus. If a scenario fails to serve as an adequate mechanism for interpretation of the text, then secondary processing takes place and a search is made for a more appropriate scenario.

Garrod and Sanford (1980) presented evidence in support of their model. They had subjects read short passages, each of which was introduced by a title. If a suitable title is used, then an appropriate scenario should be invoked by the reader; some other title would invoke an inappropriate scenario. For example, if the title "*In Court*" is used, then subsequent reference to "a lawyer" should cause no particular difficulty as this entity can be mapped directly onto an appropriate slot in that scenario. A more general title such as "*Telling a Lie*" would not be expected to invoke a courtroom scenario and any subsequent reference to "a lawyer" is likely to lead to time-consuming bridging processes. Indeed Garrod and Sanford found such an increase in reading times for those texts where the title was inappropriate.

A particular scenario can become redundant as the reader proceeds through a text. For example, the action can shift to another location or there can be a time change, either to later or earlier than the current situation. In such a case the original scenario will have become inappropriate and would make interpretation more difficult. Thus a scenario which involved having dinner in a restaurant would be expected to include a default slot for "waiter". No additional time is needed when such a person is mentioned for the first time in this context. However, a sudden temporal and spatial shift to, say, the journey home afterwards, would lead to difficulty in resolving a reference to a "taxi driver". In these circumstances the appropriate linguistic cues usually result in the reader fairly rapidly changing the scenario to one that is more appropriate. However the sentence which contains the signal for a scenario shift will be read more slowly (Anderson, 1981).

Scenarios, as used by Sanford and Garrod, refer to stereotypical events such as going to a restaurant, appearing in court, etc. The restaurant scenario is invoked by an appropriate title and provides default information about waiters, tables, menus and so on. This information makes it possible for the reader to draw inferences on-line. However, other researchers see the reader as using more general world knowledge to make elaborative inferences which embellish the text representation, eg Trabasso & van den Broek (1985), Trabasso & Suh (1993), Long & Golding (1993), van den Broek & Lorch (1993). The reader goes beyond the stereotypical but still uses generally accepted world knowledge to help interpret the text. The reader constructs a situational model derived from this world knowledge rather than a scenario in the Sanford & Garrod sense.

In the present study, a series of texts were presented in two forms, containing information which is either stated explicitly or is implied. From the title and the "setting" of the text, it is hypothesised that it should be a relatively simple matter to set up an appropriate scenario. Other of the texts draw upon more general world knowledge which one would expect to reasonably be available for each of the situations described in the texts.

Two types of inference were examined, causal and temporal. In the explicit version of a causal text, a particular link between two events in the text is openly stated; in the other version, this link is implicit and the reader has to infer it for himself. For example, one text is about driving in the country. The explicit version states that the car stops as it has run out of petrol while in the implicit version the car stops and the driver takes an empty can to the garage. From the studies mentioned so far, it would be expected that for fluent adult readers there would be no difference in reading times between explicit and implicit forms as the reader's background knowledge should provide the "missing" linking information.

The second type of inference examined texts in which a temporal shift occurs. As mentioned above, scenarios change in response to appropriate textual cues. In the present study a standard format was used, viz in the implicit form the scenario change was signalled by a time shift beyond the normally expected time span of the first scenario (Anderson, 1981). For example, in one text the first setting is in a church on a Sunday. An event then happens three hours later, which is normally much longer than most church services last and so should signal a shift to outwith the church setting. This shift is openly stated in the explicit version.

It should be noted that there are several different types of inference which can potentially be drawn when, for example, a novel is read. The reader might generate inferences about the motives behind characters' actions, their traits or emotions, the spatial relationship of certain objects, possible future events, the attitude of the writer, and so forth (Graesser & Kreutz, 1993). Genre can influence which are likely to be made, for example, spatial location might be important in solving a crime in a detective story but not in a romantic novel where characters' traits and emotions are highlighted more. In the present study, it is those inferences which are more automatic which are examined and not those more concerned with, say, the author's intentions or attitude.

Wilson (1979) found that deaf subjects took longer to respond to questions which required them to draw an inference than to questions where the information was stated explicitly in the text. This leads to a prediction for the present experiment that hearing readers infer on-line while deaf readers use bridging processes to make the inference at the time of reading the question.

6.2 METHOD

6.2.1 Subjects

Two groups were used, one experimental group of deaf children and one comparison group of normally hearing children of secondary school age.

The experimental group comprised 20 deaf children (13 boys, 7 girls) with hearing losses of 70 dB or greater in the better ear averaged over the four frequencies within the range 500Hz-4KHz. Losses ranged from 75 dB to 116 dB, with a mean loss of 100 dB (standard deviation 12 dB). All were within the normal range of intellectual ability. Ages ranged from 12 years 7 months to 18 years 9 months with a mean age of 14 years and 9 months (standard deviation 19 months). Their most recently recorded IQ had to fall within the average range of ability. All children came from homes where English is the only language spoken. Only children who were prelingually deafened were included, as it has been well established over the years that children who are deafened after the acquisition of speech usually have a great advantage linguistically over those deafened before it (eg Conrad, 1979). In addition, only those children whose parents were normally hearing were included as it is well documented that the small group of deaf children whose parents are also deaf seem to have a linguistic, social and academic advantage over other deaf children (eg Meadow, 1967,1968; Stuckless & Birch, 1966; Vernon & Kho, 1970). Children with any additional, significant handicap were also discounted as this could adversely affect the results obtained. The criteria used above to select the children for inclusion in the experimental group are broadly similar to those advocated by Meadow (1978). All the children attended special schools for the deaf or units for the hearing impaired attached to mainstream schools in Strathclyde Region, and all but two had taken part in the first study.

The comparison group comprised 20 normally hearing children (10 boys, 10 girls) in the first year of a local secondary school. Ages ranged from 12 years 3 months to 13 years 4 months, with a mean age of 12 years 9 months (standard deviation 4 months). All of the children had taken part in the first study.

Only the secondary aged comparison group was used as these children, although slightly younger than the deaf group, would be approaching their level of emotional and conceptual maturity, certainly more so than the primary aged children. Since it was the purpose of this study to examine the top-down processes involved in reading, the older hearing group could be expected to have world experiences and therefore possibly similar cognitive schemata to the deaf children than would the primary group. The subjects in each group were randomly split into two subgroups, and assigned to one of two trials (see materials section below).

6.2.2 Materials

The test material consisted of 15 short texts presented to the subjects on the visual display unit of an Apple Macintosh microcomputer. The material contained either explicitly presented information or information which required an inference to be made. The layout of the short stories was kept uniform across all texts, using the format outlined below:

Title.
First sentence.
Second sentence.
Explicit/implicit material.
Filler sentence.
Question.

Two types of inference were used, temporal and causal. The temporal items were included because these caused significantly more difficulty for the deaf in the first study than the other two types. One explanation for this fact which was mooted at the conclusion of the previous chapter was that the format used in that study could in itself have led to a poorer performance by the deaf. Therefore, in the present study the form of presentation of the temporal items was changed. In each case a time shift was signalled by stating that after a specific time interval something happened. The time interval was such that the event asked about in the question would, by implication, clearly have taken place before or after another event. For example,

FOOTBALL

James went to the football match.
It was a good game.
After the game he saw his sister.) Explicit form
Four hours later he saw his sister.) Implicit form
He kissed her.

Q/Did James kiss his sister after the game?

The causal questions were included in preference to the spatial ones because, from the previous study, it was revealed that the deaf can process spatial inferences as efficiently as hearing children matched on reading age. It was also felt that spatial texts were too simplistic in form whereas causal items gave greater scope for more subtle inference presentations and thus allowed for a more realistic pieces of text to be used.

An example of a causal inference is given below:

A DAY IN THE COUNTRY

Tom was driving in the country.

The car stopped suddenly at the side of the road.

He noticed that the petrol tank was empty.) Explicit form

He took an empty can to the garage.) Implicit form

He was cross about all this.

Q/Did the car run out of petrol?

It should be noted that for all passages, both the explicit and the implicit forms of each text are identical except for the critical sentence (sentence 3). Since the time taken to read each sentence is being measured, care was taken to ensure that both explicit and implicit forms of sentence 3 contained exactly the same number of words.

In addition to the 15 texts used in the experiment, 5 practice texts following the same layout were used.

The complete list of text items used is given in Appendix A3.

No attempt was made to restrict the vocabulary as in the first study. There it was shown that deaf children experienced difficulty with inferencing even when such factors were carefully controlled. Therefore the aim in the present study was to attempt to provide more realistic pieces of text, closer to what might be found in the reading material used in schools. Such texts also provide a richer environment for more subtle inferences to be made. Similarly the restrictive forms of syntax employed in the first study were also abandoned, eg pronoun referents were used instead of repetition of the noun phrase. The passages were scored for level of reading difficulty using the Fry graph (1977) which gave a difficulty level of 6.0 years for both trials, and the Spache formula (1953) which gave a difficulty level of 7.3 years for both trials. The difficulty level is, therefore, very similar to that of the material used in the first study and should be within the capabilities of the subjects of both groups. It should be noted, however, that from both methods of calculating difficulty level, there is a slight increase in the use of more difficult or less common vocabulary, which tends to give the passages a more realistic context. The inferred/explicit, temporal/causal, yes/no responses were randomly distributed in their presentation.

Thus each subject read 35 texts in total, the first 5 of which were practice items. Of the remaining 30 texts half were temporal, half causal. Two "trials" were run for both the deaf subjects and the controls. The first trial comprised 7 explicit and 8 implicit temporal items, and 8 explicit and 7 implicit causal items. In the second trial the explicit and implicit items were reversed, so that there were 8 explicit and 7 implicit temporal texts, and 7 explicit and 8 implicit causal texts. Half the subjects of the deaf and control groups were given the first trial while the other half were given the second trial.

6.2.3 Design of Study

A 2x2x2 factorial repeated measures design was used. The first factor, between subjects or treatments, was subject group and had two levels, deaf and hearing. The second factor, a within subjects factor, was inference type and had two levels, temporal and causal. The third factor, another within subjects, had two levels, inferred and explicit.Subjects in both the deaf and the control groups were split into two sections, each section receiving one trial of the test materials. One trial contained half explicit and half implicit items randomly distributed. The second trial contained the reverse conditions of the material, ie the implicit items in the first trial became the explicit in trial two and explicit became the implicit. Half of the questions required "Yes" responses, half "No" responses, and both these types were randomly distributed.

6.2.4 Procedure

Each subject was seated individually before the visual display unit of an Apple Macintosh microcomputer. Each text was presented one sentence at a time on the screen. As each new sentence appeared, the old sentence disappeared. Each new sentence of the text was displayed when the subject pressed the space bar on the computer keyboard, which was marked "Next". The question was presented when the subject was ready by the subject pressing a key marked "Ready". The subject then responded to the question by pressing one of two keys marked "Y" and "N". The next text was displayed by the subject again pressing the "Ready" key once more. Response times from the initial display of a sentence or question until the next keypress were recorded by the computer. The accuracy of the response to each question, ie whether it was answered correctly or incorrectly, was also recorded.

The instructions (which are given in Appendix A3) were read to the hearing subjects and signed and gestured to the deaf group. The five trial items were used to ensure that subjects fully understood what was required. This was achieved for most subjects by the end of the first practice item and for all subjects by the end of the second.

6.3 RESULTS

6.3.1 Correct responses to Questions

In order to determine whether an inference has been drawn, it is assumed that if a subject answers the question correctly, then the passage has been understood and the appropriate inference has been made.

The number of questions answered correctly for the deaf and hearing groups for the temporal and causal items under explicit and implicit conditions are shown as percentages in Table 6.1 below:

TABLE 6.1: Percentage of questions answered correctly

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	63%	49%	79%	80%
Hearing	77%	67%	92%	85%

These results were analysed using a 3 way, repeated measures analysis of variance, collapsing across questions and treating subjects as random variables, in the manner used previously. This revealed a significant main effect between the groups ($F_{1,38}$=17.463, p<0.01) with the deaf answering fewer questions correctly than the hearing subjects (68% compared with 80%). The analysis of variance also revealed a significant main effect between temporal and causal items ($F_{1,38}$=83.005, p<0.01) with fewer temporal items being answered correctly (64% compared with 84% for causal). In addition the analysis revealed a significant main effect ($F_{1,38}$=9.832, p<0.01) with more explicit than implicit items being answered correctly (78% compared with 70%).

An analysis of variance was carried out, collapsing across subjects and treating the question responses as random variables. This analysis supported the main effect between the groups ($F_{1,56}$=20.900, p<0.01), the main effect with temporal items over causal ($F_{1,56}$=22.221, p<0.01), and almost supported the main effect between explicit and implicit forms of the texts ($F_{1,56}$=3.178, p<0.08). This last result is only a trend and fails to reach significance, but the trend is in the expected direction.

On the by-subjects ANOVA there was found a trend towards an interaction between group membership and temporal/causal items ($F_{1,38}$=2.640, p<0.11). A post-hoc

Tukey test revealed that temporal items prove to be significantly more difficult for both groups ($p<0.01$) than causal items, but particularly so for the deaf. These trends were confirmed by a Tukey test on the by-materials results ($p<0.01$).

It is interesting to note that the deaf answer temporal implicit items correctly 49% of the time. This is little better than chance. Since each question can only be answered "yes" or "no", it is possible to apply the binomial distribution to the above results. The standard deviation for a binomial distribution can be calculated from $\sqrt{pqn}$, where p is the chance of making a correct guess, ie 0.5, q the chance of making an incorrect guess, also 0.5, and n is the number of attempts, in this case 100 since we are dealing with percentage scores. In other words, a score of 5% will account for one standard deviation. Therefore, a score of 60% or more will be two standard deviations above the chance level of 50%, ie almost certainly beyond guessing ($p<0.2$). Thus apart from deaf temporal implicit responses, all of the results above can be accepted as being beyond chance levels.

Summary of results:

The hearing subjects answer more questions correctly than do the deaf. Implicit texts appear to be more difficult to understand than those where the information is stated explicitly. A trend was discovered which suggests that this might operate more strongly for the deaf than for the hearing group.

Fewer temporal items were answered correctly than causal, with the deaf having particular difficulty with implict temporal items.

6.3.2 The critical sentence and the question

In the original studies by Sanford and Garrod, the unit of measurement was reading time for the critical sentence. However, in the present study this was not the most suitable measure. Although within one type of inference (ie temporal or causal) the passages are similar between explicit and implicit conditions except for the critical sentence, and then each condition contains the same number of words, the difficulty arises when comparison is made between different types of inference. The critical sentences of the temporal and causal items on average are likely to have different lengths.

Indeed for the question response times, the recorded times are actually a measure of both the time taken to read the question and the time taken to formulate and respond with an answer. If one examines the the questions it is obvious that the temporal items are on average longer than causal questions. This is necessary by the very nature of the temporal questions themselves. Since each temporal question asks whether one incident happened before or after another, they will inevitably contain two propositions (using the definition used in the first study). Causal questions usually only contain the one proposition.

Since the critical sentences and questions for temporal and causal inferences will have different average lengths, it would be invalid to use reading times for comparison. A better measure in the present study would seem to be rate of reading. Just and Carpenter (1980) suggest that a suitable measure for reading rate is syllables per second, as the syllable would seem to be a more useful unit in reading than either the single word (since words can vary greatly in length) or the single letter (since single letters are used along with diphthongs, consonant clusters etc).

6.3.2.1 Sentence 3:

The mean reading rates in syllables per second for the deaf and hearing groups on Sentence 3 for temporal and causal items under explicit and implicit conditions are shown in Table 6.2 below :

TABLE 6.2: Mean reading rates in syllables per second for Sentence 3

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	3.654	3.826	3.525	3.393
Hearing	4.897	4.677	4.278	4.016

The results are for all texts whether answered correctly or not. Since the number of texts requiring yes and no answers were proportionally equal, only one ANOVA was carried out, combining both types.

These results were analysed using a 3 way, repeated measures analysis of variance, collapsing across texts and treating subjects as random variables, with inference type as the first within subject repeated measure, explicit or inferred as the second within subject repeated measure, and deaf and hearing group as between subject grouping factor. This revealed a significant main effect between the groups ($F_{1,38}=5.586$, $p<0.05$). The deaf were significantly slower overall at reading than the controls (3.600 syllables per second compared with 4.467).

The analysis of variance also revealed a significant main effect between temporal and causal items ($F_{1,38}=31.677$, $p<0.001$), with temporal items being read faster than causal (4.264 syllables per second compared with 3.803). There was also a significant interaction effect between group membership and inference type ($F_{1,38}=4.857$, $p<0.05$). Post-hoc Tukey tests revealed that temporal items were being read significantly more quickly than causal items by the hearing subjects ($p<0.01$) but that this effect was only of borderline significance for the deaf children ($p<0.07$). It should be noted that the deaf were slower than the hearing subjects on all measures.

An analysis of variance was carried out, collapsing across subjects and treating texts as random variables. This analysis confirmed the main effect between the groups

($F_{1,56}$=150.209, $p<0.001$) with the deaf again being significantly slower overall at reading than the hearing group. The analysis also confirmed the second main effect with temporal items being significantly faster than causal items ($F_{1,56}$=9.150, $p<0.005$). The interaction effect found in the by-subject analysis was also repeated in the by-materials analysis ($F_{1,56}$=4.007, $p<0.05$). Post-hoc Tukey tests again revealed a similar pattern to the by-subjects analysis, the temporal items being significantly quicker than the causal for both groups ($p<0.01$), but this effect being more marked for the hearing subjects than for the deaf group.

Summary of findings:

The deaf read the texts more slowly than the hearing children. This might be expected considering the much documented evidence of how poorly they perform generally when reading.

Causal items are processed more slowly than temporal. This effect, while true for both groups, is greater for the hearing subjects.

It should be noted that there is no difference between explicit and implicit texts. This runs contrary to the findings of Wilson (1979).

6.3.2.2 Re-analysis of Sentence 3:

In the above analysis, rate of reading was used as the metric rather than reading times. While this might seem sensible because of the varying lengths, on average, of temporal and causal passages, there is some evidence to suggest that it might be inappropriate even to compare rates directly (Trueswell, Tanenhaus, & Garnsey,1994).

It seems reasonable to assume that as a piece of text increases in length, it takes longer to read, and that this relationship is linear. Under ideal conditions, the slope for this graph, represented by the equation $y = mx$, would pass through the origin as shown in Figure 6.1 below:

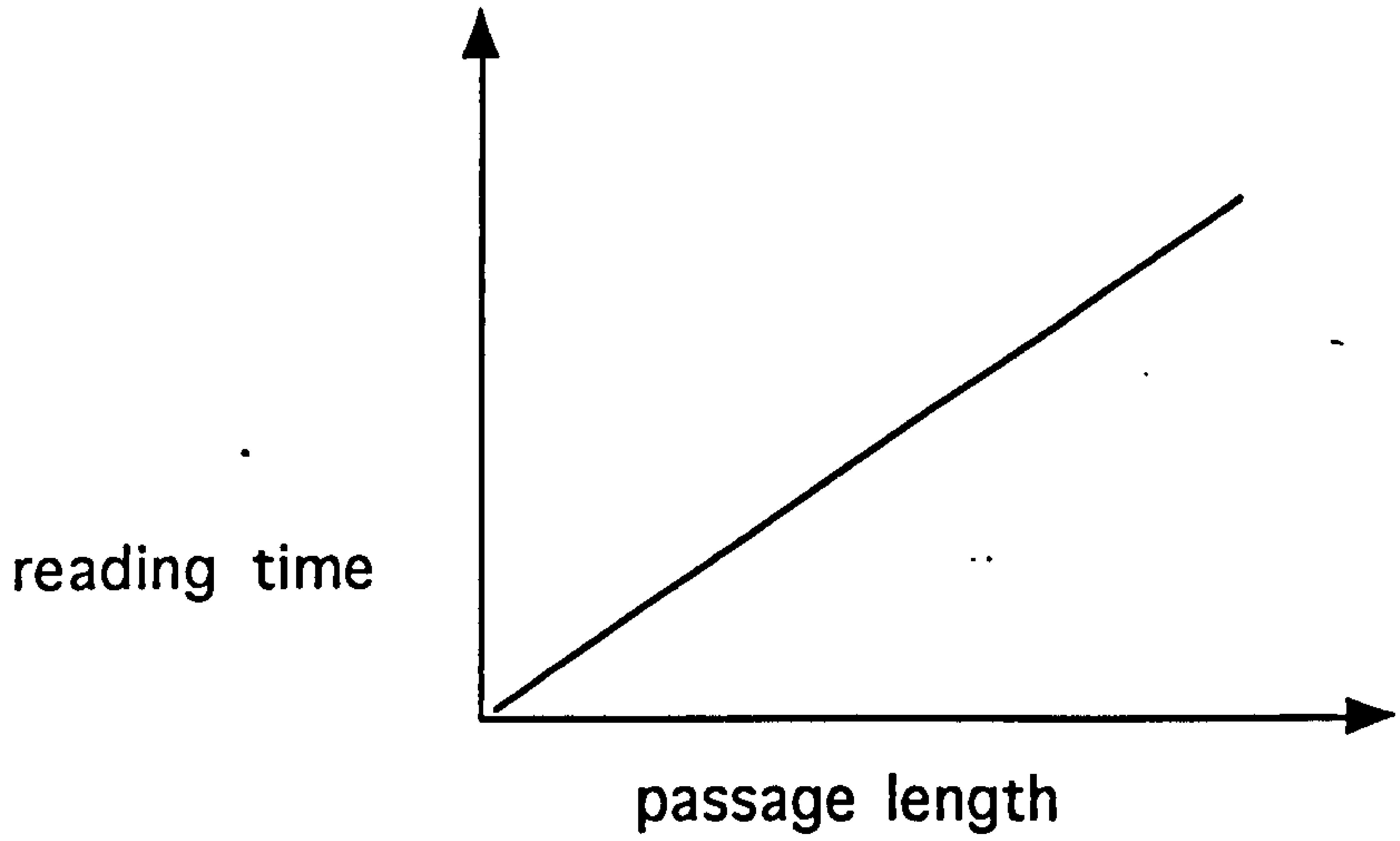


Figure 6.1: Ideal reading time plotted against passage length

However, in reality this graph does not pass through zero. There is a positive intercept on the y-axis (time), as shown below in Figure 6.2. This is represented by the equation $y = mx + c$.

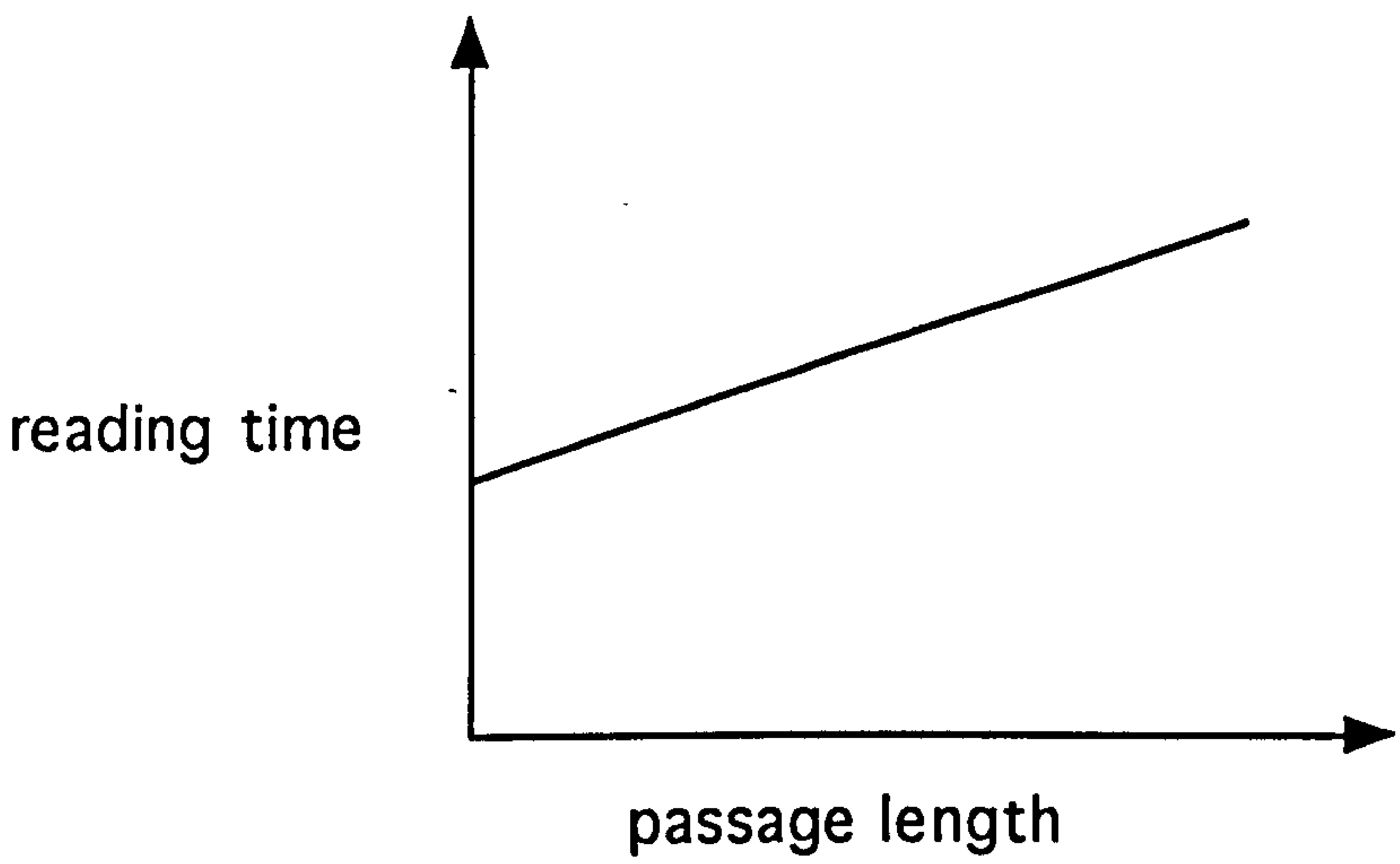


Figure 6.2: Actual reading time plotted against passage length

This means, in effect, that if reading rate (rather than reading time) is plotted against passage length, one does not obtain a constant, as would be the case if the slope passed through the origin. Rather a non-linear, inverse curve as shown in Figure 6.3 is obtained:

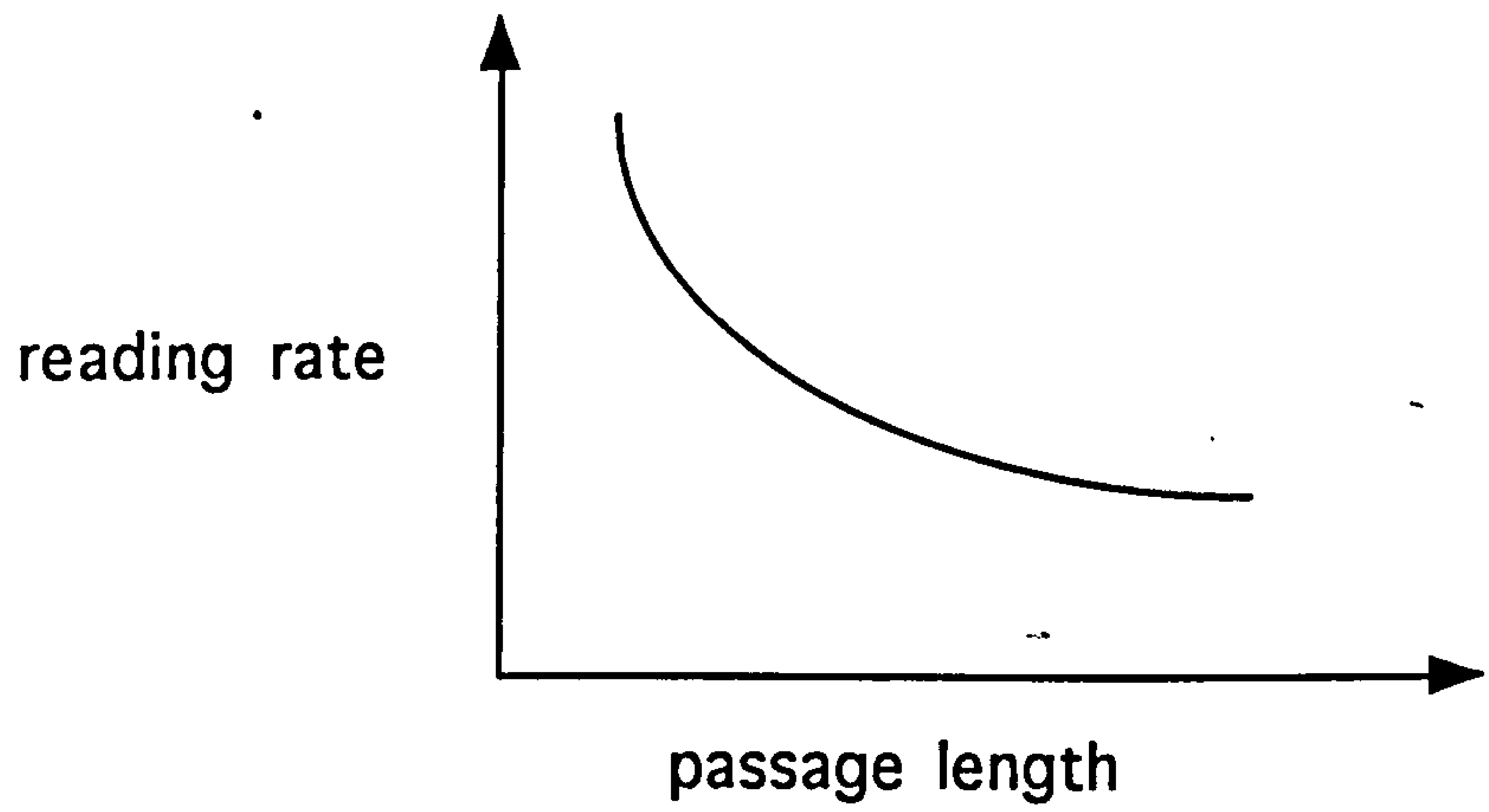


Figure 6.3: Reading rate plotted against passage length

In other words, as the length of text increases, reading rates will be reduced. As the length of text becomes shorter, reading rates are increased. One way to avoid such distortion would be to adjust subjects' reading times by subtracting the intercept, thus creating data which will pass through the origin (Trueswell et al, 1994). When reading times are converted to reading rates, the variance associated with text length is removed.

The average reading times for each passage were plotted against passage length (measured in syllables), for deaf and for hearing subjects. A regression analysis was carried out which gave a residual time of 1238 milliseconds for the deaf subjects and a residual of 909 milliseconds for the hearing subjects. These times were then subtracted from subjects' original reading times, the results converted into new reading rates, and seven outliers removed from the data. These reading rates are shown below in Table 6.3:

Table 6.3: Mean reading rates (syll/sec) for Sentence 3 after removal of residual time component

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	7.634	6.955	8.922	7.994
Hearing	9.005	7.189	7.942	8.721

These results were analysed using a 3 way, repeated measures analysis of variance, collapsing across texts and treating subject as random variables, as above. This revealed no significant factors or interactions. Similar findings were obtained when the results were analysed by collapsing across subjects and treating texts as random variables.

Thus, when the effects of transforming reading times into reading rates are considered, and the distortions which this can cause are removed, a different picture emerges.

There is no longer any difference between the deaf and hearing subjects. While the overall reading rate for the deaf might be slower, as from the unadjusted times, the actual processing speed would appear to be similar for both groups. Similarly, the adjusted rates show that there is no difference in processing speed for temporal or causal items. As before, there is no difference between explicit and implicit texts.

6.3.2.3 Sentence 4:

The times taken to read Sentence 4 were also analysed. This is the sentence following the critical manipulation within the text and it has been suggested that any additional processing time which may be required could "spill over" into this region of the text.

This is similar in effect to the "wrap up" pauses described by Just & Carpenter (1980) which reflect an integrative process at the end of a sentence. Reading times were used rather than reading rates since Sentence 4 is the same under both explicit and implicit conditions. Also the average number of words for temporal and causal types is the same (4.667).

The mean reading times in milliseconds for the deaf and control groups on Sentence 4 for temporal and causal items under explicit and implicit conditions are shown below in Table 6.4:

TABLE 6.4: Mean reading times for Sentence 4 (msecs)

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	2073	2101	1983	2220
Hearing	1639	1600	1656	1678

These results were analysed using a 3 way, repeated measures analysis of variance, collapsing across questions and treating subjects as random variables, similar to the procedure used with Sentence 3 above. This revealed a significant main effect between the groups ($F_{1,38}=6.849$, $p<0.05$) with the deaf group reading Sentence 4 more slowly than the hearing group. This by-subjects finding was replicated in the by-materials analysis ($F_{1,54}=115.337$, $p<0.001$). The by-subjects analysis also revealed a significant interaction between group membership and explicit/implicit conditions ($F_{1,38}=4.355$, $p<0.05$). A post-hoc Tukey test showed that implicit items were significantly more difficult than explicit ($p<0.05$) for the deaf group but not for the hearing subjects. However this interaction effect although not reaching significance on the by-materials analysis did reveal a trend in the same direction ($F_{1,54}=2.366$, $p<0.13$). Post-hoc testing again revealed a trend with explicit items faster than implicit for the deaf group ($p<0.10$) only. This would suggest that implicit processing is delayed until the sentence following the critical manipulation for the deaf subjects but not the hearing children. The fact that the by-materials analysis does not fully match the by-subjects would seem to reflect the greater variability in the materials themselves.

Summary of findings:

No difference was found between temporal and causal texts. It could be argued that this is to be expected since Sentence 4 is a "filler" which follows the critical sentence and would thus not be involved in the process of temporal or causal inferencing.

Again the deaf read more slowly than the hearing subjects, reflecting the general difficulty with reading.

Some implicit items were processed more slowly than the explicit texts by the deaf but not by hearing subjects. This would suggest that any extra processing time which might be required by the deaf to process implicit items is being delayed until the sentence following the critical manipulation.

6.3.2.4 Question-Response

This set of results is the response times between the appearance of the Question and the pressing of a response key, either "Yes" or "No". Thus it is a measure of both the time taken to read the question and the time taken to formulate and respond with an answer. As with Sentence 3, these reading times were converted into reading rates (measured in syllables per second), to allow for differences in question length between temporal and causal items.

The mean reading rates in syllables per second for the deaf and hearing groups on the question-response for the temporal and causal items under explicit and implicit conditions are shown below in Table 6.5:

TABLE 6.5: Mean reading rates for Question-Response (in syll/sec)

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	2.351	2.598	1.865	1.604
Hearing	3.177	2.947	2.309	2.044

The results were for correctly answered questions only as it is these texts in which the inference has been correctly made. Again since the numbers of "Yes" and "No" questions were equally balanced, only one ANOVA was carried out.

These results were analysed using a three way, repeated measures analysis of variance, collapsing across questions and treating subjects as random variables, with inference type as the first within subject repeated measure, explicit or inferred as the second within subject repeated measure, and deaf and hearing group as the between subject grouping factor. This revealed a significant main effect of groups ($F_{1,38}=8.879$, $p<0.01$). The deaf were significantly slower at reading and responding to the question than were the hearing subjects (2.105 syllables per second versus 2.619).

The analysis of variance also revealed a significant main effect of inference type, with temporal items being read more quickly than the causal, $F_{1,38}=69.507$, $p<0.001$, (2.768 syllables per second compared with 1.956).

An analysis of variance was carried out, collapsing across subjects and treating texts as random variables. This analysis confirmed the main effect between the groups found in the by-subjects analysis ($F_{1,54}$=46.127, p=0.01), with the deaf again being significantly slower overall at reading than the hearing group. The analysis also confirmed the second main effect with temporal items being significantly faster than causal items ($F_{1,54}$)= 26.367, p<0.01). No significant difference was found between the implicit and explicit versions of the texts in any of the analyses.

Summary of findings:

Using the overall reading rates (that is, the rates for reading and responding to questions), the deaf are again slower at processing than the hearing subjects.

As with Sentence 3, causal items are slower than temporal but there is no interaction effect and the finding holds equally for both groups of subjects.

There was no significant difference between explicit and implicit items.

6.3.2.5 Re-analysis of Question-Response

As was argued above when considering the critical sentence, conversion from times to rates can cause distortions if comparisons are made of texts which have different average lengths.

Thus the average reading times for each question were plotted against question length (in syllables) for deaf and for hearing subjects. A regression analysis was carried out which gave a residual time of 1926 milliseconds for the deaf subjects and 1942 milliseconds for the hearing subjects. These times were subtracted from subjects' original reading times, the results converted into new reading rates, and five outliers removed from the data.

Only the times of those questions answered correctly were included, as it is these texts in which the inference has been correctly drawn. For this reason the data for temporal texts for the deaf have been dropped from the analysis, as previous results (Section 6.3.1) have shown these to be no better than chance. The corrected reading rates are shown in Table 6.6 below:

TABLE 6.6: Adjusted reading rates (syll/sec) for Question-Response

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	-----	-----	5.431	4.558
Control	7.806	7.538	7.899	6.956

Since the deaf temporal data have been deleted, the remaining results were analysed in two parts, firstly deaf causal compared with hearing causal, then temporal compared with causal for the hearing group only.

A two way analysis of variance, with deaf or hearing group as the between subject grouping factor and explicit or inferred as the within subject repeated measure revealed that the hearing children had a faster rate than the deaf (7.427 syllables per second compared with 4.995), both by treating subjects as random variables ($F_{1,38}$= 6.734, $p<0.01$ and by treating texts as random variables ($F_{1,28}$= 5.447, $p<0.05$). No other effects or interactions were found.

Summary of findings:

When corrections are made for the distortions caused by using reading rates and different lengths of text material, the difference found previously between temporal and causal items vanishes.

The deaf again process the material more slowly than the hearing group, but, as before, there is no difference between explicit andimplicit items.

6.4 DISCUSSION

One of the main findings of this study is that the deaf consistently process the text significantly more slowly than the hearing children for the critical sentence (when uncorrected reading rate is used), the sentence following the critical sentence, and for the Question-Response. When corrected reading rates are used, there is no difference between the deaf and hearing groups at the critical sentence stage. The deaf also answer significantly fewer questions correctly than the hearing group. These results are consistent with the findings of the first study that the deaf appear to have more difficulty with inferencing. They also confirm similar findings by Wilson (1979) and Pinhas (1984). It might be thought that what is being observed is merely the fact that the deaf children, although older chronologically, have poorer general reading skills as measured on standard tests of literacy. However, the textual material used was relatively simple and well within the capabilities of both groups of subjects as regards vocabulary and syntactic structure. Some other factor must be influencing their performance.

Some studies, for example Omanson, Warren, & Trabasso (1978), show that no particular form of inferencing causes any greater difficulty than any other. The types and classifications of inference used in the Omanson study are similar to the types used in the present study and follow the taxonomy of inferences proposed by Nicholas & Trabasso (1980). However, Omanson and his colleagues were measuring only accuracy of processing inferences and not speed of processing. In the present study, temporal items apear to be more difficult to process than causal, particularly for the deaf. Yet when corrections are made for the distortions caused by different text lengths when using reading rate as the metric, there is no difference in processing time between temporal and causal texts.

The accuracy results of the present study, however, indicate that significantly fewer temporal questions are answered correctly, confirming the result from the first study that the deaf find temporal inferences the most difficult, but additionally show that the controls also find these items more difficult, although the difference is more marked in the case of the deaf. Although great care was taken to control the syntactic structures used in the texts, it is conceivable that the deaf find the particular construction used in the implicit temporal items particularly difficult to understand.

From the perspective of those theories which suggest the use of situational models (eg Sanford & Garrod, 1981; Fincher-Kiefer, 1993; Kintsch, 1988), perhaps the most interesting finding to emerge from this study is the fact that no significant difference was found between the reading rates for explicit and implicit materials for both groups for either the critical sentence or the question-response. This would suggest that both groups of readers are making elaborative inferences on-line. The situational model used by the reader brings with it a certain predictability in the persons, actions and likely outcomes so that inferences can be made as the text is read. There is no need for the reader to made time consuming, bridging inferences.

In the temporal items, one might expect that the critical sentence, which contains the implicit temporal shift, would take slightly longer to read as it involves a change of scenario (cf Anderson, 1982). However in the present study even the explicit version of the text contains a time shift, albeit overtly signalled, and this too would require the reader to change scenarios. Thus any additional time would apply equally to both conditions and this is indeed reflected in the results obtained.

For the sentence following the critical manipulation (the "filler"), the deaf again process the text more slowly than the hearing subjects, but there is no difference in processing rate between temporal and causal types. However, the deaf read this sentence more slowly in implicit passages than in those where the links are explicitly stated. The by-materials analysis indicates a trend in the same direction. This would indicate that perhaps some but not all of the passages were causing this effect. It raises the question as to whether the deaf are indeed inferring on-line, or are they making bridging inferences but delaying the process until the sentence following the critical manipulation?

An alternative explanation may be appropriate. Just & Carpenter (1980) discovered that subjects often pause at the end of a sentence to search for referents that have not been assigned while reading, to construct interclause relations and to handle any inconsistencies that could not be resolved within the sentence. This "sentence wrap-up" can occur even when comprehension occurs on-line. Aaronson & Scarborough (1976) and Mitchell & Green (1978) found that subjects tended to pause longer on the word or phrase that terminates a sentence when the subjects are allowed to self-pace the presentation of the text, as in the present study.

Just & Carpenter (1980) found that wrap-up can occur on the final word of a paragraph as well as on the final word or phrase of a sentence. They found that for fluent, adult readers an additional 71 milliseconds on average was required for wrap-up at the end of a sentence but an additional 157 milliseconds for wrap-up at the end of a paragraph. (This latter measurement also includes a small amount of time for subjects to operate a keypress). In the present study, the sentence following the critical manipulation is also the final sentence of each text. The average extra time for implicit passages was 133 milliseconds for the deaf but only 9 milliseconds for hearing subjects. Thus the deaf seem to be taking a small amount of time at the end of some of the texts at least, for general reflection and integration.

This being so, it might be expected that such "wrap-up" pauses would occur equally for the deaf and the controls. However, Jarvella (1971) and Perfetti & Lesgold (1977) have shown that verbatim memory for recently comprehended text declines after a sentence boundary, and attributes this decline to interference between sentence wrap-up processes and the maintenance of verbatim information in working memory. In the first study both deaf and secondary aged children recalled significantly more textual propositions in a verbatim or near verbatim form. However, the deaf were shown to have a weaker memory and recalled significantly fewer propositions overall than the secondary aged children. Thus the weaker memory of the deaf subjects would make them more prone to such interference and could result in a greater wrap-up time at the end of certain implicit texts.

The findings discussed above are evidence of an on-line method of processing textual material by both deaf and hearing children.

The deaf answer fewer questions correctly than the hearing group. This would be expected from the findings of the first study. However, a further finding to emerge is that significantly fewer implicit questions were answered correctly than were explicit, for both hearing and deaf subjects. This effect seems to operate more strongly for the deaf. These findings are still consistent with both groups of subjects operating an on-line, concept driven approach to text processing as only those questions which were answered correctly were included in the reading rate analysis above. In other words, both groups are using the mechanisms suggested when they infer successfully but perhaps these skills are insufficient or ineffective for some passages which may involve situations outwith their general life experience. Further and more detailed examination of particular aspects of how the subjects process text will be needed.

The above findings contradict some of the studies previously carried out in this area. For example, Wilson (1979) found that deaf children were slower in their response times to inferential questions compared with literal questions whereas there was no difference for hearing children. A later study by Pinhas (1984) predicted that hearing children would show equivalent response times when answering inferential or literal questions while deaf children would answer literal questions more quickly. In other words, hearing children would behave like constructive readers using a combination of top-down and bottom-up skills, while deaf children would behave like non-constructive readers. In fact the opposite findings to those predicted emerged. Pinhas' hearing group were slower at inferential questions and the deaf group showed no difference in response times between the two types of question.

It is likely that these differences have arisen from differences in methodology between the present study and those of Wilson and Pinhas. In the Pinhas study, passages were presented one sentence at a time on a videotape monitor, the time of presentation being determined by pre-testing. She used an external timing device to record the subjects' response times (in milliseconds). This timer was started by the experimenter as the last written word appeared. There is no detailed description of the procedure for the Wilson study but, since it is obvious from its layout etc that the Pinhas study is based on Wilson, it is likely that similar methods were used in the earlier study. However this methodology is likely to be highly inaccurate. The timer is started on the appearance of the last word by the examiner but the examiner's reaction time to the appearance of the word and her reaction time in pressing the button to start the timer is likely to vary over time due to fatigue, distractibility etc. In the present study the timer for recording the response times to the questions is started by the microcomputer as soon as the question appears on the screen. The unit of measurement was also the millisecond but the human variability of the experimenter is removed.

Also, there is no guarantee in the Pinhas study that every student has comprehended each sentence of the text. The subject has no control over the presentation rate which is an average obtained by pre-testing. The presentation rate might be too fast for some subjects. The self paced nature of the present study also leads to a more precise picture of what is actually happening as subjects read the texts. Each subject controls the rate of presentation of each sentence by pressing the appropriate button on the computer when he or she has comprehended the previous sentence. This leads to a more accurate insight of the reading process than presenting sentences at a standard rate to all subjects on the basis of pre-testing.

Wilson found that deaf children answered inferential questions less accurately than literal questions whereas there was no difference between types of question for hearing children. Pinhas found similar significant results for hearing children but only a trend for deaf children although in a similar direction to that of Wilson.In the present study, inferential questions prove to be more difficult than literal questions for both deaf and hearing groups. However, conflicting results are obtained in those studies from mainstream research which compare hearing children's literal and inferential processing. For example, Paris & Upton (1976) found that older children process both types of statement with equivalent response times and equal accuracy. In contrast Kail, Chi, Ingram, & Danner (1977) found that answers to literal questions were more accurate and faster than than answers to inferential questions in two groups of children of different ages. These discrepant findings may be attributed to differences in task demands as well as differences in the types of inferential processing required.

While measures of reading time can detect bridging inferences successfully, their use with elaborative inferences is less straightforward. In the present study, if elaborative inferences are made then no additional processing time should be necessary to respond to questions under implicit conditions. However, this means that detection of an inference having been made relies on proving the null hypothesis, ie finding no difference in reading times between implicit and explicit texts. However, the same results would be obtained if no inference was made.

One possible solution, mentioned in Chapter 4, would be to introduce a statement following that in which the inference occurs, which contradicts (or at least does not confirm) the elaboration, and so causes an increase in reading time. Doing this in the present study, however, would make the texts seem unnatural.

By using the question responses, only those texts which had been understood were included in the analysis. In ither words, if the text was understood then the inference must have been drawn. However, it could be argued that the question itself is intrusive. It is possible that the inference might not be drawn on-line and is only made in response to the probe question. If this were so, the reader would be making a backward inference by retrieving the appropriate information from memory. this would take some additional processing time compared with a text where the links are stated explicitly. However, no such difference in response time was found.

Another potential diificulty in using reading time measurements is that, for causal texts, it could be argued that in the explicit version there is the possibility of word-based, associative priming (Keenan, Potts, Golding, & Jennings, 1990). For the causal example given in Section 6.2.2 (A Day in the Country), the critical sentence in the explicit version mentions *petrol* and this could facilitate the reading of the question. However, in these texts there is an additional (filler) sentence between the critical sentence and the question. This, therefore, should reduce any effects of lexical priming (Meyer & Schvaneveldt, 1971).

From the present study it can be stated that deaf children seem to process inferential information in a similar way to hearing children. It is also likely that both groups generate knowledge-based inferences on-line during text comprehension, using situational models. In this respect, they seem to be operating in a similar manner to fluent adult readers.

The findings in the present study apply to reading. Other studies in this area have examined how deaf children draw inferences in what might be considered by some to be their native language, ie signing. It is proposed to examine this aspect in the next study.

Chapter Seven

Do deaf children draw inferences from sign as well as they do from text?

7.1 Introduction

From the results of the second study, it was noted that deaf subjects answer questions about texts where the information is implicit less accurately than when the information is explicitly stated. Also temporal items prove to be more difficult than causal items, particularly for the deaf. In addition, no extra time is required by both groups to process implicit material than to process explicit material. This latter point suggests that both groups are making elaborative inferences on-line, at the time of reading, using situational models in their analysis of discourse similar to those proposed by, among others, Sanford & Garrod (1981).

Sanford & Garrod suggest that hearing subjects also use these mechanisms for analysis of spoken discourse. This has not really been tested experimentally as in live conversation there are many opportunities for interaction between the participants. If one person fails to understand the other, he or she can interrupt and seek clarification. This may not be possible in more structured situations such as formal lectures when interruption would be viewed unfavourably. It is certainly not possible in written methods of communication, where a felicitous style is essential if comprehension is to succeed.

The question can, therefore, be asked as to whether the deaf draw inferences when using sign langauge in a similar way to that operating when they read.textual material.

Previous studies of reading times for materials which have been signed manually for deaf children (eg Wilson, 1979) show that deaf subjects take longer to respond to questions about inferred material than about explicit material. Wilson also shows that no such difference exists for his control group of hearing subjects when the material is presented to them in spoken form. The experimental set-up did not allow interaction between the communicator and the receiver as in live conversation and is, to some extent, an artificial situation. However, this would suggest that for more "natural" media of communication, viz speech for the hearing and signing for the deaf, processes similar to those described by Sanford & Garrod are operating for the hearing subjects but not for the deaf.

A recent study by Emmory and Lillo-Martin (1995) has shown that a pronoun which is signed automatically reactivates its referent which has been mentioned earlier in the discourse. Because of the unique practice of associating the sign for a person or object with a particular location in space, this operates even for "null-pronominals". Emmory, Norman, & O'Grady (1991) have shown that for overt pronouns in American Sign Language, this reactivation process is not immediate and requires some time, occuring about 1000 milliseconds downstream from where the pronoun is first encountered. This is similar to referent reactivation by a pronoun when text is read (MacDonald & MacWhinney, 1990). This would suggest that the fairly automatic process of lexical inferencing at least is occuring on-line.

The present study sets out to examine whether more global inferences are drawn when the mode of communication is sign language, and if so, whether such inferences are made on-line.

7.2 METHOD

7.2.1 Subjects

Since the construction of appropriate videotape material and electronic hardware and the writing of a suitable computer program took some considerable time, some of the subjects who took part in the second study had left school. Also since it was the intention to examine subjects inferential abilities through the medium of manual signing, only those children in the second study who were being taught in a total communication environment would be eligible for inclusion in the present study. Some of these students were used in trial runs to check the equipment and program as they were being developed. Therefore the numbers of subjects who had taken part in the second study and who were eligible for the present study were greatly reduced. Only three children from the original deaf group were used, the others coming from a school for the deaf in Lothian Region.

Only one group of deaf children was used. It comprised 16 children (12 boys, 4 girls) with hearing losses of 70 dB or greater in the better ear averaged over the four frequencies within the range 500Hz-4KHz. Losses ranged from 99 dB to 120 dB, with a mean loss of 110 dB (standard deviation 7 dB). All were within the normal range of intellectual ability (according to their most recently recorded IQ). Ages ranged from 12 years 5 months to 17 years 5 months with a mean age of 14 years 9 months (standard deviation 18 months). All children came from homes where English was the only language spoken. Only children who were prelingually deafened and whose parents have normal hearing were included in the study. These are the same conditions applying in the second study, for the same reasons mentioned therein.

The subjects were randomly split into two groups and assigned to one of two trials (see materials section below).

7.2.2 Materials

The test material consisted of the same 30 short passages and 5 practice items used in the second study. The two trials were the same as those used in that study so that inferred/explicit, temporal/causal, yes/no responses were randomly distributed in their presentation.

The passages were presented on a videotape monitor in British Sign Language (BSL) by a deaf adult who uses BSL as her natural communication medium and who also works as an auxiliary in a school for the deaf and is, therefore conversant with the way deaf children tend to communicate, the limitations (if any) of their signing, and the normal pace of delivery that such children could cope with.

It was a moot point which particular manual form of communication to use. Most schools tend to use Signed English or Sign Supported English. Signed English uses manual sign's in the correct order of spoken or written English while Sign Supported English uses manual signs and finger spelling in the same order as English. It is for this reason that schools are keen on its use. However when the textual material used in the second study was translated into either of these manual forms it proved to be awkward and pedantic. However BSL was much more elegant. In addition many people would argue that BSL is a complete language in itself with a complete morphology and rules of syntax etc, for example, Klima & Bellugi (1979), Fischer & Gough (1978), and Liddell (1975) with American Sign Language (ASL), and Brennan, Colville, & Lawson (1980) with British Sign Language (BSL). Although there are fundamental differences in the two systems, ASL can be considered to be the American equivalent of BSL. It is also, many would argue, especially the deaf community itself, the "natural" language of deaf people. For example, Dalgleish (1975) contends that sign language is the preferred mode of communication for deaf children in America, Holland, and England.

In the present study, no formal assessment was undertaken of each subject's familiarity with and fluency in BSL, but their teachers reported that they would experience no difficulty with the materials used. One teacher who signs regularly for the deaf on television, and is therefore, fairly fluent in BSL, and the native user of BSL who appeared on the videotape presentation in this study both confirmed this assessment.

7.2.3 Design of Study

A 2x2 factorial repeated measures design was used. Only one group of subjects was used. The first within subjects factor was inference type and had two levels, temporal and causal. The second within subjects factor had two levels also, inferred and explicit. Each subject was randomly assigned to one trial of the test material. One trial contained half explicit and half implicit items randomly distributed. The second trial contained the reverse conditions of the material, implicit items in the first trial became the explicit items in the second trial and the explicit became the implicit. Half the questions required "Yes" responses, half "No" responses, and both of these types were randomly distributed.

7.2.4 Procedure

The passages were presented in British Sign Language by a native user of this language on a videotape monitor. The texts were presented exactly as in the second study, ie Title, Text, Question. As the final sign of the question was presented, a 1 KHz tone was sounded on the videotape for exactly one second. The positioning of this tone is, of necessity, somewhat arbitrary but a professional tape editing machine was used for extra accuracy. The tone was initiated 15 frames before the end of the last sign used in the question. The tone thus generated was fed through the analogue input of a BBC B microcomputer and used to trigger the internal clock. The subject was then expected to respond by pressing one of two buttons ("Yes" or "No"). This would then stop the computer's internal clock and the response time would be logged together with a note of whether the response was accurate or inaccurate. Because of the nature of the computer used, response times were in centiseconds. If no response was made within 15 seconds, the program would note this fact and would reset the computer for the next passage on the videotape.

A copy of the control program is presented in Appendix A4.

7.3 RESULTS

7.3.1 Response Accuracy

The number of questions responded to correctly are shown as percentages in Table 7.1 below:

TABLE 7.1: Percentage of questions correctly answered

TEMPORAL		CAUSAL	
Explicit	Implicit	Explicit	Implicit
66.9	54.1	79.1	68.8

These results were analysed using a two way, repeated measures analysis of variance, collapsing across texts and treating subjects as random variables, with inference type as the first within subjects repeated measure, and explicit or implicit as the second within subjects repeated measure. This revealed a significant main effect between temporal and causal items ($F_{1,15}=4.439$, $p<0.05$) with fewer temporal than causal items being answered correctly (60.5% compared with 74.0%). There was also a significant main effect between explicit and implicit items ($F_{1,15}=7.091$, $p<0.05$), with fewer implicit than explicit items being answered correctly (61.4% compared with 73.0%).

An analysis of variance was carried out, collapsing across subjects and using texts as random variables. This analysis confirmed the main effect between temporal and causal items ($F_{1,14}$=12.059, p<0.01) with fewer temporal than causal being answered correctly. However, comparison of explicit with implicit forms revealed only a trend towards fewer implicit being answered correctly ($F_{1,14}$=2.984, p<0.11).

However, the score of 54.1% temporal implicit items is no better than chance (for the reasons given earlier in Section 6.3.1). If this score is removed, comparison can then be made between temporal and causal texts in which the explicitly stated. Comparison can also be made explicit and implicit versions of causal texts.

Since the deaf subjects score very poorly on implicit temporal items, it is possible that this is depressing the overall score for implicit and temporal items. A t-test comparing causal explicit with causal implicit found no significant difference. When temporal explicit items were compared with causal explicit, a borderline effect was observed, $t_{(df=15)}$ = 1.878, p= 0.08, with temporal poorer than causal (66.94% compared with 79.13%). A similar trend was also found analysing by-materials.

7.3.2.1 Response Times

Since the rate of presentation of all the material is fixed by the videotape and the timing device is started after the questions have been presented, only the response times to the questions are measured. Thus differentials between temporal and causal items which complicated matters slightly in the second study do not apply in the present study and response times can be compared directly. Only items which were answered correctly were included in the analysis. Since the number of texts requiring yes and no answers were proportionally equal, only one ANOVA was carried out.

The mean response times in centiseconds for the questions are given below in Table 7.2:

TABLE 7.2: Mean response times (csec) for Questions

TEMPORAL		CAUSAL	
Explicit	Implicit	Explicit	Implicit
169	200	206	229

The data were examined and three outlying values were removed from the by-materials distribution and four from the by-subjects distribution.

These results were analysed using a two way, repeated measures analysis of variance, collapsing across texts and treating subjects as random variables, with inference type as the first within subjects repeated measure, and explicit or implicit as the second within subjects repeated measure. This revealed a significant main effect between temporal and causal items ($F_{1,15}$= 4.812, $p<0.05$), with temporal items being responded to more quickly than causal (1.84 seconds compared with 2.18 seconds). A by-materials analysis also supported this main effect ($F_{1,56}$= 14.318, $p<0.001$). The comparison between explicit and implicit items failed to reach significance in both the by-subjets and the by-materials analyses.

7.3.2.2 Re-analysis of Response Times

The arguements used in the previous chapter concerning the comparison of texts which have different average lengths apply equally in the present study (Trueswell, Tanenhaus, & Garnsey, 1994). Since the previous study used reading rates as the metric and the present study uses reading times, it might be thought that no adjustments need to be made. However, in the present study, the times used are from the end of the question presentation to the subject's response by pressing a key. This means that the subject is reading the signed presentation and starting to formulate his or her answer before the signer has finished. Thus on longer questions the subject has more time to develop a response and so reaction times will tend to be shorter. There would appear to be no easy solution to this problem. It is not possible, therefore, to compare temporal and causal texts.

Since it was noted above that the accuracy of the responses to implicit temporal items were no better than chance, the only further comparison that can be fruitfully made is between explicit and implicit versions of causal texts, and between explicit versions of temporal and causal texts.

The results for causal texts in explicit and implicit forms were analysed using a t-test. This revealed no significant difference in either the by-subjects or by-materials analysis.

Comparison of temporal explicit with causal explicit by t-test revealed no significant difference, either by subjects or by materials.

7.4 DISCUSSION

The original aim of the present study was to examine whether or not the results found in the previous study would be duplicated when deaf children use another medium of language, namely signing. In the previous study it was shown that deaf children infer when reading from implicit material in a manner similar to hearing children. Both groups would seem to be using a situation model, top-down approach to comprehension when reading. Does this approach apply equally well in other areas of language comprehension or just when reading? Sanford & Garrod (1981) assume that this mechanism does operate for all modes of discourse with hearing subjects. However, other studies have found differences in the inferencing ability of deaf children depending on mode of presentation of material, eg Wilson (1979) and Pinhas (1984).

In the present study, it was found that response times were significantly quicker for temporal than for causal items when the deaf subjects used signing. However, as mentioned above, this is probably an artefact introduced because temporal and causal questions are, on average, of different lengths. On longer questions (the temporal texts), subjects have more time to formulate their response as the question is being signed, and tend, therefore, to respond more quickly. Because of the nature of the present experimental design, there is no way round this problem.

However, it was found that temporal items presented a significantly greater degree of difficulty than causal for deaf subjects, both for material which was read and signed. In other words, fewer temporal items were answered correctly under both presentation conditions. It will be remembered that the hearing children, when reading, also found temporal items more difficult. In the previous chapter, it was mooted whether the particular syntactic structure of implicit temporal material was causing difficulty for the deaf. Given that they are having similar difficulty when the material is presented in signed form might suggest that it is not the syntax *per se* which is causing the difficulty, since the signed presentation will follow the particular morphology of BSL, which presumably is different from that of written English.

Also, for material presented in both formats, the deaf subjects could answer more explicit than implicit items correctly. Under both presentations the by-subjects analysis was significant while the by-materials analysis revealed only trends, but in the same direction. Again this holds for hearing children when reading.

It can therefore be seen that as far as signed presentation is concerned, the deaf perform as they do when reading. This is also similar to how the hearing subjects perform when reading, except of course that the hearing group is quicker overall than the deaf.

Again, perhaps the most interesting finding to emerge is the fact that no significant difference was found between the response times for explicit and implicit questions for the deaf group. It could, therefore, be argued that Sanford & Garrod's mechanism for text comprehension already discussed applies equally well when the material to be understood is presented in signed form as it is when presented in written form.

These findings contradict the results of Wilson (1979) and Pinhas (1984). Wilson found that deaf subjects performed significantly better, with a greater degree of accuracy, when material was presented in written form than when it was signed. The explanation offered was that the written form offers a more stable format for deaf subjects than the more transient, ephemeral nature of sign. However, if the overall percentages of correct responses, averaged over both inference types and over explicit and implicit conditions, are compared for the deaf subjects for the reading presentation (the previous study) and for the signed presentation (the present study), then no difference in performance is noted. For the reading presentation, the deaf subjects responded correctly on average 67.8% of the time while for the signed presentation they responded correctly 67.2% of the time. It should be noted that the group of subjects under both presentations was substantially different. Only 3 of the original group who took part in the second study also took part in the third study, and even then it was several months later, so there would be very little chance of them remembering the material.

Pinhas predicted that deaf children would show no difference in accuracy of response between literal and inferential questions which were presented in sign. However, she only found a trend but in the predicted direction. This is the opposite of the the findings of the present study but even in the present study, although the by-subjects analysis did reveal a significant difference in favour of explicit items, the by-materials analysis revealed only a trend. Therefore these particular findings should be treated with some caution.

Pinhas also failed to show that for signed material, deaf subjects show equivalent response times for both literal and inferred questions. In fact the inferred material took longer. The present study found that no difference existed between response times for explicit and implicit material. Pinhas's findings would seem to suggest that deaf subjects use a "text" driven approach when interpreting signed material. The findings of present study, however, are consistent with those of the study in the previous chapter and suggest that the deaf use a concept driven, scenario type approach when comprehending signed as well as written/printed material.

The same arguments put forward in the second study also hold for the discrepancies noted above. Although the exact methodology of Wilson's study is unclear, it seems likely that a "live" signer was used. This certainly was the case with Pinhas. The difficulties on such an arrangement are immediately obvious. Each presentation of the material can vary due to fatigue, distractions, etc. Also it is not possible to guarantee a uniform presentation as regards exactness of signed sequences. British Sign Language and American Sign Language are active, functional languages and are liable to slight variations in, perhaps the order of signs or substitutions of similar but subtly different signs. As in speech presentation, a particular sign (or word) can be emphasised slightly more or less on different occasions, thus adding slight nuances and possibly changing meaning in very subtle ways. It puts a great burden on the signer to ensure exact similarity of presentation each time. The method used in the present study of a videotaped presentation ensures uniformity. Each subject sees exactly the same thing. Both trials were carefully checked for uniformity before use in the experiment.

Also, in the present study, the triggering of the timer was much more accurately controlled. The trigger tone was exactly the same for each subject and the timing process was under computer control. With Pinhas and probably for Wilson, an external timer was started by the examiner after the last sign was presented. This would be most likely to lead to variations in reaction time and judgement of the examiner. Thus the present study brings much more precision to the problem.

In the previous study, it was argued that, although they draw correct inferences less frequently, the deaf can infer in a similar manner to hearing children. It was also noted that their rate of reading was generally slower overall, even when reading text which was not critical to the inference process. However, in the present study, the deaf still experience the level of difficulty with certain forms of implicit material, even when no reading is required.

This would suggest that the difficulties may not result from problems with reading *per se*. Since they occur in more than one modality of communication, they may arise from more general language processing or conceptual problems.

The present findings would suggest consistency over reading and sign interpretation and similarity of mechanisms to those employed by hearing subjects when reading. This is a much more optimistic standpoint as it emphasises the "normality" of the deaf in this respect. However, for some reason the deaf seem to have more difficulty in drawing inferences than their hearing counterparts. A closer look at the mechanism involved when interpreting inferred material is the subject of the next study.

Chapter Eight:

Can deaf children make instrumental inferences?

8.1 Introduction

The results of the previous studies showed that deaf children draw inferences from textual material in a similar manner to hearing children.The deaf, it seems, also employ such a mechanism when drawing inferences which have been presented to them in the from of sign language.

The present study moves from the perspective taken previously of examining the narrative from a global overview to a more detailed molecular viewpoint. Comprehension involves, at least in part, the formation of a coherent representation of the text in the mind of the reader by relating each new phrase in the text to what has already been read. One commonly used and convenient aspect for study is anaphoric reference, in which a noun phrase or pronoun relates to an entity mentioned earlier in the text, for example:

Herb got some beer out of the trunk.
The beer was warm.

The word beer in the second sentence can be read more rapidly in the above instance than if it is referred to indirectly as in the example below:

Herb unpacked the picnic supplies from the trunk.
The beer was warm.

The reason given by Haviland and Clark (1974) for the example above is that beer can be one of several members of the class "picnic supplies", so that a bridging link between the two sentences has to be made, taking extra processing time. In the first instance there was a direct antecedent and no extra time was required.

Several others, eg Garrod & Sanford (1977, 1978), Sanford & Garrod (1977) and Kennedy (1979) have shown that the time taken to map one noun phrase onto another is a function of the class-membership relation between the noun antecedent and anaphor. Since in the example from Haviland & Clark (op cit) quoted above, the mental representation of "picnic supplies" does not seem to readily or directly include an entity for beer, Garrod & Sanford (1981) ponder the extent of complexity which can be ascribed to the meaning representation of a sentence. They consider in particular the case of verbs.

Garrod & Sanford quote the work of, for example, Norman, Rumelhart & LNR (1975), Schank (1973) and Just & Clark (1973) on their descriptions of verbs as being related sets of atomic concepts that together make up the complete meaning of the verb. Examples would be:

KICK..............MOVEMENT of the FOOT
GIVE..............TRANSFER something to a RECIPIENT
PROHIBIT.......ALLOW with a NEGATIVE component

Garrod & Sanford speculate as to whether or not a sentence is represented in a "decomposed" from with the various elements already available or whether such decomposition only takes place when it is needed. They quote evidence from Kintsch (1974) which concludes that decomposition does not take place at time of encoding, as more complex decompositions would be more difficult to retrieve from memory. However Gentner (1981) has shown that what is important is how well connected the entities of the decomposition are, and concluded that decomposition does indeed occur. Carpenter & Just (1977) have also shown that subjects treat verbs like "forget" as NOT-remember, thus offering support for the decomposition hypothesis.

If verbs are represented in decomposed form in the mind of the reader, then this will affect how subsequent sentences are processed. Incoming text could be mapped directly onto the elements of the decomposition, including those elements which are implied by the verb, for example "car" being an element implied by the verb "drive". Therefore the time taken to read a target sentence containing the anaphor "car" following on from an implicit antecedent would take no longer than that following on from an explicit antecedent. Another example is the verb "dress" and the anaphor "clothes":

Mary dressed the baby. (implicit)
The clothes were made of pink wool.

Mary put the clothes on the baby. (explicit)
The clothes were made of pink wool.

Early studies, for example Anderson & Ortony (1975) with adults, and Paris & Lindauer (1976) with children, suggested that such instrumental inferences are routinely drawn. Subjects read passages which contained sentences such as "*The container held the apples*". At the end of the passage, a cue word was presented ("*basket*" or "*bottle*" for the example given), and subjects were asked to recall the original sentence. It was found that "*basket*" leads to a quicker recall of the sentence than "*bottle*", this being interpreted as evidence that the specific type of container (ie a basket) is inferred at the time of reading the sentence.

However, the memory technique used in these studies has subsequently been criticised (Alba & Hasher, 1983; McKoon & Ratcliff, 1981, 1986; Singer, 1979). It is possible that the findings could be due to differential ability in finding "*container*" in memory at the time of recall. Singer (1980) used instrumental cues which have different association strengths with verbs. For example, "*spoon*" has a strong forward association with "*stirring soup*" but a weak reverse association. "*Ladle*" has a strong reverse association with this verb phrase but a weak forward association. By showing that it was the backward association from the instrumental cue to the verb phrase (rather than the forward associations) which predicts the ease of recall, Singer indicated that in the study by Anderson & Ortony (1975), the cued recall task was tapping search processes at the time of recall, rather than the inference having been drawn at the time of reading.

Garrod & Sanford (1981) used reading times in their study and found that instrumental inferences are made as the text is read. They found that it took subjects no longer to read a sentence containing, for example, a reference to "*the car*' when it followed a sentence containing the verb "*drive*" (and no explicit mention of "*car*") than when a car was mentioned explicitly in the preceding sentence.

Singer (1979), using a similar experimental design to Garrod & Sanford, found that subjects did in fact take longer to read a sentence which contained an instrument, if this instrument had only been implied by use of a verb in the preceding sentence. For example, it took longer to read "*The shovel was heavy*" following on from the first sentence below than the second:

"*The boy cleared the snow from the stairs*" (Implied)
"*The boy cleared the snow with a shovel*" (Explicit)

One possible explanation for the discrepant results from the two studies quoted above is put forward by Cotter (1984). By examining the dictionary definitions of the verbs used in each study, and using a word recognition task, she suggests that, in effect, two different subsets of verbs are used. For those used by Garrod & Sanford, the instruments seem to be a more inherent part of the definition of the verbs than those used by Singer. She also felt that these differences were tapped by the reading task, where subjects are under a degree of pressure, but not detected at the more relaxed pre-testing stage when subjects completed a questionnaire to define each verb.

Examination of the textual material used in the Garrod & Sanford study and the Singer study suggests a possible alternative explanation. Garrod & Sanford use passages containing three sentences while Singer uses passages which contain only two sentences. O'Brien, Shank, Myers, & Rayner (1988), using eye movement data, have shown that subjects do make elaborative inferences at the time of reading if the context is sufficiently constraining, or if they are encouraged to do so by a demand sentence which focusses attention on the inferred concept. It could be argued that the Garrod & Sanford texts, because of their greater length, offer a slightly richer and therefore more constraining context. However, in a subsequent paper, Garrod, O'Brien, Morris, & Rayner (1990) reanalysed the results of O'Brien et al (1988) and conducted an experiment which showed that inferences were drawn on-line in constraining contextual conditions only where a definite anaphoric relationship exists between the anaphor and its antecedent. In other words, an inference would only be drawn automatically when reference was made to a definite instrument (eg *the shovel*) rather than an indefinite reference (eg *a shovel*).. Yet it would appear that in both the Singer and the Garrod-Sanford materials it is definites which are used.

Corbett & Dosher (1978) used the Stroop technique (described earlier in Chapter 4) to investigate whether instrumental inferences are made. Subjects read sentences similar to that used by Singer, eg

The boy cleared the snow from the stairs.

This was followed by a probe which was either a semantic associate of "*shovel*" or an unrelated word. The probe was written in coloured ink and subjects had to name the colour. If the target word is activated, then because of interference, it should take longer to name the colour. Corbett & Dosher, however, found no evidence for instrumental inferences occurring, unless subjects were instructed to make such inferences.

A recent study by van Meter & Pressley (1994) used a recognition probe consisting of a word fragment based on the implied instrument to examine whether children (aged 10 to 14 years) make instrumental inferences on-line when reading single sentences. However, results indicated that such inferences were not made spontaneously but were made when subjects were instructed to do so.

The evidence as to whether instrumental inferences are generated during the reading process is equivocal. It is possible that such inferences may occur with particular verbs but not with others. Since the present study seeks to use essentially the same verbs as were used by Garrod & Sanford and a similar experimental design, it would be expected that hearing subjects would generate instrumental inferences at the time of reading. Since, from the evidence of the studies in the two previous chapters, deaf children appear to infer in a similar manner to hearing children, it is further hypothesised that the deaf will also generate instrumental inferences on-line.

8.2 METHOD

8.2.1 Subjects

Two groups were used, one experimental group of hearing impaired children and one comparison group of normally hearing children of secondary school age.

The experimental group comprised 19 deaf children (14 boys, 5 girls) with hearing losses of 70 dB or greater in the better ear averaged over the four frequencies within the range 500 Hz to 4 KHz. Losses ranged from 75 dB to 116 dB, with a mean loss of 99 dB (standard deviation 11 dB). Ages ranged from 12 years 3 months to 16 years 11 months, with a mean age of 14 years 9 months (standard deviation 14 months). All the subjects were within the normal range of intellectual ability, as determined by their most recently recorded IQ. All children came from homes where English is the only language spoken. Also, only children who were prelingually deafened were included, for the same reasons given in the second study, viz children who have been deafened after the acquisition of speech are generally considered to have a great advantage linguistically over those deafened after this stage (eg Conrad, 1979). Additionally, only those children whose parents were normally hearing were included as it is well documented that the small group of deaf children whose parents are also deaf seem to have linguistic, social and academic advantage over other deaf children (eg Meadow, 1967, 1968; Stuckless & Birch, 1966; Vernon & Kho, 1970). Children with any additional, significant handicap were also discounted as this could adversely affect the results obtained. All the children attended schools for the deaf or units for the hearing impaired in Strathclyde Region, and all but four had taken part in the second study. The four new subjects were substituted for subjects who had left school or were ill when the present study was run.

The comparison group comprised 20 normally hearing children (12 boys, 8 girls) in the second year of a local secondary school. Ages ranged from 12 years 11 months to 13 years 9 months, with a mean age of 13 years 4 months (standard deviation 3 months). All but five of the children had taken part in the second study.

8.2.2 Materials

The test material consisted of 20 short texts presented to the subjects on the visual display unit of an Apple Macintosh microcomputer. These were preceded by four practice items using the same format.

The material contained those verbs used by Garrod & Sanford (1981) in their study on verb decomposition. Only 13 of the original 16 verbs were included, the other 3 being discarded as it was felt they may be inappropriate for children. Those verbs not used were:

grow....(height)
kiss.....(lips)
sell......(money)

The reasons given by Garrod and Sanford themselves as to possible confusion with these verbs were thought to possibly operate more strongly in the case of children. For example, "grow" could refer to human height or the cultivation of flowers or vegetables; "kiss" could imply lips or mouth; "sell" would suggest firstly supplying goods, with receiving money for this service as a secondary albeit important component.

In order to obtain more possible verbs which could be included with the experimental material, the same procedures used by Garrod and Sanford were adopted. A list of verbs with one word of context, if necessary, were presented to a group of subjects who were asked to define the verbs, based on the example: *CARVE: Use a knife to slice meat*. The material was pretested on a group of 14 adults (teachers, psychologists, civil servants) who were unaware of the purpose of the study.

This list is given below in Table 8.1, with the one context word in brackets, where one was given, and the expected noun phrase. The percentage of subjects who responded with the expected noun phrase (instrument) is shown for each verb.

TABLE 8.1: Percentage of responses mentioning the expected noun phrase.

PAINT	brush	57%	
SWIM	water	100%	*
EAT	food	93%	*
READ	book	36%	
SING	song	14%	
SHOOT (rabbit)	gun	64%	*
CLIMB (to roof)	ladder	7%	
SIT	seat	14%	
BOIL (water)	kettle	14%	
TIE (shoe)	laces	79%	*
DIG (garden)	spade	86%	*
CLEAN (teeth)	toothbrush	71%	*
WASH (hands)	soap	50%	
STIR (tea)	spoon	79%	*

Garrod and Sanford used the criterion for inclusion of material that at least 70 % of the subjects included the implied entity in their definition. Those items marked with an asterisk in the above list were included; this includes all items over 70%, with only one verb ("shoot") achieving slightly less than this level (64%).

The full set of materials is listed in Appendix A5.

The layout of the texts followed a pattern similar to that used by Garrod and Sanford. This comprised three sentences followed by a question. An example of the type of passage used is shown below:

Tom was going to London.

He was taking his car.	(Explicit form.
He was driving there.	(Implicit form.
The car had been fixed.	(Noun phrase.
It had just been fixed.	(Pronoun.

Q/(*Did Tom go to London by car?* Y
(*Did Tom go to London by train?* N

The *first sentence* sets the scene for the short narrative.

The *second sentence* contained either an implicit version, where the entity or instrument is left unstated and is only implied by use of a verb, or an explicit version, where a paraphrase is used with the noun phrase explicitly mentioned.

The *third sentence* was the target sentence containing the entity or instrument as an anaphoric referent. This was mentioned explicitly as a noun phrase as in the Sanford-Garrod experiment, or indirectly as a pronoun. It should be possible to correctly assign meaning to the pronoun following explicit mention of the entity or instrument in the second sentence but much more difficult following the implicit, verb version. For the example noted above, it should be fairly straightforward, even in the implicit version, to assign the subsequent noun phrase "*the car*" to the vehicle which was being driven. However, if a pronoun follows an implicit sentence, there is a conflict (Sanford, Garrod, Lucas, & Henderson, 1983). Subjects would need to make a greater effort to resolve the pronoun, and make more bridging inferences to try to connect it to something in the previous text. Although this should be more difficult, it should still be possible for hearing subjects at least (Sanford, 1990, p524). It may, however, be too difficult for deaf subjects, given the problems they seem to have with pronouns (eg Quigley et al, 1974, 1976, 1977).

The *question* then asks if the particular entity or instrument was used or mentioned in the narrative, this requiring an affirmative answer. For one in every five trials, an alternative entity or instrument was used which would require a negative answer, this procedure being used to ensure that subjects read for meaning.

Measures of reading times were taken for the third sentence and for the question response. Whether a question was answered correctly or not was also noted.

8.2.3 Design of Study

A 2x2x2 factorial repeated measures design was used, with subject group as the between subject grouping factor, implicit/explicit condition as the first within subject repeated measure, and noun/pronoun as the second within subject repeated measure.

Subjects were given one of four trials. Texts were presented in either explicit or implicit form, with a noun phrase or pronoun anaphor. One text in each group of five required a negative response to the question, this being done as a check to ensure that subjects read for meaning. The explicit/implicit, noun/pronoun, yes/no conditions were distributed randomly throughout the trials.

8.2.4 Procedure

Each subject was seated individually before the VDU of an Apple Macintosh microcomputer. Each text was presented one sentence at a time on the screen. As each new sentence appeared, the old sentence disappeared. Subjects controlled the rate of presentation and gave their responses to the questions by pressing selected keys on the computer keyboard. The response times between the presentation of one sentence and the appearance of the next and the responses to the questions (yes/no) were recorded by the computer.

8.4 RESULTS

Measures were taken of (a) the percentage of questions answered correctly, (b) the time (in milliseconds) to read the third sentence, which follows the critical manipulation of the material and contains the anaphoric referent in either noun phrase or pronoun form, and (c) the time (in msecs) taken to read and respond to the question.

8.4.1 Number of Correct Responses

The percentages of questions which were answered correctly under each condition by each group are shown in Table 8.2 below. Only those questions which required an affirmative response were included, the "no" responses being included merely as a check that subjects were reading for meaning. While Garrod and Sanford used responses to all questions in their analysis, their subjects were undergraduates, who would be expected to have a high level of literacy. From the results of the studies reported in earlier chapters, it is known that the subjects in the present series of studies, particularly the hearing impaired, are poorer at drawing inferences successfully. Subjects could respond correctly to questions which require a negative response but it really remains unknown whether they have drawn the expected inference. In the example given above about Tom driving to London, it is possible to answer the question "*Did Tom go to London by train?*" correctly by answering "no". However, it could just be possible that the reader might think he travelled by plane or helicopter. It is only for those questions which require an affirmative response that there can be any certainty that the correct inference has been drawn.

TABLE 8.2: Percentages of questions correctly answered.

	EXPLICIT		IMPLICIT	
	Noun	Pronoun	Noun	Pronoun
Deaf	88%	87%	82%	70%
Hearing	91%	96%	89%	91%

These results were analysed using a 3 way repeated measures analysis of variance, collapsing across texts and treating subjects as random variables. This revealed a significant main effect between groups ($F_{1,37}$= 13.375, $p<0.001$) with the deaf answering fewer questions correctly than the hearing subjects (81.6% compared with 91.9%). The analysis also revealed a significant main effect between explicit and implicit conditions ($F_{1,37}$= 6.292, $p<0.02$) with fewer implicit texts than explicit being answered correctly (82.8% compared with 90.6%). In addition, the analysis revealed a trend in the interaction between group membership and noun/pronoun presentation ($F_{1,37}$= 3.791, $p<0.06$). Post-hoc simple effects analysis indicated that the deaf were significantly poorer at dealing with the pronoun condition than were the hearing group ($p< 0.001$), while there was a trend for the deaf to make more errors with pronouns than noun phrases.

The results were then analysed by collapsing across subjects and treating texts as random variables. This ANOVA replicated the main effect between groups ($F_{1,60}$=

11.072, $p<0.005$), with the deaf again poorer than the hearing subjects. The second main effect between explicit and implicit material was also replicated ($F_{1,60}= 6.066$, $p<0.05$), with implicit again being more difficult than explicit. The analysis also replicated the trend in the interaction between group membership and noun/pronoun presentation ($F_{1,60}= 1.424$, $p<0.07$), with post-hoc testing revealing a similar pattern as in the by-subjects analysis.

Summary of findings:

The deaf make fewer correct responses than hearing subjects.

Fewer implicit items are answered correctly by both hearing and deaf subjects.

The deaf find pronoun referents more difficult to understand than hearing subjects.

8.4.2 Third Sentence

The reading times for Sentence 3 are shown in Table 8.3 below:

Table 8.3: Reading times (msecs) for Sentence 3.

	EXPLICIT		IMPLICIT	
	Noun	Pronoun	Noun	Pronoun
Deaf	2296	2185	2363	2381
Hearing	1778	1781	1982	2125

The results are for all texts whether answered correctly or not. They were analysed using a 3 way, repeated measures analysis of variance, collapsing across texts and treating subjects as random variables, with explicit or implicit as the first within subject repeated measure, noun or pronoun version as the second within subject repeated measure, and deaf and control group the between subject grouping factor. This revealed a borderline significant main effect between the groups ($F_{1,37}$= 3.610, $p<0.06$). Therefore the deaf are significantly slower overall at reading than hearing subjects (2306 milliseconds compared with 1917). The analysis of variance also revealed a significant main effect between explicit and implicit items ($F_{1,37}$= 6.754, $p<0.01$), with implicit items being read more slowly than explicit (2213 milliseconds compared with 2010). There were no other significant main effects or interactions.

An analysis of variance was carried out, collapsing across subjects and treating texts as random variables. This analysis confirmed the main effect between groups ($F_{1,76}$= 46.650, $p<0.001$) with the deaf again being significantly slower overall at reading than the comparison group. The analysis also confirmed the second main effect with implicit items being read significantly more slowly than explicit ($F_{1,76}$= 3.815, $p<0.05$). Again there were no other significant main effects or interactions.

Summary of findings:

The deaf read the texts more slowly than hearing subjects, this finding being consistent with the results found in previous studies and reflecting how poorly the deaf perform when reading.

Implicit items are processed more slowly than explicit by both groups. This finding is contrary to what one might expect given the findings of Sanford & Garrod's verb study.

8.4.3 Question-Response

Only those questions which were answered correctly were analysed. The times (in milliseconds) taken to read and respond to the question are shown in Table 8.4 below:

TABLE 8.4: Question-Response times (msecs)

	EXPLICIT		IMPLICIT	
	Noun	Pronoun	Noun	Pronoun
Deaf	3170	3315	3592	3656
Hearing	2395	2506	2642	2735

These results were analysed using a 3 way, repeated measures analysis of variance, collapsing across texts and treating subjects as random variables, similar to the procedure used with Sentence 3 above. This revealed a significant main effect of Group ($F_{1,37}$= 12.273, $p<0.001$) with the deaf group reading and responding to the question more slowly than the hearing group (3433 milliseconds compared with 2569). This analysis also revealed a significant main effect of inference type ($F_{1,37}$= 11.483, $p<0.002$) with implicit items being read and responded to more slowly than explicit items (3156 milliseconds compared with 2846).

An analysis of variance was carried out, collapsing across subjects and treating texts as random variables. This analysis replicated the main effect between groups ($F_{1,60}$= 47.040, $p<0.001$) with the deaf again reading and responding more slowly than the controls. However the main effect between explicit and implicit texts was not reproduced in this by-materials ($F_{1,64}$= 2.405, $p<0.14$). This would seem to indicate that only some of the texts are causing the significant main effect between explicit and implicit conditions on the by-subjects analysis.

The text were examined, particularly those verbs which were introduced in the present study but not included in the original Sanford & Garrod experiment. In light of Cotter's (1984) findings, particular attention was paid to the presence of any so-called Singer verbs. Cotter's "dictionary exercise" was used, ie the six "new" verbs were looked up in the Oxford English Dictionary to see if and how often an instrument was cited as part of the dictionary definition. Two such Singer verbs were discovered, viz "tie" and "clean". These verbs were removed from the data array and the new raw data re-analysed.

This re-analysis revealed a significant main effect between groups in both the by-subject analysis ($F_{1,38}$= 15.670, p<0.001) and the by-materials analysis ($F_{1,56}$= 64.500, p<0.001) as well as a significant main effect between explicit and implicit texts, again in both by-subjects ($F_{1,38}$= 12.344, p<0.005) and by-materials analyses ($F_{1,56}$= 5.520, p<0.05).

This last finding, however, runs contrary to what one might expect. It would seem more likely to find a significant difference between explicit and implicit texts for the Singer verbs and not for the Garrod-Sanford material. In effect, the difference is enhanced, not reduced, by the removal of the two Singer verbs.

Summary of findings

The deaf read and respond to the questions more slowly than hearing subjects. For some of the.textual material, implicit texts would seem to be read more slowly than explicit.

8.5 Eliciting definitions for verbs: hearing subjects

Garrod & Sanford pretested their list of verbs on a group 15 subjects by having them define each verb. The criterion used for selecticn was that a high percentage (in effect 70% or greater) of responses included the the expected noun phrase, for example, "*clothes*" for the verb "*dress*". However, the subjects used by Garrod and Sanford in this selection process were adults.

The materials used in the present study were essentially those used by Garrod & Sanford, with the addition of seven new verbs. These additions were selected using the same procedure, pretesting on a group of adults. However, it is possible that children might produce different responses if they were asked to define the list of verbs.

The pool of verbs, from which the seven were selected, were presented post-hoc, along with those chosen from Garrod-Sanford, to a group of 42 secondary pupils of the same age as the control group. They were asked to define each verb following the Garrod-Sanford procedures. The percentage of responses which contained the expected instrument or noun phrase are shown in Table 8.5 below:

TABLE 8.5: Percentage of responses for expected noun phrase: hearing children.

PAINT	36%		DRIVE	86%	*
SWIM	83%	*	CUT (hair)	71%	*
EAT	64%*		FREEZE	57%	
READ	40%		PHOTOGRAPH	76%	*
SING	12%		POST (letter)	67%	*
SHOOT (rabbit)	71%	*	RIDE	64%	*
CLIMB (to roof)	7%		BAKE	67%	*
SIT	33%		SAIL	69%	*
BOIL (water)	17%		DRESS	71%	*
TIE (shoe)	52%		SMOKE	60%	
DIG (garden)	36%		BUY	79%	*
CLEAN (teeth)	64%	*	UNLOCK	81%	*
WASH (hands)	48%		FLY	71%	*
STIR (spoon)	86%	*			

It can be seen that five of the seven new verbs which scored over 64% response with the adults also scored over 64% with the hearing children, viz *swim, eat, shoot (rabbit)*, and *clean (teeth)*. Only two of the new verbs failed to reach the desired cut off point, *tie (shoe)* and *dig (garden)*.

For those verbs selected from the Garrod-Sanford list, all except two (*freeze* and *smoke*) have 64% or over of responses which contain the expected instrument or entity.

Although Garrod and Sanford use a 70% cut off as the criterion for including a verb in their materials, they do not specify the exact percentage response for each verb. Therefore it is not possible to compare the children's responses directly with those of the adults. However, from the pool of new verbs, from which seven were selected, there is a similar pattern between child and adult responses(compare Table 8.5 with Table 8.1). It seems as though there is broad agreement between the two groups, with the children's responses being perhaps slightly more diverse (the adults rated 9 verbs with 80% or greater agreement, the children only rated 2 verbs at this level of agreement).

8.6 Discussion

The above results confirm those from earlier studies, that the deaf children read more slowly than the hearing subjects at both the critical sentence and question response. They also answer fewer questions correctly than the controls. Apart from this finding there is no difference between the deaf and hearing children, other than the greater difficulty experienced by the deaf when dealing with pronouns. This is to be expected as it has been well documented by several authors that pronominalisation is one of the syntactic structures of written English which presents some difficulty for the deaf, eg Quigley, Power, & Steinkamp (1977).

What is more interesting is the fact that elaborative inferences appear not to be made by both groups of subjects. It would seem that the verbs are not stored in a decomposed form as was postulated by Garrod & Sanford (1981). Implicit items take longer to read for both groups at both the critical sentence and question response. Therefore, subjects appear to be making backward bridging inferences rather than forward elaborative inferences. No evidence was found in the present study for the instruments associated with each verb. The findings are more consistent with those of Singer (1979).

Cotter (1984) re-examined the material used in both the Singer and the Garrod & Sanford studies. She felt that the different results which were obtained were due to two different subsets of verbs being used in each experiment. Those used by Garrod & Sanford had a stronger semantic relationship between verb and associated instrument, as was evident from the dictionary definition of the verb, which cites the instrument as part of the definition. Those used by Singer did not.

In the present study, the majority of verbs used were those from Garrod & Sanford (1981) with seven additions. These seven were selected by the same procedures used by Garrod & Sanford, being rated by a group of adults. When these seven verbs were subsequently rated by a group of hearing adolescents, only five of them qualified using the Garrod-Sanford criteria. Of the others chosen from those used by Garrod and Sanford in their original study, all but two met the necessary criterion. However, the children's responses seemed to be more diverse than those of the adults, with responses during the rating exercise reaching a slightly lower level of agreement. Perhaps children use different cognitive strategies when dealing with verbs than do adults. While this is one possible explanation, other factors could be operating.

A more likely hypothesis is to be found in the layout of the textual passages themselves. A series of studies by O'Brien, Shank, Myers, & Rayner (1988) showed that as the context of a passage became increasingly constraining, and therefore increasingly richer, there was more likelihood that readers would make elaborative inferences on-line. Readers tend to take a cautious approach by not making any inference which is not necessary at the time (eg Perfetti, 1993). They do not commit themselves to an interpretation which might prove to be incorrect later on. Within a fairly rich context, elaborations can be made more safely as the passage is unlikely to throw up any surprises or unexpected twists.

As mentioned in Section 8.1, a subsequent study by Garrod, O'Brien, Morris, & Rayner (1990) showed that inferences should be drawn automatically in a constrained context only when definite reference is made to the instrument. However, all texts used in the present study contained definite references to the instruments. None had an indirect reference, yet inferences were not drawn on-line. This would also appear to be true for the Singer (1979) and the Corbett & Dosher (1978) studies cited earlier. Therefore, the particular linguistic constraints stated by Garrod et al (1990) would not appear to be operating in these cases.

This could explain the differences found between the Garrod & Sanford and Singer studies. The Singer passages consisted of only two lines of text whereas Garrod & Sanford used three lines. This additional line might be enough to add sufficient context to allow elaborative inferences to occur on-line.

However it would seem that this one additional line of text is insufficient for the younger readers who participated in the present study. Both the hearing and deaf subjects failed to draw elaborative inferences on-line. In the earlier on-line study in Chapter 6 and in the signing study in Chapter 7, subjects were inferring on-line. However, the material which was used in both these studies was much richer in content, comprising four lines of text as well as a title to help set up a more specific context. It is possible that younger readers are more cautious than adults and require even more context before committing themselves to elaboration.

The adult subjects in the Garrod-Sanford and Singer studies were undergraduates, of above average ability and with (presumably) well developed literacy skills. The children used in the present study were within the average ability range and could reasonably be expected to have less well developed literacy skills (for example, metalinguistic skills - see Section 1.7). It is possible that the children lack the confidence of the adults and for this reason are less likely to commit themselves to making firm elaborations when reading passages which do not have a rich and, therefore, relatively constrained context.

Thus the findings of the present study are consistent with the current consensus of opinion regarding instrumental inferences. It is now widely accepted, even by proponents of the constructionist viewpoint who see the reader as making elaborative inferences, that instrumental inferences are not normally generated on-line (eg Graesser, Singer, & Trabasso, 1994).

It has been suggested in previous chapters that the deaf have some difficulty with drawing inferences from implicit material when this is presented in either written or signed form. However, when they do draw inferences correctly they do so on-line in a manner similar to that of their hearing counterparts. The question may be asked, therefore, as to why the deaf infer successfully only some of the time. It was suggested earlier that this might be because they lack certain experiences and, as a result, have difficulty in calling upon appropriate scenarios and background knowledge when necessary. It is towards this consideration that we turn our attention in the next study.

Chapter Nine

How do subjects construct scenarios?

9.1 Introduction

The studies in previous chapters have shown that deaf children are less successful at drawing inferences than hearing children, but that when they do, they do so in a manner similar to their hearing counterparts. It seems that some textual passages may cause greater difficulty for the deaf than others.

It was postulated that perhaps the deaf have impoverished life experiences because of their communication difficulties. They are less likely than hearing children to absorb information from, say, television and videos. When a hearing child watches a programme on TV, there is likely to be discussion of the programme content with parents or siblings which extends the child's knowledge base. For a deaf child, communication levels with those around him are likely to be less fluid, with, possibly, a lower level of elaboration.

This is only one source of possible reduced input. A deaf person is much less likely to read for information or pleasure. Cornett (1975), noticed when he came to work at Gallaudett University (a university for the deaf) how few students read for pleasure. Cornett estimated the figure to be as low as 5 percent, yet these are very bright deaf students.

The purpose of this study was to attempt to examine the knowledge base of deaf children, at least with regard to those scenarios or stereotypical situations used in the earlier studies.

9.2 Estimates of scenario duration

9.2.1 Rationale:

The purpose of this part of the study was to explore how subjects would estimate typical time limits for a variety of scenarios. In the on-line study described earlier, the time shifts used in the texts were devised by the experimenter in consultation with a small number of colleagues. Times were chosen so as to be clearly outwith what was seen as the "usual" time span for various situations. For example, one text describes a football match and goes on to mention a time shift, "four hours later", which is clearly outwith the normal time span of such an event (usually about 1 hour 45 minutes). It might, of course, be argued that events such as football matches are biased towards males. However, the time shift chosen (4 hours) is of such a length that it should intuitively seem to be too long for playing any game. Most games, other than perhaps something like marathon running, would not take so long; any game or sport lasting 4 hours would demand considerable stamina.

While all texts were constructed in the manner described above, with the intention of clearly signalling a shift in scenario, the following study attempts to measure more precisely how children measure the various situations in terms of time span.

Given that the deaf subjects are likely to have less enriched experience of life due to their communication difficulties, it is possible that they will generate time estimates which are substanatially different from those of the hearing group, for at least some of the scenarios used in the on-line and signing studies. It might be expected that the hearing subjects would estimate time spans similar to those given by the adult raters.

9.2.2 Subjects:

The experimental group comprised 19 deaf subjects with hearing losses of 70 dB or greater in the better ear averaged over the four frequencies within the range 500Hz-4KHz. Losses ranged from 82 dB to 115 dB, with a mean loss of 99 dB (standard deviation 9 dB). All were within the average range of ability. Ages ranged from 12 years 9 months to 17 years 0 months, with a mean age of 14 years 6 months (standard deviation 14 months). The comparison group comprised 46 hearing children. Ages ranged from 12 years 3 months to 13 years 9 months, with a mean age of 13 years 4 months (standard deviation 4 months).

9.2.3 Method and Materials:

The subjects were given a sheet of paper with a list of the various scenarios used for the temporal items of the on-line study. These were not the titles of those texts but rather a more general and fuller description of each scenario. Along side each was space for the subject to estimate the time.

The instructions given were:

> " Different things take different times to do. For example,
> (a) eating a bag of crisps would take about 5 minutes.
> (b) going to the dentist might take half an hour.
> (c) watching a video could take 2 hours.
> How long do the following things usually take? ...."

9.2.4 Results:

Mean times for each scenario for deaf and control groups is shown in Table 9.1 below, along with the time shift used in the text:

Table 9.1: Estimates of scenario duration

Scenario	Mean Times		Range (SD)		Time in Text
	Deaf	Hearing	Deaf	Hearing	
Church	63min	52min	10-150min (34min)	40-60min(8m)	3hr
Funfair	47min	6min	1-180min (50min)	2-15min (3min)	30min **
Ferry	1.9hr	2.3hr	5m-8hr (1.7hr)	25min-6hr (1.4hr)	3hr
Bath	23min	28min	5-60min (16min)	10-60min (10min)	2hr
Friend's house	1.4hr	1.8hr	2min-3hr (0.8hr)	10min-4hr (1hr)	5hr
TV Repair	48min	37min	1min-4hr (56min)	10-90min (20min)	3hr
Football	1.6hr	1.7hr	5min-5hr (1.1hr)	1-3hr (0.4hr)	4hr
Swim at seaside	38min	62min	3min-2hr (31min)	5min-4hr (65min)	2hr
Train to school	54min	12min	15min-2hr (40min)	1-40min (9min)	2hr
Dance	2hr	2.1hr	5min-7hr (1.9hr)	10min-4hr (1hr)	5hr
Shop window	12min	4min	3sec-1hr (15min)	5sec-10min(3m)	30min
Train moving	25min	4min	3-85min (26min)	10sec-20m (4m)	30min *
Letter	48min	26min	5min-3hr (54min)	10-60min (12min)	1hr *
Film	2.1hr	2.1hr	7min-4hr (0.8hr)	30min-4hr (0.7hr)	5hr
Tea/cafe	26min	24min	4-90min (24min)	10-60min (11min)	1hr

Note: min = minute(s); hr = hour(s); SD = standard deviation of sample
** greater than the time used in text; * borderline

9.2.5 Discussion:

All the mean time estimates by the control group are comfortably within the time shifts used in the texts of the on-line study. Only for two scenarios does the range of estimates fall beyond that used in the text. However, for the deaf group, the mean time estimate for one scenario (*Funfairride*) was well over the time shift in the text while two others (*time between getting on a train and it's moving off*; and *writing a letter*) could be considered borderline. For three scenarios (viz, *having a bath*; *taking tea in a cafe*; and *watching a film*) the range of estimates for both groups were very similar. Apart from these, the deaf always had a much wider range of estimates for each situation.

Perhaps the wide ranges estimated by the deaf indicates the limited extent of their experience of real life events or that their experiences are less consistently labelled linguistically by the deaf child.

It is unusual that the deaf group, while smaller in number, shows greater variability than the hearing group. In order to examine the variability within each group, the time estimates for each scenario were compared with the time shifts used in the texts for each subject using the Sign test. For the hearing group, this showed that the estimates by all subjects were significantly less than those used in the texts. For the deaf group, similar results were found for all subjects but one. This particular subject had a wide range of estimates which in six of the fifteen narratives were beyond those used in the texts.

When the Sign test was used to compare the average time estimates for each scenario by the deaf group with those of the hearing group, no significant difference was found.

9.3 Eliciting of scenarios for causal texts

9.3.1 Rationale

Just as for temporal texts, it is possible that the deaf and hearing children invoke substantially different scenarios for causal texts.

The purpose of this part of the study was to investigate what constituted various scenarios for the groups of subjects. If given a free hand to describe various scenarios, what characters, activities, objects, etc would they include?

If the deaf do have a more impoverished representation for certain situations, this should be reflected when they attempt to describe entities associated with each scenario. It is hypothesised that fuller and wider ranging descriptions will be produced by the hearing subjects.

9.3.2 Method

Subjects were given a number of sheets of paper at the top of which was printed a title. The rest of the sheet was left blank. Titles were those of the causal texts used in the online study. Subjects were asked, for example:

> "Please list as many things you can think of (people, objects, things you might do) that are usually to do with:
>
> EATING DINNER: "

It was possible to combine together scenarios for more than one textual passage. For example, the scenarios for "*Waiting at the bus stop*", "*Going home*", and "*Going to* *school*", were incorporated into one sheet, entitled "*Waiting for a bus*". However, there were still twelve scenarios to be completed. It is possible that subjects could tire towards the end and so give fewer responses than to earlier items. In order to avoid such a systematic error, the response sheets were randomised in each bundle, so that no one scenario or grouping of scenarios was coming consistently at the end.

9.3.3 Subjects:

The experimental group comprised 15 deaf subjects, aged from 12 years 9 months to 15 years 11 months, with a mean age of 14 years 3 months (standard deviation 12 months). All were within the normal range of ability. Hearing losses in the better ear over the frequencies 500 Hz-4KHz ranged from 85 dB to 115 dB, with a mean loss of 100 dB (standard deviation 9 dB).

The comparison group comprised 36 hearing subjects. Ages ranged from 12 years 8 months to 13 years 9 months, with a mean age of 13 years 4 months (standard deviation 3 months).

Some subjects participated in all three parts of this study. However, the various parts were administered on different occasions, each several days apart. Most importantly, sections 9.1 and 9.2 were administered before section 9.3 which is described next. Section 9.3 uses a technique which requires subjects to continue a story by adding a sentence to part of the texts used in the online experiments. It was necessary to do this after subjects had given their own time estimates or continued a causal text in their own fashion. If they had been exposed to section 9.3 first, it would have been possible that their own estimates etc would have been contaminated.

9.3.4 Scoring of responses:

A content analysis was carried out on all responses. Each separate entity properly associated with a particular scenario which was mentioned was totalled A certain degree of subjectivity is unavoidable but subjects' written responses were grouped under the following four broad categories:

People : This would include all people normally associated with that scenario, for example, mum and dad and other family members would normally be associated with "*eating dinner*" since dinner is usually eaten at home. However, if dinner was eaten at a restaurant then it is reasonable to mention a waiter. Any of these entities would be accepted in the scoring.

Considering another scenario, "*ambulance at scene of accident*", it is reasonable to expect firemen, ambulance personnel and the police would be associated with such a situation. Other people likely to be mentioned would be the driver, and possibly pedestrians or witnesses. Family members would not usually be included (although they could be, of course, if they happened to be travelling in the car, but not as a general rule). It was noticed that the deaf tended to include family members more often than hearing subjects in various scenarios where the connection was tenuous. Such entities were not included in the analysis.

Objects : This is more straightforward and includes all objects which could normally be associated with various scenarios. For example, "*eating dinner*" is likely to include knives, forks, plates, table, food, etc. The "*car crash*" might include tow truck, ambulance, bandages, broken glass, etc.

Actions : This includes all actions, activities etc normally associated with a particular scenario. For example, eating, chewing, and talking might be mentioned for "*dinner*", while rescue, help, or scream could be mentioned for "*car crash*".

Descriptive: Credit was given for additional, appropriate descriptive details associated with a scenario, eg for,say "*A Walk in the Country*", adjectives like "muddy", "cold and miserable", or "warm and sunny" would be scored in this category.

9.3.5 Results:

The average number of entities mentioned per scenario is shown below for each group in Table 9.2. The lower figure in brackets is the percentage of the total number of entities which were in that category.

Table 9.2: Mean number of entities in each category mentioned for causal scenarios

	Persons	Objects	Activities	Descriptive	TOTAL
Deaf	0.86	8.33	1.27	0.74	11.31
	(8%)	(74%)	(11%)	(7%)	(100%)
Hearing	0.72	4.81	1.60	0.45	7.60
	(9%)	(63%)	(21%)	(6%)	(99%)

It can be seen that the deaf on average write down 49% more entities associated with the listed scenarios than do the controls (an average of 11.31 items per scenario compared with 7.60). This might be thought surprising considering the well documented difficulties the deaf have with written work and what might be thought of as a consequent reluctance to put pen to paper. None of the administrations of the tests (by class teachers) was time limited; subjects were given enough time to write as much or as little as they wanted.

There are remarkable similarities in the percentages of persons and descriptive terms listed by each group. However, while both groups tended to list objects most frequently, the controls mentioned fewer objects than the deaf group but did list almost twice as many activities.

There is large variability in the responses from both the deaf and the control group. However, the range of responses is very similar. For example, for "*Eating Dinner*" most subjects in both groups tended to list items of food, pieces of cutlery, family members who would be present, etc. However, some subjects in each group gave an extended list. For instance, both groups mentioned waiters/waitresses, cafes and so on. One person in the deaf group included going to the butchers to get meat for the dinner, while a few in the control group described setting the table and washing up afterwards.

Eating dinner would be regarded a fairly normal, mundane activity compared to, say, horse riding which fewer people are likely to have experience of. Surprisingly, even here there were similar responses from both groups. Apart from the obvious responses (eg horse, saddle, ride, jump, etc), a few subjects in each group mentioned more specialised terms associated with this activity. For example, the deaf listed people such as grooms and jockeys, equipment such as jodhpurs, bit, and reins. The controls listed people such as jockeys and instructors, equipment such as bridle and whip, and actions such as canter, trot and gallop.

The relative richness or otherwise of the various scenarios as generated by each group is discussed later in Section 9.5.

9.4 Continuation of stories

9.4.1 Rationale of Study:

The purpose of this study was to investigate how subjects would continue a set of short stories (the texts used in the earlier on-line and verb studies) given the first part of the stories, up to and including the critical sentence, ie the sentence where the experimental manipulation takes place. For the temporal items in the on-line study this manipulation involves stating a time shift and, by implication, a new scenario. For the causal items, the critical sentence implied some event or situation which would take place as a consequence of what was stated. For the verb items, the critical sentence used a verb which should imply the use or inclusion of a particular instrument or object.

Would subjects continue the stories in a similar manner to the originals or in some other way? Would the story structure constrain their responses or suggest a plethora of alternatives?

9.4.2 Method and Materials:

Each of the subjects was given a booklet containing each of the texts used in the on-line experiment along with those used in the verb experiment. The passages were randomly mixed. The texts from the on-line study were presented with titles (as in the original) while those from the verb study were not (again, as in the original). Beneath the last line of each text was a blank line. Subjects were asked to write the next line for each story.

Examples of the materials are given below:

GOING TO SCHOOL

John was on his way to school.	CAUSAL:	Implication is that John
He went by bus.		is late for school.
He had to wait a long time.		
....		

IN CHURCH

It was Sunday. Ian went to church. Three hours later he saw his sister.	TEMPORAL:	Implication is that when Ian saw his sister they were not in church.
Peter was in the garden. He was digging up some flowers.	VERB:	The implication is that Peter was using a spade.

9.4.3 Subjects:

The comparison group comprised 50 children coming to the end of their first year of secondary education in a large comprehensive school in west of Scotland. Ages ranged from 12 years 9 months to 13 years 9 months, with a mean age of 13 years 3 months (standard deviation 4 months). They were picked randomly from two typical classes.

The deaf group was rather small, only 8 subjects. They met the conditions mentioned above with regard to hearing loss, level of ability, age of onset of deafness etc. Losses ranged from 85 dB to 115 dB, with a mean loss of 101 dB (standard deviation 9 dB). Ages ranged from 13 years 5 months to 15 years 11 months, with a mean age of 14 years 6 months (standard deviation 14 months).

9.4.4 Analysis of Responses:

Each story can be continued in several ways. Using the classification described below, subjects responses were scored according to these categories and converted to percentages. The analysis of responses is, to an extent, subjective, but examples of how responses were scored are given in the next section.

Temporal Items:

The story could either continue logically or not. If the story continued logically then what was written can refer to either the "old" scenario, ie the scenario in the first part of the text, or it could refer to the "new" scenario, ie the scenario suggested by the time shift. Alternatively, the story could continue logically but be ambiguous in relation to the scenario operating at that stage.

Causal Items:

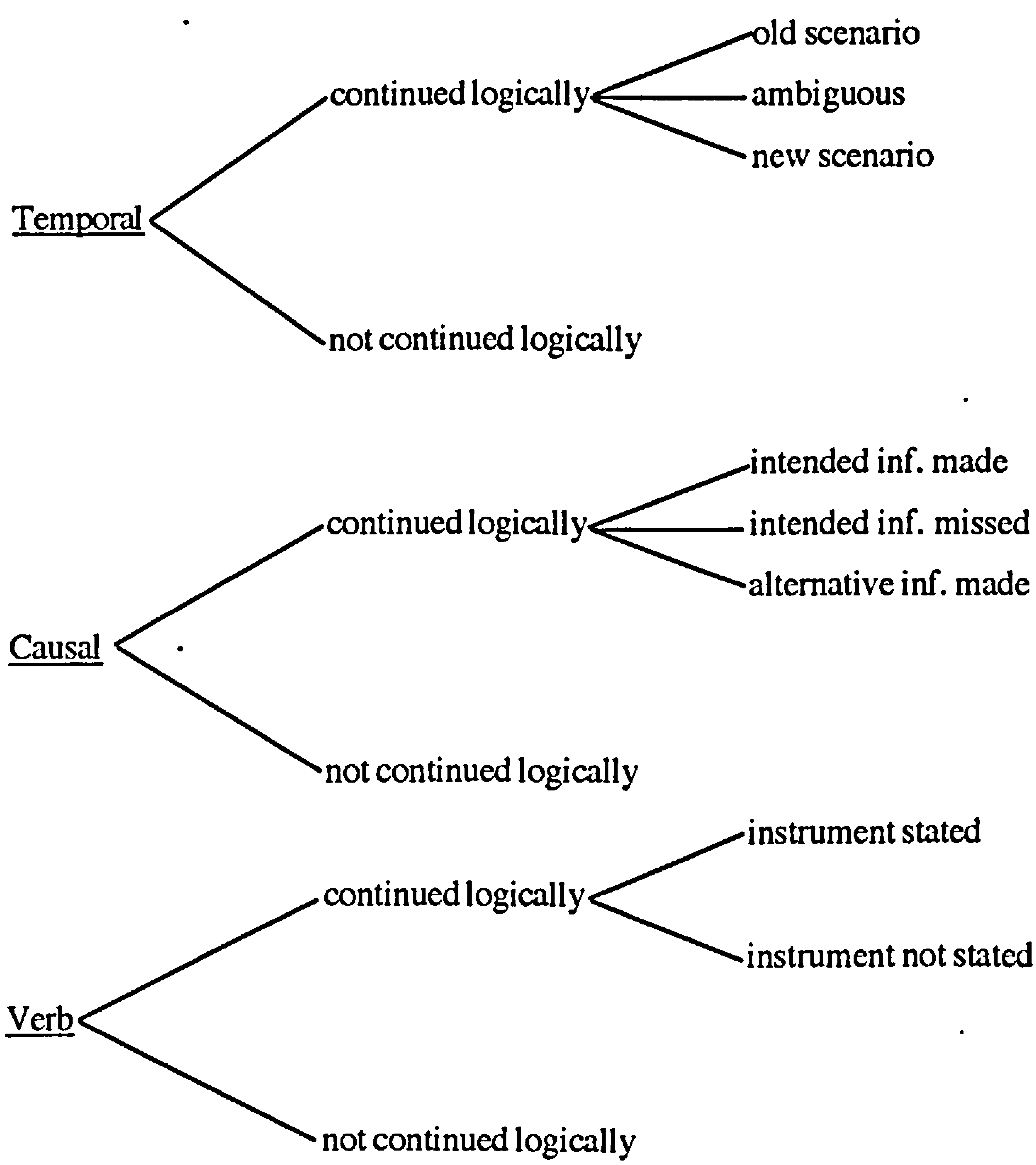
As above, the story can continue in a logical manner or it may not. If the continuation is logical then what is written can show that the intended inference is indeed made. It could also show that this inference is missed. A third possibility is that an alternative inference to the one intended is made; this can be so such that it is not possible to state whether the originally intended inference is made or not.

Verb Items:

Again the story can be continued logically or not. If it is continued logically then the instrument associated most commonly with that verb (as conjectured in the original experiment) can be mentioned in the continuation or it can be omitted.

These alternative routes for continuing a story are shown diagrammatically in Figure 9.1:

Figure 9.1: How stories could be continued



9.4.5 Results:

From the control group of fifty, only 44 subjects were used in the analysis. The remaining 6 were discarded mainly because these subjects had not completed the task and had many blank sections on their returns.

For the deaf group, the returns from half of the subjects had to be discarded, again because of failure to complete the task. Also a large portion of the written responses of these children contained mere repetition of parts of the text, with no attempt being made to continue the story. Therefore because of the small number of subjects, these results must be interpreted with caution.

Examples of subjects' responses for temporal and causal texts and how they were scored are given below. The scoring of verbal texts is much more straightforward; the instrument or entity is either mentioned or it is not.

<u>TEMPORAL</u>	"In Church"
Old scenario:	"And went to the church hall for a cup of tea."
Ambiguous:	"He said hello."
New scenario:	"Then they went to the shops."
<u>CAUSAL</u>	"Going to School"
Original inf. mentioned	"And he was late for school."
" inf not mentioned	"But soon the bus came."
Alternative inference	"The bus had a flat tyre."

The results for the both groups are shown in Table 9.3 below.

Table 9.2 Continuation of Stories

		Cont	Deaf
	N =	44	4
	Continued logically - new scenario	26%	23%
TEMPORAL	Continued logically - old scenario	23%	38%
	Continued logically - ambiguous	51%	38%
	Not continued logically	0%	0%
	Continued logically - inf. mentioned	44%	48%
CAUSAL	Continued logically - inf. not "	4%	5%
	Continued logically - alternative inf	50%	43%
	Not continued logically	2%	4%
	Continued logically - instr. named	14%	11%
VERB	Continued logically - instr. not named	86%	86%
	Not continued logically	1%	3%

9.4.6 Discussion:

From the above results it can be seen that for the control group, the vast majority of responses are directed towards continuing the storyline in a logical manner. This offers cohesion to the text. For those very small number of instances where this did not take place, it did seem that the subjects themselves made some kind of connection but it was not obvious to the scorer.

It can be seen that for temporal items, the control subjects tend to generate a storyline which is ambiguous as far as the scenario which is operating at that point is concerned (about 50% of the time). What is perhaps most surprising is that where a particular scenario can be identified in the continuation, the subjects chose to sustain the old scenario almost as often as they move to a new one (both about 25% of the time) in spite of the flagging of a time shift in the critical sentence. When they are asked to estimate the times that the various activities involved in the temporal texts would take (see above), in every instance they estimated an average time well within the time shift used in the actual text.

For causal items, the controls continue most often with an alternative inference to that in the original text (about 50% of the time) but they do mention the intended inference in almost as many instances (44% of the time).

For the verb items, the controls tend not to mention the associated instrument or object (only about 14% of the time). This would seem to be the natural thing to do. It would seem to be more essential to continue with a logical story line than to be sidetracked into talking about a narrow and specific aspect related to the action just described. These results would seem to be consistent with the findings of O'Brien et al (1988). They found that subjects will generate simple elaborative inferences if they are given sufficient context. Hence in the temporal and causal texts used above, subjects are given a title and three lines of text and asked to continue with a fourth line. This amounts to a fair a degree of context which can constrain how subjects will respond. However in the verb texts, subjects are not given a title to orientate them and are only given two lines of text; they are asked to continue with the third line. This is insufficient context to constrain subjects and they have more freedom to continue as they wish. Most opt for a continuation of the narrative which progresses the story line and does not remain static, describing the instrument used in the last given sentence.

This would suggest one possible reason why subjects in both the deaf and control groups in the verb study described earlier replicated Singer's pattern of results rather than those of Sanford & Garrod.

9.5 Comparison of quality of scenario generation with performance in the on-line study

9.5.1 Preamble

During the on-line study it was mooted that perhaps the deaf had particular difficulty with certain texts. They appeared to use similar mechanisms during the inferencing process as the hearing subjects but this might have failed when the scenario depicted in the text was beyond their experience.

The estimates of duration for temporal texts and the generation of facets of causal scenarios noted above was an attempt to measure how enriched each situation was for both groups. Some scenarios were better than others and so a comparison was made of subjects' attempts at eliciting the various scenarios with their performance in the on-line study.

For the temporal texts, all mean time estimates by the control group were well within the time shift used in the texts. On two scenarios the range of estimates exceeded the time used in the text, but even then, the majority of subjects were still within the time shift.

The deaf showed a wider range of time estimates than the hearing subjects generally, with the range often going beyond the time shift used in the texts. However, if mean times are considered, the deaf overestimated on one scenario (the funfair ride) and were borderline, although just with the time shift, for two other situations (a train moving off and writing a letter). As a comparison, the difference between the reading rates for the deaf of explicit and implicit conditions of Sentence 3 (where the critical manipulation takes place) were calculated. One might expect the three situations where the deaf have some difficulty in estimating the time shifts accurately to be those texts were the implicit reading rate is slower than the explicit. If the deaf draw on a version of a scenario which is impoverished, then on-line processing should break down and either a successful inference does not take place or the subjects have to use an alternative mechanism to comprehend which will involve more time. The differences in rates is shown in Table 9.4 below:

TABLE 9.4:

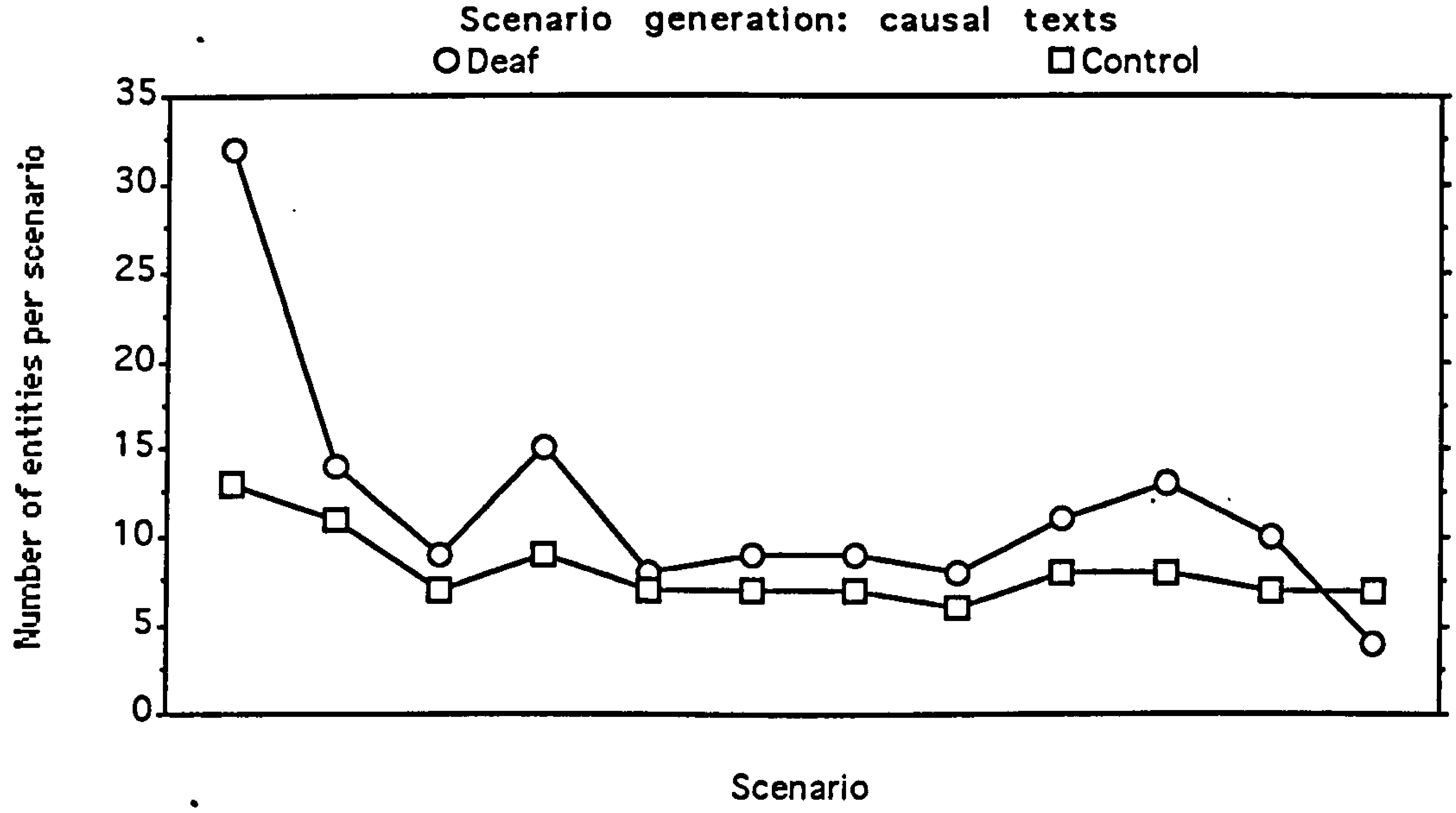
Temporal Scenario	Difference in Reading Rates (in syll/sec): (implicit minus explicit)	
Going to church	9.909	
Funfair ride	2.157	*
Ferry (going on hols)	0.276	
Bath	3.003	
Visit friend's house	-2.768	
TV repair	-12.702	
Football match	- 6.120	
Swim at seaside	3.663	
Train home from school	3.378	
Going to a dance	- 3.217	
Looking in shop window	4.570	
Train moving off	-0.270	*
Writing a letter	-7.428	*
Watching a film	3.812	
Taking tea in cafe	1.020	

No consistent pattern was found between reading rates and those scenarios which were less enriched for the deaf group on average.

If a similar comparison is made between the time estimates made by the deaf and how accurately questions were answered, again no consistent pattern is found. Only one scenario (time between boarding a train and the train moving off) was found to have both a low accuracy score and a poor time estimate.

Using the average number of entities mentioned for each scenario that were obtained in section 9.2 above, a similar comparison was made for the causal scenarios for both the deaf group and the hearing group. Firstly the relative richness of the scenarios was compared (this is shown in Figure 9.2 below). The deaf generate more items for each scenario than the controls but the pattern of relative richness/impoverishment is remarkably similar. Although the deaf generate more items than the comparison group, contrary to expectations given the well documented evidence of their difficulty with written output, this is a quantitative measure and may not reflect the true depth of the cognitive model held for each scenario. From section 9.2 it was noted that the hearing group tended to mention a wider range of items (eg more activities and more people) associated with the various scenarios than the deaf. This might be a small indication that although they put less down when listing such items, they reflect a better qualitative approach than the deaf. A Pearson product-moment correlation was calculated giving $r= 0.8760$, $df=10$, $p< 0.01$. Thus the pattern of relative difficulty for each causal text/scenario is the same for both deaf subjects and controls.

Figure 9.2: Comparison of extent of enrichment for scenario generation for deaf and hearing subjects



Key to scenarios:

1 Eating dinner, 2 In the swimming baths, 3 Skating on thin ice on a pond, 4 Travelling by car, 5 An ambulance at the scene of an accident, 6 Waiting for a bus, 7 Horse riding, 8 A hungry dog in the kitchen, 9 Making sandwiches, 10 A long drive in the country, 11 Walking across a field on a rainy day, 12 Shopping with purse at top of shopping bag.

If those texts where the implicit condition produced a slower reading rate for the critical sentence than explicit, and hence where the inference seemed to be made more slowly, are plotted on the graph in Figure 9.1, it can be seen that they correspond with 7 out of eight of the most poorly elicited scenarios. Indeed all seven are below the mean number of items generated for all causal scenarios. However, if the reading rates for the critical sentence under the implicit condition are compared with the number of items generated for each causal scenario, the Pearson product-moment correlation is positive but fails to reach significance ($r = 0.385$, $df = 10$, $p > 0.05$). Therefore, although the reading times for the implicit condition of the critical sentences are slower for those scenarios which are most impoverished, the difference is not great enough to reach statistical significance.

Similar comparisons for the number of questions answered correctly, and the reductions in reading rate for the questions in the on-line experiment with the richness of each scenario fail to reach significance.

9.5 Discussion

It would seem that the deaf show a broadly similar pattern in their scenario generation to their hearing peers. For temporal texts, all but one scenario is within the time shifts used in the on-line study, with another two borderline but still within the time shifts. For causal texts, the deaf generate more items, actions, etc associated with the various scenarios than the hearing subjects and this may seem unexpected given all the evidence of how poor the deaf's written output tends to be. It may be that although the controls produce less, nonetheless what they produce is of better quality. The hearing group tend to mention a broader range of items, eg they mention more "activities" and "people" than the deaf who tend to mention more inanimate objects. However there is a large degree of overlap between the groups. There is a remarkable similarity in the relative difficulty of the causal scenarios, being almost identical for both groups.

With the deaf group it is for those causal scenarios where they produce fewest items in their listings that they seem to have longer times in the on-line study under implicit conditions. Given the greater variability within the deaf group, it is possible that some scenarios cause particular difficulties for some subjects. Graesser. & Kreuz (1993, p150) argue that it is critical for readers to possess a sufficient knowledge base if they are to be able to generate inferences on-line. This might suggest, therefore, that the deaf subjects' unfamiliarity with these scenarios leads to a situation where an inference cannot be drawn on-line and the subjects have to use secondary processing.

Chapter Ten:

The Eyetracker Study

10.1 Introduction:

Although the results from the previous studies appear to follow a consistent pattern, it could be argued from the evidence presented in the forth chapter that the methodologies used, like all approaches in this area, have certain limitations. Keenan, Potts, Golding & Jennings (1990) are of the opinion that the only way to overcome such restrictions is to use more than one methodological approach. The strengths of one approach will compensate for the deficiencies of another and vice versa, so that if a range of approaches offers converging results, there can be greater confidence in the findings.

The main experimental paradigms used thus far have been those of cued recall and on-line reading or response times. These have the disadvantage that they only enable testing of inference generation somewhat downstream from where this actually occurs. The sentence reading time is a measure for the whole sentence, not what is happening from moment to moment as a text is read. The memory probe occurs even longer after the inference may have been made.

In the present study we use the approach of tracking the movements of the eyes as they scan over text, noting which words are actually fixated, in which order and the length of time that the eyes dwell on each word. The process is sufficiently accurate to consider which letter (or letters) within a word is fixated. This methodology is non-intrusive and is considered by some (eg Rayner & Pollatsek, 1989, p111) to be the most reliable, offering the richest insight into the processes occurring during the act of reading textual material.

A fuller discussion of what is generally known about eye movements during the reading process and what measurements are particularly useful is given in Section 4.10. However, it is generally accepted that measures of gaze duration (that is how long words are fixated on the first sweep) are good estimates of speed of lexical access, while regressive movements reflect higher order processes (Rayner & Pollatsek, 1989).

Pollatsek & Rayner (1990), however, urge a certain amount of caution when interpreting eye movement data. Fixation times, for example, may not reflect a totally accurate estimate of speed of lexical access. Some studies, eg Rayner, Well, Pollatsek, & Bertera (1982) have shown that parafoveal processing can take place. That is, as the eye fixates one word, partial information about the following word can be extracted on that same fixation.

This technique has been used successfully to study automatic inferencing in adults. Clark & Sengul (1979) found that if the surface distance in a text between an anaphor and its antecedent increased, then the time to read the sentence containing the anaphor increased. Ehrlich & Rayner (1983) used eye movement data to show that the search for the antecedent happens immediately the anaphor is encountered. They discovered longer fixation times on the anaphor as the surface distance increased. They found that if there were three or more intervening lines of text, then the search continues beyond the anaphor, with increased length on the two following fixations, indicating an increasing effort at resolution.

More recently, Garrod, Freudenthal, & Boyle (1994) used eye-tracking studies to explore the rather complex interaction of pragmatic aspects of a text, current focus of the reader's attention (eg who the main character is), and the lexical specificity of the pronoun (ie its gender) in the resolution of anaphors.

Eye movement studies of children's reading are rarer. A study by Taylor (1965) showed that as children get older their fixation times on words decrease, their saccades become longer, and the number of regressions they make is reduced.

There appear to be few studies of the eye movements of deaf children as they read. The only study traceable in the literature is by Wood, Wood, Griffiths and Howarth (1986, p109). These authors found that deaf subjects fixate each word in turn as they read through a text. This is in contrast to the findings of Beggs, Breslaw, and Wilkinson (1982), quoted in Wood et al (1986), that hearing children tend to skip over certain words or groups of words. This is rather surprising given the evidence from recent work (eg Rayner, 1995) that fluent readers tend to fixate the majority of words on a page. However, the discrepant results of Beggs et al (1982) might be explained by their use of less sophisticated eye tracking technology than is currently available.

Therefore, the use of an eyetracking paradigm in the present study is a departure from the traditional approach. The investigation is, to a large extent, breaking new ground and is, *ipso facto*, of the nature of a pilot study.

The purpose of this study is to re-examine deaf and hearing children as they draw temporal and causal inferences, using the same textual materials as in the on-line and signing studies. In those studies the deaf were found to consistently read more slowly than the hearing controls. A study of eye movements could determine whether the deaf fixate words for a longer than the hearing subjects. Do the deaf have slower lexical access?

Another problem with the on-line study was that it relies on establishing the null hypothesis to prove that elaborative inferences have occurred. Since the deaf had a much wider range of times when estimating various temporal scenarios, it may be that they are making bridging inferences, but that the additional time needed is occurring when they read the phrase containing the time shift and they read the remainder of the sentence more quickly. Since the on-line study measures overall times for reading whole sentences, it would be insensitive to such mechanisms.

Similarly, since eye movements give a moment to moment picture of reading behaviour, it should be possible to note which parts of the text are difficult for the reader and cause him/her to make regressive sweeps.

10.2 METHOD:

10.2.1 Subjects:

The experimental group initially comprised 6 severely or profoundly deaf children each with an average hearing loss in the better ear of 70dB or greater, all functioning within the normal range of intelligence, with no additional handicap and with hearing parents. All were prelingially deafened. No subject wore glasses and none had any medical condition of the eye. However, not all the subjects could be tracked satisfactorily and two of the six had to be dropped. Subjects who had long eyelashes, for instance, proved to be poor trackers as the beam would often lock onto the eyelash instead of the eye. For the four remaining, the average hearing loss was 104dB, ranging from 101dB to 106dB, SD = 2dB. Ages ranged from 14 years 1 month to 15 years 8 months (mean = 14 years 7 months, SD = 9 months). These four subjects were well motivated pupils who were considered by their teachers to be better readers than the average deaf child of that age.

The comparison group initially comprised 16 children chosen at random from the second year of a local mainstream secondary school. All were considered to be within the average range of ability. None wore glasses and, as far as medical records showed, none had any medical condition of the eye. However, as with the experimental group, not all subjects could be tracked satisfactorily and so 6 had to be dropped. The remaining 10 subjects had and average age of 13 years 6 months (SD = 3 months).

10.2.2 Materials:

The materials used were those of the online experiment described in Chapter 2. The texts were presented in the same order and had the same number of practice items.

Each text was presented in its entirety with title, four lines of text and question, spaced as shown below. An example of a temporal and a causal text are given.

TITLE
First line
Second line
Third line
Fourth line
Question

<u>Temporal</u>

FOOTBALL

James went to the football match.
It was a good game.

After the game he saw his sister.	Explicit form
Four hours later he saw his sister.	Implicit form

He kissed her.

/Did James kiss his sister after the game?

Causal

A DAY IN THE COUNTRY

Tom was driving in the country.
The car stopped suddenly at the side of the road.
He noticed that the petrol tank was empty. Explicit form
He took an empty can to the garage. Implicit form
He was cross about all this.

Q/Did the car run out of petrol?

When subjects had read the material, they pressed a button on the computer keyboard which then cleared the narrative from the screen and presented the question, to which a yes/no push-button response was required.

This presentation is different from that in the on-line experiment where each line is presented singly. The present situation is more "natural" in that subjects can retrace their steps if difficulties arise as they read the text.

All textual material was presented in lower case with capitals being used when appropriate, ie proper names, beginnings of sentences).

10.2.3 Design of Study:

This was similar to the online experiment and is a 2X2X2 factorial repeated measures design. The between subjects factor was subject group and had two levels, deaf and control. The first within subjects factor was inference type and had two levels, temporal and causal. The second within subjects factor had two levels, inferred and explicit.

Subjects in both the deaf and the control groups were split into two sections, each section receiving one trial containing half explicit and half implicit items randomly distributed. The second trial contained the reverse conditions of materials, ie explicit texts in the first trial became implicit in the second.

10.2.4 Procedure:

A Stanford Research Institute Purkinje Dual Image Generation 5.5 Eyetracker was used in this study. It had an angular resolution of 10' arc. This is linked to a Vanilla 386 personal computer which controls the experiment and records data. Figure 10.1 (below) shows the layout of the various pieces of equipment.

Subjects were seated before a video monitor which was located at a distance of 70 centimetres. Their heads were kept in a constant position by use of a chin rest and a support for the forehead. A strap with velcro ties around the back of the head minimised head movements. Consideration was given to using a bitebar but as some subjects felt.that this was distasteful, it was not used.

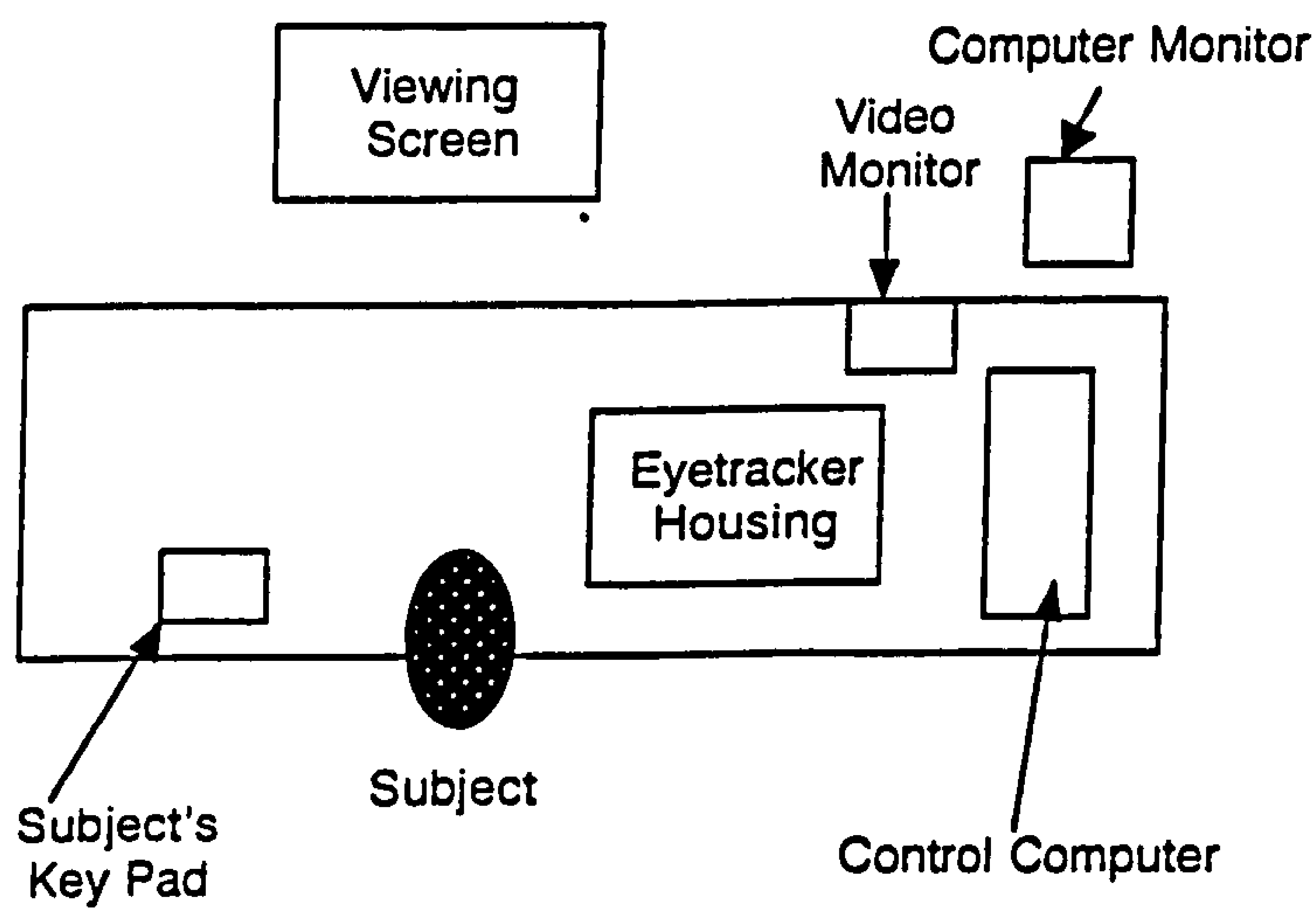


Figure 10.1: Layout of Eyetracker Equipment

To the right hand side of the subject's head was the eyetracking equipment itself. This comprised a device which shone a narrow beam of infrared light through a series of mirrors onto the pupil of the subject's eye, and another device for detecting the reflected image. Since the light is within the infrared range, it is not visible to the human eye and so subjects are unaware of it. The whole of this equipment is mounted on a movable platform such that as the subject's eye moves (to look at text on the monitor), the infrared device can move also to keep the beam locked onto the subject's pupil.

The eyetracker is fine tuned and calibrated as outlined in Appendix A6.

The instructions given to subjects are noted in Appendix A7. These were given to the hearing subjects by the experimenter but were signed and explained to the deaf subjects by their teacher.

The principle on which the equipment works is that when light is shone onto a human eye, the light is reflected in a series of images at zones of optical discontinuity within the eye (the Purkinje-Sanson images). The first and brightest image, reflecting about 2.5% of the incident light, is from the anterior of the cornea (Trevor-Roper,1955); the second from the posterior surface of the cornea (0.024% of incident light); the third and fourth from the anterior and posterior respectively of the lens surfaces (both 0.036% of the incident light). Only the first and fourth images are used by the eyetracker. Since the forth image shows no parallactic movement in relation to the iris as the eye rotates, the relative positions of the first and forth images can be used as an indicator of where the eye is pointing. During the calibration phase, a series of dots whose coordinates are known by the computer are displayed on the monitor. The subject is asked to focus on each dot in turn and then make a keypress. This allows the computer to feed the detected positions of both Purkinje images for each known coordinate position into a complex algorithm so that during the experiment the computer can calculate the exact coordinates being fixated by the eye as the text is being read.

In the experiment proper, before each text is presented, a small white dot appears at the top left hand corner of the screen. This focuses the subject's attention to this point and when the computer detects that the coordinates being fixated correspond to the white dot, then and only then is the text presented (the first word of the text appears just to the right of the white dot). This ensures that subjects are looking at the top left of the screen and are ready to start reading from the beginning of the text. This avoids the unfortunate situation where a subject might look first at the middle of a text and then jump back to the beginning. In this case it would be impossible to say whether the subject had absorbed aspects of the text up to this point. In this case it would be difficult to know which fixations were first fixations and which would be classed as second pass fixations.

10.3 Results:

As mentioned above, it is possible from the recorded data to tell which words are fixated and in which order, which letter(s) within words are fixated, and how long each fixation lasts. From these data can be calculated the mean fixation duration, how many regressions are made, and the mean number of fixations per word (this last giving an estimate of perceptual span).

However, the data as gathered above has to be "cleaned up". This process is referred to as eyewashing. As the eyetracker scans the reader's eye, there is sometimes drift in the vertical plane. This is because after making a saccade, the landing position of the eye is frequently just above or just below the line. This is due to the imperfect control of the ocular motor system. While the horizontal position of each eye fixation remains true, the vertical position often moves up or down by one or two lines and this has to be corrected before the data can be used for analysis.

Correction of drifting in the vertical co-ordinates is done manually. Multiline text, such as that used in the present study, requires subtle evaluation. One way of doing this is to monitor the horizontal positions of the eye's movement. An example to illustrate this is given below:

12345678......etc ....32 HORIZONTAL CO-ORDINATE

1 xxxx xx xxxx xxx xx

2 xx xxx xx xxxx xxx xxx xxxx xxxx

3 xx xx xxxx xxx

A series of fixations which may start on line 2 may then suddenly change to vertical co-ordinates for line 3. However, if the horizontal co-ordinates continue in an increasing sweep up to position 32 or thereabouts, then this is most likely a drift, and the true vertical fixation remains on line 2, and we do not have any real movement to line 3.

Occasionally a fixation may be totally inconsistent and may be regarded as "rogue". These are usually infrequent and are ignored in the final analysis.

Another function of the eyewashing process is to combine very small fixation durations, usually of less than 80 milliseconds. This is done automatically by the computer programme.

The time required to complete this eyewashing process varies from subject to subject but on average takes between about two and a half and five hours, depending on the complexity of the data. Raw data files could contain as much as 43,000 bytes of information. Therefore the process is rather time consuming.

It should also be noted that the computer programme used to analyse the raw could not distinguish between those questions which were answered correctly and those which were not. This is because the programme was designed for use with undergraduate subjects who would be expected to comprehend most of what they read. As was evident from the on-line study, children are less successful. The results are for all texts whether they were answered correctly or not, ie whether the inferences were understood or not. Consequently, any significant findings are likely to be diluted since the analysis will include data from certain passages where the inference was not made.

10.3.1 Analysis of complete texts:

The measures noted below are calculated for complete texts including titles, exactly as the subject sees them.

10.3.1.1 Mean fixation durations:

All fixations are included in these calculations, first fixations, second fixations, and second pass fixations (ie first fixation duration, gaze duration and total viewing time).

The results are shown below in Table 10.1:

TABLE 10.1. Mean fixation durations (msecs):

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	277	262	263	283
Hearing	229	226	231	228

These results were analysed by analysis of variance, the between group factor being subject group (deaf vs hearingl), the first within factor being inference type (temporal vs causal), and the second within factor being inference (inferred or explicit).

Collapsing across texts and treating subjects as random variables we find that the deaf fixate for significantly longer than hearing subjects (271 msec compared with 228 msec), $F_{1,12}$= 10.017, $p<0.01$. There was also a significant interaction between inference type and explicit/implicit, $F_{1,12}$= 9.898, $p<0.01$. Post-hoc Scheffé tests showed that causal-implicit items had significantly longer fixations (255 msec) than temporal-implicit (244 msec) at $p<0.01$. There was also a significant 3 way interaction, $F_{1,12}$= 10.846, $p<0.01$. Post-hoc Scheffé tests showed that for deaf subjects, under the implicit condition, causal items were significantly longer than temporal ($p<0.05$). From inspection of the mean fixation durations in Table 10.1 above, it can be seen that there is virtually no difference between temporal implicit and causal implicit for the hearing group (226 msecs compared with 228 msecs). Therefore the main interaction between inference type and explicit/implicit is being caused by the deaf subjects. This would be expected, given the results of the earlier studies which showed the degree of difficulty experienced by deaf subjects when reading implicit temporal texts.

Collapsing across subjects and treating texts as random variables, all significant factors and interactions above are replicated. The deaf are again significantly longer than controls, $F_{1,28}$= 169.277, $p<0.001$. There is also a significant interaction between inference type and explicit/implicit, $F_{1,28}$= 4.990, $p<0.05$. The three way interaction is also supported, $F_{1,28}$= 10.313, $p<0.005$. Post-hoc Scheffé tests for both interactions showed similar patterns to the by-subjects analysis.

10.3.1.2 Percentage regressions:

In calculating the mean percentage regressions, all regressions were included in the analysis. Rayner & Pollatsek (1989) discount the small corrective regressions which often follow return sweeps. However this procedure was not followed in the present study since, while subjects did show such corrective regressions, they did not do so consistently. They would often fixate on the middle of the new line or even almost to the extreme right after the return sweep, and follow this with several regressive sweeps towards the start of the new line.

The mean number of regressions for each group under each condition is show in Table 10.2 below:

TABLE 10.2: Percentage regressions

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	26.4	27.7	27.4	27.0
Hearing	25.3	25.9	25.0	26.1

There were no significant differences on any of the main factors and no significant interactions. However both the deaf readers and the hearing group make many more regressions (about 26% and 25% respectively) than the fluent adult readers quoted by Rayner & Pollatsek (1989), who made about 10% regressions. This most likely reflects the relative complexity for the children of the textual materials used in the present study.

10.3.1.3 Number of fixations per word:

The mean number of fixations per word gives an estimate of length of saccade during the reading process. The more fixations per word, the shorter the average length of saccade.

The mean number of fixations is shown below in Table 10.3:

TABLE 10.3: Number of fixations per word

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	1.282	1.505	1.257	1.287
Hearing	1.060	1.175	1.108	1.179

Collapsing across texts and treating subjects as random variables, an analysis of variance revealed a significant main effect with implicit passages having more fixations per word than explicit (1.236 compared with 1.137): $F_{1,12}$= 8.881, $p<0.01$. This was true for both subject groups. There was also a significant interaction between group and inference type: $F_{1,12}$= 6.860, $p<0.02$. Simple effects analysis shows that for the deaf group, there were significantly more fixations on temporal items than for causal (1.3934 compared with 1.2719).

Collapsing across subjects and treating materials as random variables, the ANOVA revealed a significant main effect with implicit texts having more fixations, $F_{1,28}$= 6.018, $p<0.02$, which supports the by-subjects finding. The interaction found above is also supported in the by-materials analysis, $F_{1,28}$= 4.827, $p<0.03$.

An additional significant main effect was found in the by-materials analysis with deaf subjects having more fixations per word than than hearing subjects, $F_{1,28}$= 28.623, $p<0.0001$. This effect, however, was not replicated in the by-subjects analysis. Therefore, this effect is likely to be the result of having few subjects (particularly in the deaf group), but many textual passages.

10.3.2 Content words versus function words:

A distinction is often made in psycholinguistics between two classes of words, known as "function words" and "content words". Function words include prepositions, conjunctions, articles, and pronouns. They usually mean little if read in isolation, eg "of", "the". Content words, which include nouns, verbs and adjectives, are, on the other hand, more meaningful in isolation, eg "tree", "dog", etc.

Adverbs present some difficulty as they could be accommodated in either class, depending on the context of their use. In the analysis outlined below, they were counted for the most part as content words. Only a handful were classified as function words, again because of the context of the textual material.

In addition, function words are usually regarded as a closed class, that is, there is a relatively small number of them in the language; content words, however, are an open class, with a limitless number of words available and new ones being added regularly to the language.

One view is that content words are the bricks or building blocks of language and function word are the cement that holds them together. As a result, function words tend to be among the most frequently used words in the language. Because of this high frequency, function words tend to be highly predictable and rapidly encoded. They are frequently skipped (Rayner & Pollatsek, 1989). Indeed, while about 90% of content words are fixated, only about 30% of function words are fixated by the eyes (Rayner, 1995). However, Wood (1984b) has argued that the deaf children are poorer than hearing children in their use of diectic words. Quigley et al (1976) have shown that the deaf have problems with various syntactic structures. The question is raised then as to whether the deaf fixate a larger percentage of function words than do the hearing controls.

The percentage of the total number of fixations which are on function words was calculated for all subjects on every text. These results were analysed using a one-way analysis of variance, collapsing across texts and using subjects as random variables, with subject group as the between group factor. This failed to show any significant difference between groups ($F_{1,12}= 0.428$, $p> 0.05$), with the deaf subjects having 18.3% and the controls 19.6% of their fixations on function words.

The analysis was repeated, collapsing across subjects and using texts as random variables. This also failed to reveal any significant differences between groups ($F_{1,30}= 2.295$, $p> 0.05$).

Summary of Results:

Fixations:

When reading, the deaf have longer mean fixation times.
Causal items had longer fixation times than temporal under implicit conditions by the deaf.

Regressions:

There were no differences in number of regressions between groups.
Both groups had considerably more regressions than adults.

Number of Fixations:

There were more fixations per word under implicit conditions than explicit.
For the deaf subjects only, there were more fixations for temporal than causal items.

Function vs content words

There were no differences between the deaf and hearing subjects in the types of word fixated, ie the deaf behave like the controls and tend to skip function words, fixating them only about 19% of the time.

10.3.3 Critical Regions of the text:

The above results are based on eye fixations over the complete text. The next sections looks at what may be happening in specific parts of the text. Some parts of the narrative may be considered more critical than others to the comprehension of the inference.

In the temporal texts in the third sentence, a change of scenario is signalled clearly in the explicit version (eg "After he went home,..."). In the implicit version, the change is signalled by mention of a lapse of time which is outside the normally accepted limits for the activity described in the first scenario (eg "Three hours later,..."). It is this change which is being probed in the question, whether the subject has realised that the scenario has in fact changed. As Anderson (1982) has shown, when a time-shift is encountered in a narrative, the sentence which contains the time-shift takes longer to read when the time-shift is beyond the normal time boundary expected for that scenario. The additional processing time is necessary as alternative scenario structures or other bridge-building operations are set up. Therefore, this phrase in both the explicit and implicit versions would seem to be the most critical in letting subjects know that a change has taken place.

In causal texts the situation is not so clearly delineated. The third sentence is still critical in conveying the information to be inferred but because of the nature of the inference, no one part of that sentence does this alone. The whole of the third sentence is necessary if the inference is to be understood.

Therefore, it was decided to examine these critical regions of the texts to see whether subjects did anything different from others parts of the material.

10.3.3.1 First pass fixations

The first pass fixation durations for the critical regions of the text as outlined above (ie the time shift phrase in the temporal items and the whole of the critical sentence in the causal items) were analysed using a 2X2X2 analysis of variance. The mean fixation times are given below in Table 10.4.

TABLE 10.4: First pass fixations (msecs)

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	274	266	331	274
Hearing	216	210	219	229

Collapsing across materials and using subjects as random variables, the deaf were found to have significantly longer fixations than the hearing subjects (286 msecs compared with 218 msecs), $F_{1,12}= 11.093$, $p< 0.01$. Causal texts also had longer fixation durations than temporal (263 msecs versus 241 msecs), $F_{1,12}= 4.720$, $p< 0.05$.

Collapsing across subjects and treating texts as random variables, the analysis revealed the deaf to have significantly longer fixations than hearing subjects ($F_{1,28}= 12.739$, $p<0.002$) but failed to failed to find any difference between causal and temporal items ($F_{1,28}= 1.108$, $p> 0.05$).

10.3.3.2 Second pass fixations

The mean times for second pass fixations is given in Table 10.5 below:

TABLE 10.5: Mean second pass fixation times (msecs)

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	160	203	170	163
Hearing	121	98	118	144

Collapsing across texts and treating subjects as random variables, The deaf are again found to have longer fixation times than the hearing group, but of borderline significance (174 msecs versus 120 msecs), $F_{1,12}= 4.490$, $p< 0.06$.

Collapsing across subjects and treating texts as random variables, the deaf have significantly longer fixation times than the hearing group ($F_{1,28}= 19.414$, $p< 0.0001$), thus supporting the by-subjects finding. However, in the by-materials analysis an interaction effect was found between temporal/causal items and group membership. Simple effects analysis indicated that temporal items had longer fixation times than causal. This effect was not replicated in the by-subjects analysis.

10.3.3.3 Total gaze duration

The mean times for total gaze duration are given in Table 10.6 below:

TABLE 10.6: Mean times for total gaze duration (msecs)

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	267	267	329	281
Hearing	222	211	223	231

Collapsing across texts and treating subjects as random variables, the analysis revealed the deaf to have longer fixation times than hearing subjects (286 msecs versus 221 msecs), $F_{1,12}= 8.426$, $p< 0.02$. Causal texts had longer fixation times than temporal (266 msecs versus 242 msecs), $F_{1,12}= 5.304$, $p< 0.04$.

Collapsing across subjects and treating texts as random variables, deaf were found to be longer than the hearing group ($F_{1,28}= 10.756$, $p< 0.003$). No significant difference was found between causal and temporal items ($F_{1,28}= 1.739$, $p> 0.05$).

10.3.3.4 Percent regression

The mean percentages regressions are shown in table 10.7 below:

TABLE 10.7: Percentage regressions

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	28.9	25.6	20.7	19.0
Hearing	25.1	24.7	22.6	21.4

Collapsing across texts and treating subjects as random variables, temporal items had significantly more regressions than causal (26 % compared with 21 %), $F_{1,12}$= 7.713, $p< 0.02$.

Collapsing across subjects and treating texts as random variables, temporal items were again found to have a greater percentage of fixations than causal items ($F_{1,28}$= 7.252, $p< 0.01$).

10.3.3.5 Fixations per word

The mean number of fixations per word is shown below in Table 10.8:

TABLE 10.8: Mean number of fixations per word

	TEMPORAL		CAUSAL	
	Explicit	Implicit	Explicit	Implicit
Deaf	1.374	1.665	1.167	1.085
Hearing	1.109	1.277	1.003	1.076

Collapsing across texts and treating subjects as random variables, temporal items were found to have more fixations than causal (1.356 versus 1.083), $F_{1,12}$= 26.253, $p< 0.0003$. An interaction effect was also found between group membership and temporal/causal type ($F_{1,12}$= 5.074, $p< 0.04$). Post-hoc simple effects analysis revealed temporal items have more fixations than causal especially for the deaf.

Collapsing across subjects and treating texts as random variables, temporal items were again found to have more fixations than causal ($F_{1,28}$= 28.316, $p<0.0001$) and the interaction between group membership and temporal/causal items was also significant

($F_{1,28}$= 7.974, p< 0.01). However, several other significant effects were found in the by-material analysis which were not supported in the by-subjects analysis: Implicit items had more fixations than explicit ($F_{1,28}$= 7.240, p< 0.01), deaf had more fixations than the hearing group ($F_{1,28}$= 17.446, p< 0.001), temporal-implicit had more fixations than temporal-causal ($F_{1,28}$= 6.694, p< 0.02), and there was a significant interaction between group, temporal/causal, explicit/implicit ($F_{1,28}$= 5.082, p< 0.03).

Most of the significant effects and interactions are supported in both the by-subjects and by-materials analyses. However, a few are significant only in one or other. According to Clark's (1973) reasoning we require significance in both analyses to be able to confidently accept the effects as real. In the present study, the sample size is, of necessity, small. As a result variability within each subject group becomes more of a problem. With a bigger sample one would expect such individual variability to be evened out. Therefore, in the summary below only those effects which are significant in both the by-subjects and by-materials analyses are quoted.

Summary of results:

First pass fixations:

The deaf have longer fixation times in the critical regions than hearing subjects.

Second pass fixations:

The deaf again have significantly longer fixations than the controls.

Total gaze duration:

The deaf once more have longer fixation durations than the controls.

Percent regressions:

Temporal items have more regressions within the critical regions of the texts than causal items.

Fixations per word:

Temporal items have more fixations per word, ie have shorter saccade lengths, than than causal items, this effect operating more strongly for the deaf.

10.4 Discussion:

The studies reported in earlier chapters used either cued memory probes or on-line methods. Each has its disadvantages, which have already been discussed. The present study is an attempt to monitor the eye movements of subjects as they read text which is presented in a "natural" form, ie the text is presented in total, as it would appear on the page. The reader is free to move forwards or backwards, to scan over the text at will. The technique is unobtrusive yet can monitor what is happening from moment to moment during the reading process.

The most obvious finding of this study is that the deaf subjects have significantly longer fixation times than hearing subjects. This is true for all fixations, for the mean overall fixation times for entire texts as well as for first pass, second pass, and total gaze duration for the critical regions of texts. One possible reason for this could be that the deaf have slower lexical access.

There is no difference between deaf and hearing subjects in the percentage of regressions found in the overall text. However, both groups make significantly more regressions than fluent adult readers (about 25% compared with 10%). This is more in line with the findings of Taylor (1965) who found the average number of regressions for hearing 14 year olds to be about 20%.

When the critical regions are examined, more regressions were found for temporal items than for causal. This may reflect the fact that the time shift phrase for temporal items is relatively self contained and is pivotal to the inference. With causal texts, although the whole of the critical sentence contains the information necessary for the inference, it is part of a longer causal chain of events, a series of causal antecedents and consequences, which may possess more forward momentum than is is possible with a temporal text. Therefore regressions are less likely.

The number of fixations per word is a rough indication of mean saccade length; the more fixations per word, the shorter the saccade. For complete texts, implicit items had shorter saccade lengths than explicit items. This would seem logical in that, if implicit text is more difficult to comprehend, then the reader should make a more concentrated effort, in this case by jumping a shorter distance each time the eyes are moved. In the present study no difference is found between mean fixation times for explicit or implicit texts. Rayner and Pollatsek (1989) have shown that as the difficulty of text increases, readers make longer fixations, smaller saccades, and more regressions, resulting in slower overall reading rates. In the present study, subjects do seem to be making shorter saccadic jumps but are not making longer fixations, and do not make more regressions. Since there is no difference between explicit and implicit in mean fixation durations, the fact that there are more fixations per word for implicit texts means that the implicit texts must have slower overall reading rates.

It should be remembered that in the on-line experiment no difference was found in reading times between explicit and implicit texts. However these results were for those measures taken, ie for the critical sentence, the filler sentence, and the question-response, and not for the entire text. Also the experimental procedure was different in that only one sentence was presented at a time. Each sentence disappeared as the next one came on screen. The presentation rate was under the control of the subject. This is less natural than the present study where the entire text is presented on screen together and subjects can jump backwards or forwards as necessary.

It is possible, therefore, that in the on-line study, where the text is presented one sentence at a time and erased as each new sentence replaces it, subjects have to be more cautious. Without the possibility of looking back to check any uncertainties, subjects have to be sure that they completely understand what has been read before committing themselves to the next sentence. Given that each presentation contained both explicit and implicit items, subjects may treat both conditions with equal caution, taking extra time to read all texts, and this resulting in no difference between the two conditions in overall reading times.

With the more normal situation of the eye tracking study, the text is presented in its entirety, subjects have more freedom and flexibility. They can afford to read what they believe to be an explicit text more quickly since they can always backtrack if there are any difficulties, for example, if the text turns out to contain implicit information.

Support for the above argument can be found when the overall reading times for the two experimental paradigms are compared. If the mean reading times for the critical sentence in the on-line study and for the complete text in the eye movement study are prorated for a ten word sentence, direct comparison of overall reading times is possible. It is found that the hearing subjects do indeed take longer to read the text, for both inference types under explicit and implicit conditions, in the on-line study than in the eye movement study. Therefore, under the conditions of presentation used in the on-line study (ie, one sentence at a time with no possibility of backward checks), the hearing subjects are more cautious, with longer overall reading times. It should be noted that, for the deaf children, mean overall reading times for the on-line study are not longer than overall reading times for the eyetracking study. The possible reasons for this are discussed later.

Under "normal" conditions, ie when reading simple, explicit material, it is known that adult subjects can and do make use of information from parafoveal input. Whole words to the right of the fixated word can be identified, or more likely, partial information about the next word can be garnered (Rayner & Pollatsek, 1989). It has been shown by a number of studies, eg Rayner, Well, Pollatsek, & Bertera (1982), Rayner & Bertera (1979), Rayner (1975a), Underwood & McConkie (1985), Balota, Pollatsek & Rayner (1985), McConkie & Hogaboam (1985), and Pollatsek, Rayner & Balota (1986), that subjects can extract useful information up to about 15 character spaces to the right of the fixated word (in English). The next word or two to the right of the fixated word can often be identified. It is more likely that some information can be gained about the next word to the right but there is also likely to be an element of guessing from the context. Readers can use this information to help them determine where to fixate next.

Rayner (1986) used a moving window technique to investigate the perceptual span of young readers. This method uses an eyetracking technique in which a small region of the text around the fixation point is presented as normal but the surrounding text is made unreadable (by, for example, replacing every letter by the letter "x"). When the reader's eye moves on to a new fixation, the text around this area appears as normal text and the new surrounding text is obliterated. This happens so rapidly that it is not noticed by the reader. As the window size is reduced to less than the subject's perceptual span, reading should be disrupted.

Using this technique, Rayner found that beginning readers (aged about 7 years) have a slightly smaller perceptual span than skilled readers, about 11 character spaces compared with 15. However, after about a year these very young readers are directing much of their attention to the right of the fixation, as do skilled readers. As the window is reduced to about 5 character spaces, the reading rate of younger readers is disrupted less than that of older children who in turn are disrupted less than adults. This would suggest that the smaller perceptual span of the youngest readers is due to their devoting more attention during a fixation to foveal word processing than do more skilled readers. Rayner also found that if the difficulty of the text is increased, the older children (aged about 9 years) reduce their perceptual span to that of the younger children. These findings were for normally hearing children. There would appear to be no such studies involving deaf children.

Possibly what is happening in the present study is that subjects are being more cautious with implicit texts and are not relying as much on this parafoveal input. They are fixating for a similar amount of time overall but are concentrating much more on foveal input. They are, in effect, behaving as younger readers and reducing their saccadic jumps as the text becomes more difficult, or at least less certain.

The cautious approach, which the hearing subjects seem to be adopting when using the on-line protocol, does not appear to be used by the deaf. They have a longer overall reading time (on average) for the eye movement study than for the on-line study. The deaf, it seems, are using a different process.

For causal texts, the deaf were found to have longer fixation times than for temporal items. This makes sense in light of the findings from the on-line study where causal items were found to be processed more carefully (ie at a slower rate but more accurately) than temporal items.

However, for the deaf, temporal texts have shorter saccade lengths than causal texts, both over the entire narrative and within the critical regions. Again, if temporal items are more difficult (especially for the deaf), then one might expect that readers would make shorter saccades (Rayner & Pollatsek, 1989).

It should be noted that temporal items have shorter saccade lengths for the control group also. This confirms the findings of the on-line study that temporal items proved to be more difficult than causal for both groups of subjects, but particularly so for the deaf.

When the types of word which are fixated are examined, no difference is found between deaf and hearing subjects. For both groups, 81% of fixations made are on content words and only 19% on function words. Again this is what would be expected from previous research (Rayner, 1995).

The findings of the present study expand on those of the previous experiments. From the on-line study it was concluded that, with sufficient context, the deaf make elaborative inferences as the implicit material is read, in a similar manner to hearing children. Only when they seemed to lack the relevant knowledge or experience (direct or indirect) of the situation described in the text did they fail to process on-line and revert to bridging strategies. With a less enriched context (as in the verb study) instrumental elaborations were not made, confirming the general consensus view that such inferences are not generally made on-line (eg Graesser & Kreuz, 1993). From the signing study, it seemed to be the case that these mechanisms were independent of modality of communication, reflecting a more generalised language difficulty which extended beyond the process of reading.

The on-line study relies on proving the null-hypothesis, with the attendant difficulties of this methodology, which have already been discussed. By studying eye movements during the reading process, it has been possible to examine more precisely what is happening from moment to moment. It has been discovered that both deaf and hearing readers process implicit material more slowly, by making shorter saccadic sweeps. This, however, is still reconcilable with the view that inferences are being made on-line. It has been stated earlier that it is generally agreed that first pass fixations are good estimates of speed of lexical access while regressive movements reflect difficulties with higher order processing (Rayner & Pollatsek, 1989). The slower approach used by both groups for implicit material, ie shorter saccadic sweeps, is not regressive. The movement continues to be forward through the text. However, the slower processing speed might reflect the fact that it takes time to make an elaborative inference.

Magliano et al (1993) showed that causal antecedent inferences were generated on-line but only when readers were given sufficient time to do so. They estimated that such inferences may be generated after a delay of about 700 milliseconds but could not say for certain when exactly the inference is made. These authors did not find any evidence to support the generation of causal consequence inferences on-line but they themselves admit that their textual materials may not have been constraining enough.

The technique used by Magliano and his co-workers involved a predetermined rate of presentation of text and a lexical decision task (unlike the studies in this thesis).

Other estimates of the time taken to generate particular classes of inference automatically on-line are around this magnitude, eg 650 milliseconds (Kintsch, 1988; McKoon & Ratcliff, 1992).

Fincher-Kiefer (1993) proposes that at the situation model level, inferences are made, although these might be incomplete or tentative, but they lag slightly behind the reader's textbase construction.

The overall picture is that deaf children are not so very different from their hearing counterparts when they read. They tend to have a slower overall reading rate but this may reflect problems with lexical access. However, when attempting to comprehend implicit material, the deaf use their general world knowledge in a similar manner to hearing children, but because of lack of appropriate experience, direct or indirect, this background knowledge may not be as enriched as it is for hearing children.

The particular difficulties experienced by both groups (for example, the greater difficulty in making temporal inferences) seem to be greater for deaf children. It was stated at the beginning of this chapter that this eyetracking study was of the nature of a pilot study. Because of the difficulties in conducting this particular experiment, and the relatively small numbers of deaf subjects used, the results should be viewed as promising but not definitive.

Chapter Eleven:

Overview

Deaf children's reading has been extensively studied. However, like research into the reading of hearing children (eg Harrison, 1996), much of this work has centred on the early stages of learning. Less attention has been directed towards comprehension skills, those aspects of the reading process which go beyond the initial stages of decoding and word identification. These higher order reading skills are none-the-less very important. They are crucial to the skilled reader.

While the purpose of the present thesis was to examine those aspects of reading concerned with comprehension, particularly inferencing skills, account was taken of those factors of lower order processing as they apply to deaf children. The fact that the deaf have more limited vocabularies than hearing children (eg Hatcher & Robbins, 1978; Paul, 1984), have problems with English syntax (eg Quigley et al, 1977), and have difficulties with some aspects of figurative language (eg Conely, 1976; Orlando & Shulman, 1989), all influenced the design of the textual materials used in the present series of studies.

It has been shown that deaf children do bring their own world knowledge (or schema) to the reading process (eg Gaines et al, 1981), but that some perhaps use these broad, top-down skills to compensate for the poorer bottom-up skills described earlier. While the deaf tend to develop similar discourse structures and story grammars as hearing children, it has been shown that they are less skilled at dis-embedding the most important information from a story structure (Coggiola, 1983).

Very little work has been directed towards inferencing skills in the deaf, yet it is at the very level of complexity where text relies more on implicit content (ie about the eight year old level) that deaf children reach a plateau, and make very little progress.

The main study in this area was by Wilson (1979). He found that deaf children were as competent in literal interpretation as hearing children but much poorer when an inference had to be drawn. He also found this to be true when the the inference had to be made from a signed presentation. Wilson concluded that the deaf were non-constructive readers, ie that they were not using a schema driven approach to text interpretation.

In contrast to the paucity of work directed towards inferencing skills in the deaf, there is a very large volume of studies examining this process in hearing adults. Much of this work has has occurred over the last twenty or so years and concerns itself with what type of inferences are made (forward or backward) and when such inferences take place (at the time of reading or a some later point). These many studies have used a wide variety of experimental techniques, each with its own advantages and disadvantages. The type of inference, the length of text (eg single sentences, paragraphs, longer segments), how the text is presented, and the measures used to determine whether an inference has been made have varied across studies. This has led to differing results, with obvious difficulties in obtaining a unified interpretation of what is occurring. The current consensus seems to be that only those inferences which are necessary for local cohesion of the text are made automatically as the text is read. Only when the context is so constraining as to eliminate alternative interpretations will an elaborative inference be made on-line. However, exactly what inferences are drawn depends also on the genre of the text, the reader's motivations, and the intentions of the author.

The present series of studies examined the inferencing skills of severely/profoundly deaf adolescents of average intellectual ability, using a variety of experimental techniques, ranging from memory probes, through on-line measures of reading time, to measurement of eye movements as subjects read. Three types of inference were examined: spatial temporal, and causal.

The first study showed that the deaf have more difficulty interpreting implicit material than the hearing subjects, even those younger subjects matched on reading ability. It was found that the deaf adolescents were as good at making spatial inferences as hearing children approximately matched on reading ability. However, temporal inferences proved to be the most difficult for both hearing and deaf subjects, but particularly so for the deaf.

The on-line study confirmed that the deaf are poorer at drawing temporal and causal inferences but when they did so, they appeared to be making elaborative inferences automatically on-line just like the hearing subjects. Temporal narratives seem to be particularly problematic for the deaf. It seemed that these inferencing skills were independent of modality of communication and operated in a similar way when the information was presented in signed form, indicating that the difficulties were not purely reading based. It was postulated that perhaps the deaf lacked the necessary experience and background knowledge of some situations and so were unable to make inferences on-line in these cases. When the deaf readers do have the relevant background knowledge, they perform much like hearing children.

A further reading time study, which examined instrumental inferences (a particular type of elaboration in which a verb can imply a particular, associated instrument, eg stir-spoon), showed that deaf children are like hearing children in that neither group draws these inferences automatically on-line. This finding supports the consensus view that fluent adult readers do not draw such inferences at the time of reading unless the context is sufficiently constraining (eg Rayner & Pollatsek, 1989)

There was a remarkable similarity between the deaf and hearing children's constructions of the various scenarios described in the textual materials used in the experiments. (Only some of the deaf children took part in the on-line studies). There seemed to be similarities in the relative richness of the various scenarios for both groups. However, the deaf did show a more variable range of time estimates for the temporal scenarios, while the hearing subjects tended to give richer descriptions for the causal scenarios. However, no correlation was found between those scenarios with less enriched background, and those which the deaf found more difficult to interpret.

The eye movement study allowed a more precise examination of what was happening during the reading process. It seemed that the deaf and hearing subjects were actually taking slightly longer to read implicit material but that this was caused by closer scrutiny of the text by subjects. It had not been possible to detect this with the on-line study. However, it is still likely that both the hearing and deaf readers are drawing inferences on-line and the extra time is that needed to make the necessary elaborations. It should be noted that the type of inference demanded by the texts used in these studies should be fairly automatic but will still rely on subjects using their general world knowledge. It was also found that deaf subjects probably take longer to access meaning from words.

Pedagogical implications

These findings are very encouraging. They show the normality of deaf readers, that they use similar mechanisms to hearing children when comprehending text. Yet Wood (1980) contends that the teaching experiences of deaf and hearing children are very different. Conversations with deaf children tend to concentrate on the literal and pragmatic functions of language and ignore the more metaphorical and imaginative aspects. There is a strong emphasis towards "bread-and-butter" language.

Wood argues that deaf children are also much more dependent on the teacher. If they encounter a difficult word they look to the teacher, who invariably gives the deaf child the full answer straight away, as it is often assumed that the child does not know the meaning of the word as well as being unable to read it. Explanations tend to be concrete in nature and are very disruptive to the flow of the reading process. This is in contrast to teachers of hearing children, who are assumed to know the meaning of the word and are, therefore, given prompts to help them decode more accurately.

An empirical study by Howarth et al (1981) confirms that deaf children are indeed stopped more frequently by their teachers than are hearing children, and that these extra stoppages occurred because the deaf child appeared not to know the meaning of a word. Because of this, the reading rates of the deaf children were much slower than those of the hearing children. Indeed, they were so slow as to be below the normally accepted rate at which spoken language can be understood.

A study by Allington (1980) of hearing children found that poor readers are interrupted more often by their teachers than are good readers.

Durkin (1978-79) found that teachers of hearing children expend much more effort in assessing comprehension, with less than 1% of reading lessons being devoted to comprehension instruction. What children experience is a massive dose of unguided practice and little help in how to actively acquire comprehension skills.

One feature of reading lessons for the deaf is the fact that the child is most often expected to read aloud. In this way the teacher can monitor the accuracy of the child's performance. Since the child's speech is likely to be unclear, the teacher cannot be sure how accurately the child is performing. This in turn leads to more breaks to check certain words and again causes disruption.

Perhaps another way of monitoring performance would encourage greater fluency. The child could read silently (the way most skilled readers do) and questions could answered afterwards to check the level of comprehension.

A few studies have examined the efficacy of directly teaching children comprehension skills. Hansen (1981) showed that seven year old hearing children, with age appropriate reading skills, who are trained to integrate material from the text with their prior knowledge, and who are encouraged to predict what might happen next in the story, show better comprehension than children who are not given such training. However, their improved performance related only to the texts used in the training and did not transfer to new material.

In a refinement of the above study, Hansen & Pearson (1983) compared such training techniques on hearing nine year olds who were good or poor comprehenders. They found that the training helped the poorer readers, not only to understand the original passages, but to transfer this skill to new texts. There was no effect for the more competent readers. The findings of both these studies suggest that encouraging subjects to make inference can be effective for young children or for older children who have poor levels of comprehension. Older and more competent readers seem to make inferences spontaneously. More recent studies by Yuill & Joscelyne (1988) and by Yuill & Oakhill (1988) would seem to confirm these findings. However, Oakhill & Garham (1988) emphasise that all such training can only be effective if the children have the requisite background knowledge to start with, or if it is provided for them prior to reading.

A similar study by Maeder (1979) with deaf children reading at the eight year level or above, found that written or signed questions inserted during reading to enhance comprehension had little or no effect.

It seems, therefore, that deaf children are treated very much like hearing children who are poor comprehenders. It has been shown that direct instruction in comprehension techniques can be effective in such circumstances, yet the present series of studies show that, with the appropriate background knowledge, older deaf children can draw inferences as effectively as their hearing counterparts. What they seem to lack is a breadth of background knowledge and experience. Thus, while the process of making inferences is similar for both groups, the deaf read more slowly with lower level of comprehension. This obviously has a wide ranging impact on all areas of the curriculum, where the pupil is reading to learn.

Deaf children perhaps need to be encouraged to expand their experiences of life. Instead of concentrating too much on developing the deaf child's proficiency in, say, English syntactic construction, it may be more profitable to encourage broader discussion around whatever is being read, to empower the deaf child to explore beyond that solely presented on the page. This is in line with Wood (1980) who advocates encouraging "abstract, imaginative, poetic uses of language" during deaf children's reading lessons. This sentiment is echoed by Webster (1986, p249) when he suggests that "*...accepting the linguistic limitations of deaf children inspires an approach whereby more efficient, more active questioning of text is encouraged.*"

Since the present set of studies shows that the deaf use similar mechanisms to hearing children when drawing inferences, more exploration of various text genre could be incorporated into the curriculum of hearing impaired children. As Zwaan and van Oostendorp (1993) have shown that the text genre can affect the types of inference made (more spatial inferences in detective stories, for example), investigation of different story types should be encouraged, perhaps with explicit instruction on how to extract information that goes beyond the literal.

If we now know that deaf readers are capable of a greater level of comprehension than was previously suspected, there should possibly be a shift away from the stifling insistence on correct grammatical interpretation and precise decoding. By encouraging greater fluency, deaf children will be encourage the read more for pleasure. The situation described by Strassman (1992) where deaf children see reading merely as one of the things which one does at school, or Cornett's (1975) statement that as few as 5% of young, intelligent deaf adults read for pleasure, could be turned around and deaf children be given the opportunity to fully enjoy books and literature which until now has been denied to so many of them.

Suggestions for further work

The attempt to compare different types of inference, such as temporal and causal, by using reading rates rather than direct comparison of reading times, led to difficulties in the analysis of the results, since the relationship between reading rate and sentence length is not linear (see Section 6.3.2.2). This could be accomodated by reanalysis of the data in the on-line reading study, but no such adjustments could be made with the data from the signing study. In any further work in this area, it would seem sensible to design the textual materials much more tightly so that direct comparison can be made across different inference types, using reading times as the metric.

The series of studies described in this thesis should be viewed merely as a starting point. Further work is needed to explore more thoroughly the various types of inference encountered in everyday textual materials, for example, causal antecedents and causal consequences. The work already begun with hearing adults could profitably be extended to deaf adolescents. Further studies using a range of techniques described earlier, but particularly those involving the monitoring of eye movements, are called for. Given the variability within the deaf community, further studies using eyetracking techniques with a much larger sample size would seem to be necessary.

Any findings from such further work should lead to a more appropriate approach to enhancing deaf children's comprehension skills.

References:

Aarons, L. (1971). Subvocalisation: aural and EMG feedback in reading. *Perceptual and Motor Skills, 33*, 271-306.

Aaronson, D. & Scarborough, H. S. (1976). Performance theories for sentence coding: Some quantitative evidence. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 56-70.

Abbott, V. & Black, J. B. (1986). Goal-related inferences in comprehension. In J. A. Galambos, R. P. Abelson, & J. B. Black (Eds.), *Knowledge structures*. Hillsdale, NJ: Erlbaum.

Alba, J. W. & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93*, 203-231.

Allington, R. (1980). Teacher interruption behaviours during primary-grade oral reading. *Journal of Educational Psychology, 72*, 371-374.

Anderson, A. (1982). Text comprehension - the influence of temporal information on processing and reading rate. Unpublished Ph.D. thesis, University of Glasgow.

Anderson, A., Garrod, S., & Sanford, A. (1983). The accessibility of pronominal antecedents as a function of episode shift in narrative text. *Quarterly Journal of Experimental Psychology, 35*, 427-440.

Anderson, R. (1981). A proposal to continue a centre for the study of reading (Technical Proposal, 4 Vols.). Urbana, IL: University of Illinois, Centre for the Study of Reading.

Anderson, R. C & Ortony, A. (1975). On putting apples into bottles -A problem of polysemy. *Cognitive Psychology, 7,* 167-180.

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schollert, D. L., Stevens, K. V., & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behaviour, 15,* 667-679.

Arlin, P. (1977). The function of Piagetian operational levels in the preference and production of metaphors. Paper presented to the Society for Research in Child Development, New Orleans.

Ayers, T. L. (1984). Silent reading time for tongue-twister paragraphs. *American Journal of Psychology, 97,* 605-609.

Baddeley, A. (1976). *The psychology of memory.* London: Harper & Row.

Baddeley, A. (1979). Working memory and reading. In P. Kolers & M. Wrolstad (Eds), *Processing of visible language,* Vol 1. New York: Plenum Press.

Baddeley, A & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation,* Vol. 8. New York: Academic Press.

Baddeley, A. & Lewis, V. (1981). Inner active processes in reading: The inner voice, the inner ear, and the inner eye. In A, M. Lesgold & C. A. Perfetti (Eds), *Interactive processes in reading.* Hillsdale, NJ: Erlbaum.

Baggett, P. (1975). Memory for explicit and implicit information in picture stories. *Journal of Verbal Learning and Verbal Behaviour, 14,* 538-548.

Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d' Arcais & K. Rayner (Eds.) *Comprehension processes in reading.* Hillsdale, N. J.: Lawrence Erlbaum Associates.

Balota, D., Boland, J., & Shields, L. W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language, 28*, 14-36.

Balota, D. & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 340-357.

Balota, D., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*, 364-390.

Banks, J., Fraser, R., Fyfe, J., Grant, C., Gray, M., MacAuley, M., & Williams, S. (1989). Teaching deaf children to read: A pilot study of a method. *Journal of the British Association of Teachers of the Deaf, 13*, 129-141.

Banks, W. P., Oka, E., & Shugarman, S. (1981). Recoding of printed words to internalised speech: Does recoding come before lexical access? In O. J. L. Tzeng & H. Singer (Eds), *Perception of print.* Hillsdale, NJ: Erlbaum.

Barnitz, J. (1980). Syntactic effects on the reading comprehension of pronoun-referent structures by children in grades two, four, and six. *Reading Research Quarterly, 15*, 268-289.

Baron, J. (1973). Phonetic stage not necessary for reading. *Quarterly Journal of Experimental Psychology, 25*, 241-246.

Bartlett, F. (1932). *Remembering.* Cambridge: Cambridge University Press.

Beggs, W. & Breslaw, P. (1982). Reading retardation or linguistic deficit? III: a further examination of response strategies in a reading test completed by hearing-impaired children. *Journal of Research in Reading, 6*, 19-28.

Beggs, W., Breslaw, P., & Wilkinson, H. (1982). Eye movements, reading achievement and reading test validity. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements.* Amsterdam: North Holland Publishing Co.

Bell, B. J. & Torrance, N. (1986). Learning to make and recognise inferences in the early grades. Paper presented to the 67th Annual Meeting of the American Educational Research Association (San Francisco, California: April, 1986).

Bellugi, U., Klima, E., & Siple, P. (1974). Remembering in signs. *Cognition, 3*, 93-125.

Bench, J. & Bamford, J. (1979). *Speech-hearing tests and the spoken language of hearing-impaired children.* London/New York/San Francisco: Academic Press.

Billow, R. (1975). A cognitive developmental study of metaphor comprehension. *Developmental Psychology, 11*, 415-423.

Black, J. B. & Bower, G. H. (1980). Story understanding as problem solving. *Poetics, 9*, 223-250.

Bloom, C. P., Fletcher, C. R., van den Broek, P., Reitz, L., & Shapiro, B. P. (1990). An on-line assessment of causal reasoning during comprehension. *Memory and Cognition, 18*, 65-71.

Bormuth, J. (1970). *On the theory of achievement test items.* Chicago: University of Chicago Press.

Bransford, J & Johnson, M. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behaviour, 11*, 717-926.

Brennan, M., Colville, M. D., & Lawson, L. K. (1980). *Words in hand: A structural analysis of the signs of British Sign Language*. Edinburgh: Moray House College.

Brewer, W. (1975). Memory for ideas: Synonym substitution. *Memory and Cognition, 3*, 458-464.

Brewer, W. (1980). Literary theory, rhetoric, stylistics: Implications for psychology. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Brimer, A. (1972). *Wide-span reading test*. London: Nelson.

Brooks, P H. (1978). Some speculations concerning deafness and learning to read, in Liben, L S (Ed), *Deaf Children: Developmental Perspectives*, (pp 87-101). New York, Academic Press.

Brown, A. (1975). Recognition, reconstruction and recall of narrative sequences by pre-operational children. *Child Development, 46*, 156-166.

Brown, A. (1980). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, N. J.: Erlbaum.

Brown, A, Campione, J. & Day, J. (1981). Learning to learn: On training students to learn from texts. *Educational Research, 10*, 14-21.

Brown, A. & Day, J. (1980). Strategies and knowledge for summarizing texts: The development of expertise. Unpublished manuscript, University of Illinois, Centre for the Study of Reading, Urbana, Illinois.

Brown, A & Murphy, M. (1975). Reconstruction of arbitary versus logical sequences by pre-school children. *Journal of Experimental Child Psychology, 20,* 307-326.

Bryant, P & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature, 232,* 456-458.

Campbell, F. W. & Wurtz, R. H. (1978). Saccadic omission: Why we do not see a grey-out during a saccadic eye movement. *Vision Research, 18,* 1297-1303.

Carpenter, P. A. & Just, M. A. (1977). Reading comprehension as eyes see it. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension.* Hillsdale, NJ: Erlbaum.

Chafe, W. (1970). *Meaning and the structure of language.* Chicago: University of Chicago Press.

Chafe, W. (1972). Discourse structure and human knowledge. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* Washington: Winston.

Chai, D. (1967). Communication of pronoun reference in ambiguous English sentences for children and adults. University of Michigan, Centre for Human Growth. (ERIC Document Service, ED 012 889).

Chen, K. (1976). Acoustic images in visual detection for deaf and hearing college students. *Journal of General Psychology, 94,* 243-246.

Chomsky, N. (1957). *Syntactic Structures.* The Hague, Netherlands: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N (1976). *Reflections on language.* London: Temple-Smith.

Chumbley, J. I. & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory and Cognition, 2,* 130-138.

Clark, E. V. (1978). From gesture to word: On the natural history of deixis in language. In J. S. Bruner & A. Garton (Eds.), *Human growth and development.* Oxford: Oxford University Press.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behaviour, 12,* 335-359.

Clark, H. H. (1975). Bridging. In R. Schank & B. Nash-Webber (Eds.), Theoretical issues in natural language processing. Proceedings of a conference at the Massachusetts Institute of Technology.

Clay, M. (1977). *Reading: the patterning of complex behaviour.* London: Heinemann.

Coggiola, D. C. (1983). The identification and use of levels of importance in text learning by hearing impaired college students. Unpublished doctoral dissertation, University of Rochester, New York.

Collett, D. & Lewis, T. (1976). The subjective nature of outlier rejection procedures. *Applied Statistics, 25,* 228-237.

Cometa, M. & Eson, M. (1978). Logical operations and metaphor interpretations: A Piagetian model. *Child Development, 49,* 649-659.

Conely, J. (1976). Role of idiomatic expressions in the reading of deaf children. *American Annals of the Deaf, 121*, 381-385.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory and Cognition, 2*, 130-138.

Conrad, R. (1977). The reading ability of deaf school-leavers. *British Journal of Educational Psychology, 47*, 138-148.

Conrad, R. (1979). *The deaf schoolchild.* London: Harper & Row.

Corbett, A. & Dosher, B. (1978). Instrument inferences in sentence encoding. *Journal of Verbal Learning and Verbal Behaviour, 17*, 479-491.

Corcoran, D. W. J. (1966). An acoustic factor in letter cancellation. *Nature, 210*, 658.

Cornett, O. (1975). Cued speech: What and why? In P. Henderson (Ed.), *Methods of communication currently used in the education of deaf children*, pp 41-44. London: Royal National Institute for the Deaf.

Cotter, C. A. (1984). Inferring indirect objects in sentences: Some implications for the semantics of verbs. *Language and Speech, 27*, 25-45.

Davey, B, LaSasso, C, & Macready, G. (1983). Comparison of reading comprehension task performance for deaf and hearing readers. *Journal of Speech and Hearing Research, 26*, 622-628.

Day, J. (1980). Training summarization skills: A comparison of teaching methods. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

DiFrancesca, S. (1972). Academic acheivement test results of a national testing program for hearing-impaired students-United States, Spring (Series D, No 9). Washington, DC: Gallaudet College, Office of Demographic Studies.

Dixon, K, Pearson, P, & Ortony, A. (1980). Some reflections on the use of figurative language in children's textbooks. Paper presented at the annual meeting of the National Reading Conference, San Diago.

Doctor, E. A. & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory and Cognition*, *8*, 195-209.

Dodd, B. (1980). The spelling abilities of profoundly pre-lingually deaf children, in U. Frith (Ed), *Cognitive processes in spelling*, pp 423-440. London: Academic Press.

Dodd, B. (1987). Lip-reading, phonological coding and deafness. In B. Dodd & R. Campbell (Eds). *Hearing by eye: The psychology of lip reading*. London: Erlbaum.

Dodd, B & Hermelin, B. (1977). Phonological coding by the pre-lingually deaf. *Perception and Psychophysics*, *21*, 413-417.

Dopkins, S. & Myers, J. L. (1989). Goal relationships during the processing of text. Unpublished manuscript, University of Massachusetts, Amherst.

Dosher, B. A. & Corbett, A. T. (1982). Instrument inferences and verb schemata. *Memory and Cognition*, *10*, 531-539.

Duffy, S. A. (1986). Role of expectations in sentence integration. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *12*, 208-219.

Durkin, D. (1978-79). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly*, *14*, 481-533.

Dyer, M. G. (1983). *In-depth understanding: A computer model of integrated processing for narrative comprehension*. Cambridge, MA: MIT Press.

Edfeldt, A. W. (1960). *Silent speech and silent reading*. Chicago: University of Chicago Press

Ehrlich, K. & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behaviour, 21*, 296-306.

Engle, R. W., Cantor, J., & Turner, M. (1989). Modality effects: Do they fall on deaf ears? *Quarterly Journal of Experimental Psychology, 41*, 273-292.

Erickson, M. E. (1987). Deaf readers reading beyond the literal. *American Annals of the Deaf, 132*, 291-294.

Ewoldt, C. (1981). A psycholinguistic description of selected deaf children reading in sign language. *Reading Research Quarterly, 17*, 58-89.

Fillmore, C. (1968). The case for case. In E Bach & R Harms (Eds), *Universals in linguistic theory* (pp 1-88). New York: Holt, Rinehart and Winston.

Fincher-Kiefer, R. (1993). The role of predictive inferences in situation model construction. *Discourse Processes, 16*, 99- 124.

Fischer, S. & Gough, B. (1978). Verbs in American Sign Language. *Sign Language Studies, 18*, 17-48.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*, 221-233.

Fletcher, C. R. (1986). Strategies for the allocation of short-term memory during comprehension. *Journal of Memory and Language, 25*, 43-58.

Fletcher, C. R. & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language, 27*, 235-244.

Forster, K. I. (1981). Priming and the effects of sentence and lexical context on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33*, 465-495.

Fraise, P. (1963). *The psychology of time*. New York: Harper & Row.

Frith, U. (1979). Reading by eye and writing by ear. In P. A. Kolers, M. E. Wrolstad & H. Bouma (Eds). *Processing of visible language*, Vol. 1. New York: Plenum.

Fry, E. (1977). Fry's readability graph: clarification, validity and extension to level 17. *Journal of Reading, 21*, 242-252.

Furth, H. (1966). A comparison of reading test norms of deaf and hearing children. *American Annals of the Deaf, 111*, 461- 462.

Fusaro, J & Slike, S. (1979). The effect of imagery on the ability of hearing-impaired children to identify words. *American Annals of the Deaf, 124*, 829-832.

Gaines, R, Mandler, J, & Bryant, P. (1981). Immediate and delayed story recall by hearing and deaf children. *Journal of Speech and Hearing Research, 24*, 463-469.

Garnham, A. (1979). Instantiation of verbs. *Quarterly Journal of Experimental Psychology, 31*, 207-214.

Garrod, S., O'Brien, E., Morris, R., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 250-257.

Garrod, S. & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behaviour, 16*, 77-90.

Garrod, S. & Sanford, A. (1978). Anaphora: A problem in text comprehension. In R. N. Campbell & P. T. Smith (Eds.), *Recent advances in the psychology of language.* New York: Plenum Press.

Garrod, S. & Sanford, A. (1980). Bridging inferences and the extended domain of reference. Paper presented at 9th International Attention and Performance Conference, Cambridge, England.

Garrod, S. & Sanford, A. (1981). Bridging inferences in the extended domain of reference. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp 331-346). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gentner, D. (1981). Verb semantic structures in memory for sentences: Evidence for componential representation. *Cognitive Psychology, 13*, 56-83.

Giorcelli, L. (1982). The comprehension of some asspects of figurative language by deaf and hearing subjects. Unpublished doctoral dissertation, University of illinois at Urbana-Champaign.

Goetz, E. T. (1977). Inferences in the comprehension of memory for text. Urbana, ILL: University of Illinois, Centre for the Study of Reading (ERIC Document and Reproduction Service No. ED 150 548).

Goggiola, D. (1983). The identification and use of levels of importance in text learning by hearing impaired college students. Unpublished doctoral disertation, University of Rochester, New York.

Golding, J. M., Graesser, A. C., & Millis, K. K. (1990). What makes a good answer to a question? Testing a psychological model of question answering in the context of narrative text. *Discourse Processes, 13*, 305-326.

Goodman, K. (1969). Analysis of oral reading miscues: applied psycholinguistics. *Reading Research Quarterly, 5*, 9-30.

Gormley, K & Franzen, A. (1978). Why can't the deaf read? Comments on asking the wrong question. *American Annals of the Deaf, 123*, 524-547.

Gough, P. (1972). One second of reading. In J. Kavanagh & I. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp 331-358). Cambridge, MA: MIT Press.

Graesser, A. C. (1981). *Prose comprehension beyond the word.* New York: Springer-Verlag.

Graesser, A. C. & Clark, L, F. (1985). *Structures and procedures of implicit knowledge.* Norwood, N. J.: Ablex.

Graesser, A. C., Haberlandt, K., & Koizumi, D. (1987). How is reading time influenced by knowledge based inferences and world knowledge? In B. K. Britton & S. M. Glynn (Eds.), *Executive control processes in reading.* Hillsdale, NJ: Erlbaum.

Graesser, A. C. & Kreuz, R. J. (1993). A theory of inference generation during text comprehension. *Discourse Processes, 16*, 145-160.

Graesser, A. C., Lang, K. L., & Roberts, R. M. (1991). Question answering in the context of stories. *Journal of Experimental Psychology: General, 120*, 254-277.

Graesser, A. C. & Murachver, T. (1985). Symbolic procedures of question answering. In A. C. Graesser & J. B. Black (Eds.), *The psychology of questions.* Hillsdale, NJ: Erlbaum.

Graesser, A. C., Robertson, S. P., & Anderson, P. A. (1981). Incorporating inferences in narrative representations: A study of how and why. *Cognitive Psychology, 13,* 1-26.

Graesser, A. C., Robertson, S. P., Lovelace, E., & Swinehart, D. (1980). Answers to why-questions expose the organisation of story plot and predict recall of actions. *Journal of Verbal Learning and Verbal Behaviour, 19,* 110-119.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371-395.

Green, D & Shallice, T. (1976). Direct visual access in reading for meaning. *Memory and Cognition, 41,* 753-758.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics, Vol. 3: Speech Acts.* New York: Seminar Press.

Griswold, E & Cummings, J. (1974). The expressive vocabulary of preschool deaf children. *American Annals of the Deaf, 119,* 16-28.

Haber, R, N. & Haber, L. R. (1982). Does silent reading involve articulation? Evidence from tongue-twisters. *American Journal of Psychology, 95,* 409-419.

Haberlandt, K. & Bingham, G. (1978). Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triples. *Journal of Verbal Learning and Verbal Behaviour, 17,* 419-425.

Hammermeister, F. (1971). Reading acheivement in deaf adults. *American Annals of the Deaf, 116*, 25-28.

Hamp, H. (1975). *The picture aided reading test.* Northampton: Hamp.

Hansche, L. & Gordon, G. (1983). An investigation of the relationship of story schema to reading ability and grade level. In J. Niles & L. Harris (Eds.), Searches for meaning in reading/language processing and instruction. Thirty-second yearbook of the National Reading Conference (pp 255-259). Rochester, N. Y.: National Reading Conference.

Hansen, J. (1981). The effects of inference training and practice on young children's reading comprehension. *Reading Research Quarterly, 16*, 391-417.

Hansen, J. & Pearson, P. D. (1983). An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers. *Journal of Educational Psychology, 75*, 821-829.

Hanson, V. L. (1982). Short-term recall by deaf signers of American Sign Language: Implications of encoding strategy for order recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 572-583.

Hanson, V. L. (1986). Access to spoken language and the acquisition of orthographic structure: Evidence from deaf readers. *Quarterly Journal of Experimental Psychology, 38*, 193-212.

Hanson, V. L., Goodell, E. W., & Perfetti, C. A. (1991) Tongue-twister effects in the silent reading of hearing and deaf college students. *Journal of Memory and Language, 30*, 319-330.

Hanson, V. L., Liberman, I. Y., & Shankweiler, D. (1984). Linguistic coding by deaf children in relation to beginning reading success. *Journal of Experimental Child Psychology, 37*, 378-393.

Hanson, V. L. & Lichtenstein, E. L. (1990). Short-term memory coding by deaf signers: The notion of primary lenguage coding reconsidered. *Cognitive Psychology, 22*, 211-224.

Hardyck, C & Petrinovich, L. (1970). Subvocal speech and comprehension level as a function of the difficulty level of reading material. *Journal of Verbal Learning and Verbal Behaviour, 9*, 647-652.

Hardyck, C , Petrinovich, L & Ellsworth, D. H. (1966). Feedback of speech muscle activity during silent reading: Rapid extinction. *Science, 154*, 1467-1468.

Harrison, C. (1996). *Methods of teaching reading: Key issues in research and implications for practice.* Edinburgh: Scottish Office Education and Industry Department, Interchange Document No. 39.

Hatcher, C & Robbins, N. (1978). The development of reading skills in hearing-impaired children. Cedar Falls: University of North Iowa. (ERIC Document Reproduction service No ED 167 960).

Haviland, S. & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behaviour, 13*, 512-521.

Houck, J. (1982). The effects of idioms on reading comprehension of hearing impaired students. Unpublished doctoral dissertation, University of Northern Colorado.

Howarth, S., Wood, D., Griffiths, A., & Howarth, C. (1981). A comparative study of the reading lessons of deaf and hearing pry school children. *British Journal of Educational Psychology, 51*, 156- 162.

Huey, E. B. (1908/1968). *The psychology and pedagogy of reading*. Cambridge, M. A.: MIT Press.

Ikeda, M. & Saida, S. (1978). Span of recognition in reading. *Vision Research*, *18*, 83-88.

Iran-Nejad, A, Ortony, A, & Rittenhouse, R. (1981). The comprehension of metaphoric uses of English by deaf children. *Journal of Speech and Hearing Research*, *24*, 551-556.

Israelite, N. (1981). Direct antecedent context and comprehension of reversible passive voice sentences by deaf readers. Unpublished doctoral thesis, University of Pittsburgh.

Jarvella, R. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Behaviour*, *10*, 409-416.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, Massachusetts: Harvard University Press.

Johnson-Laird, P. N. & Stevenson, R. (1970). Memory for syntax. *Nature, 227*, 412.

Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329-354.

Just, M. A. & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Rockleigh, NJ: Allyn & Bacon.

Just, M. A. & Clark, H. H. (1973). Drawing inferences from the presuppositions and implications of affirmative and negative sentences. *Journal of Verbal Learning and Verbal Behaviour, 12*, 21-31.

Kail, R. V., Chi, M. T. H., Ingram, A. L., & Danner, F. W. (1977). Constructive aspects of children's reading comprehension. *Child Development, 48*, 684-688.

Keenan, J., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behaviour, 25*, 115-126.

Keenan, J., Golding, J., Potts, G., Jennings, T., & Aman, C. (1990). Methodological issues in evaluating the occurrence of inferences. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation* (Vol. 25). Orlando, Florida: Academic Press.

Keenan, J. & Kitsch, W. (1974). The identification of explicitly and implicitly presented information. In W. Kitsch (Ed.), *The representation of meaning in memory* (pp 153-166). Hillsdale, N. J.: Lawrence Erlbaum Associates.

Keenan, J., Potts, G., Golding, J., & Jennings, T. (1990). Which elaborative inferences are drawn during reading? A question of methodologies. In D. A. Balota, G. B. Flores d' Arcais, & K. Rayner. (Eds.), *Comprehension processes in reading*. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Keenan, J., Potts, G., Jennings, T., & Golding, J. (1988). Can naming detect inferences? Paper presented at the annual meeting of the Psychonomic Society, Chicago, Illinois.

Kennedy, R. A. (1979). Eye movements and reading. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory. London: Proceedings of the International Conference.

King, C & Quigley, S. (1985). *Reading and deafness*. London: Taylor & Francis.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W..& Mross, E. F. (1985). Context effects in word identification. *Journal of Memory and Language, 24*, 336-349.

Kintsch, W. & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review, 85*, 363-394.

Kleiman, G. (1975). Speech recoding in reading. *Journal of Verbal Learning and Verbal Behaviour, 14*, 323-339.

Klima, E. & Bellugi, U. (1979). *The signs of language*. Cambridge, Massachusetts: Harvard University Press.

Kluwin, T, Getson, P, & Kluwin, B. (1980). The effects of experience on the discourse comprehension of deaf and hearing adolescents. *Directions, 1*, 49.

Kosslyn, S. M. & Matt, A. M. (1977). If you speak slowly, do people read your prose slowly? Person-particular speech recoding during reading. *Bulletin of the Psychometric Society, 9*, 250-252.

Kretschmer, R. (1976). Judgments of grammaticality by 11, 14, and 17 year old hearing and hearing-impaired youngsters. Unpublished doctoral dissertation. University of Kansas.

Kretschmer, R. (1982). Reading and the hearing-impaired individual: Summation and application. *Volta Review, 84,* 107-122.

Kyle, J. (1980). Reading development of deaf children. *Journal of Research in Reading, 3,* 86-97.

Lake, D. (1978). Syntax and sequential memory in hearing-impaired children. In H. Reynolds & C. Williams (Eds), Proceedings of the Gallaudet conference on reading in relation to deafness. Washington, DC: Gallaudet College.

LaSasso, C. (1980). The validity and reliability of the cloze procedures as a measure of readability for prelingually, profoundly deaf students. *American Annals of the Deaf, 125,* 559-563.

Lesgold, A. M. (1974). Variability in children's comprehension of syntactic structures. *Journal of Educational Psychology, 66,* 333-338.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, Massachusetts: MIT Press.

Levy, B. A. (1975). Vocalisation and suppression effects in sentence memory. *Journal of Verbal Learning and Verbal Behaviour, 14,* 304-316.

Levy, B. A. (1977). Reading: Speech and meaning processes. *Journal of Verbal Learning and Verbal Behaviour, 16,* 623-638.

Leybaert, J. & Alegria, J. (1993). Is word processing involuntary in deaf children? *British Journal of Developmental Psychology, 11,* 1-29.

Lichtenstein, E. (1980). A proposal for the experimental investigation of recoding strategies used by deaf readers during reading comprehension. Centre for the Study of Reading, University of Illinois, Urbana.

Liddell, S. (1975). *Restrictive relative clauses in American Sign Language*. San Diego, CA: Salk Institute and University of California.

Linde, C. & Labov, W. (1975). Spatial networks as a site for the study of language and thought. *Language, 51*, 924-939.

Locke, J. (1978). Phonemic effects in the silent reading of hearing and deaf children. *Cognition, 6*, 175-187.

Locke, J & Locke, V. (1971). Deaf children's phonetic, visual and dactylic coding in a graphemic recall task. *Journal of Experimental Psychology, 89*, 142-146.

Long, D. L. & Golding, J. M. (1993). Superordinate goal inferences: Are they automatically generated during comprehension? *Discourse Processes, 16*, 55-74.

Long, D. L., Golding, J. M., & Graesser, A. C. (1992). A test of the on-line status of goal-related inferences. *Journal of Memory and Language, 31*, 634-647.

Long, D. L., Golding, J. M., Graesser, A. C., & Clark, L. F. (1990). Goal, event, and state inferences: An investigation of inference generation during story comprehension. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation: Inferences and text comprehension* (Vol. 25). New York: Academic.

Lovie, P. (1986). Identifying outliers. In A. D. Lovie (Ed.), *New developments in statistics for psychology and the social sciences*. London & New York: BPS-Methuen.

Lucas, M. M., Tanenhaus, M. K., & Carlson, G. N. (1989). Inferences in sentence comprehension: The role of constructed representations. Proceedings of the Ninth Annual Cognitive Science Society Meeting (pp 566-574). Hillsdale, N. J.: Lawrence Erlbaum Associates.

Lucas, M. M., Tanenhaus, M. K., & Carlson, G. N. (1990). Levels of representation in the interpretation of anaphoric reference and instrument inference. *Memory and Cognition, 18*, 611-637.

McCawley, J. (1968). The role of semantics in a grammar. In E. Bach & R. Harms (Eds), *Universals in linguistic theory* (125-169). New York: Holt, Rinehart and Winston.

McConkie, G. W. & Hoaboam, T. W. (1985). Eye position and word identification in reading. In R. Groner, G. W. McConkie, & C. Menz (Eds.), *Eye movements and human information processing*. Amsterdam: North-Holland Press.

McConkie, G. W. & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17*, 578-586.

McCutchen, D., Bell, L. C., France, I. M., & Perfetti, C. A. (1991). Phoneme-specific interference in reading: The visual tongue-twister effect revisited. *Reading Research Quarterly, 26*, 87-103.

McCutchen, D. & Perfetti, C. A. (1982). The visual tongue-twister effect: Phonological activation in silent reading. *Journal of Verbal Learning and Verbal Behaviour, 21*, 672-687.

McGill-Franzen, A & Gormley, K. (1980). The influence of context on deaf readers' understanding of passive sentences. *American Annals of the Deaf, 125*, 937-942.

McGuigan, F. J. (1967). Feedback of speech muscle activity during silent reading: Two comments. *Science*, *157*, 579-580.

McGuigan, F. J. & Bailey, S. C. (1969). Longitudinal study of covert oral behaviour during silent reading. *Perceptual and Motor Skills*, *28*, 170.

McKoon, G. (1988). Word identification and elaborative inference: The relation between theory and empirical measurement. Unpublished manuscript.

McKoon, G. & Keenan, J. (1974). Response latencies to explicit and implicit statements as a function of the delay between reading and text. In W. Kintsch (Ed.), *The representation of meaning in memory* (pp 166-176). Hillsdale, N. J.: Lawrence Erlbaum Associates.

McKoon, G. & Ratcliff, R. (1981). The comprehension processes and memory structures involved in instrumental inferences. *Journal of Verbal Learning and Verbal Behaviour*, *20*, 671-682.

McKoon, G. & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *12*, 82-91.

McKoon, G. & Ratcliff, R. (1989a). Assessing the occurence of elaborative inference with recognition: Compatibility checking versus cue theory. *Journal of Memory and Language*, *28*, 547-563.

McKoon, G. & Ratcliff, R. (1989b). Semantic associations and elaborative inference. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 326-338.

McKoon, G. & Ratcliff, R. (1989c). Inferences about contextually defined categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 1134-1146.

McKoon, G. Ratcliff, R. (1990). Dimensions of inference. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation: Inferences and text comprehension* (Vol. 25). New York: Academic.

McKoon, G & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440-466.

Maeder, J. (1979). *Comparative effectiveness of inserted postquestions in prose learning on the reading comprehension of hearing-impaired children.* Washington, DC: Model Secondary School for the Deaf.

Magliano, J. P., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedent and causal consequence inferences. *Discourse Processes*, *16*, 35-54.

Magliano, J. P. & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, *29*, 193-232.

Mandler, J. (1978). A code in the node: The use of a story schema in retrieval. *Discourse Processes*, *1*, 14-35.

Mandler, J. & Johnson, N. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*, 111- 151.

Mani, K. & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory and Cognition*, *10*, 181-187.

Marschark, M., Mouradian, V., & Halas, M. (1994). Discourse rules in the language productions of deaf and hearing children. *Journal of Experimental Child Psychology*, *57*, 89-107.

Masson, M. E. J. & Freeman, L. (1990). Fluent identification of repeated words. *Journal of Experimental Psychology: Learning, Memory an Cognition*, *16*, 355-373.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, *81*, 899-917.

Meadow, K. P. (1967). The effect of early manual communication and family climate on the deaf child's development. Unpublished doctoral dissertation, University of California, Berkeley.

Meadow, K. P. (1968). Early manual communication in relation to the deaf child's intellectual, social, and communicative functioning. *American Annals of the Deaf*, *113*, 29-41.

Meyer, D. E. & Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227- 234.

Millis, K. K. (1989). The time course of constructing bridging and expectation knowledge-based inferences during the comprehension of expository text. Unpublished doctoral dissertation, Memphis State University.

Mitchell, D. C. & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*, 609-636.

Moores, D. (1967). Applications of "cloze" procedures to the assessment of psycholinguistic abilities of the deaf. Unpublished doctoral thesis, University of Illinois at Urbana-Champaign.

Morrison, R. E. & Inhoff, A. W. (1981). Visual factors and eye movements in reading. *Visible Language, 15*, 129-146.

Morrow, D., Bower, G., & Greenspan, S. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language, 28*, 292-312.

Morrow, D., Greenspan, S., & Bower, G. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language, 26*, 165-187.

Moulton, R & Beasley, D. (1975). Verbal coding strategies used by hearing-impaired individuals. *Journal of Speech and Hearing Research, 18*, 559-570.

Murray, J. D., Klin, C. M., & Myers, J. L. (1991). Forward inferences about specific events during reading. Paper presented at the Psychonomic Society Conference, San Francisco.

Myers, J. L. (1990). Causal relatedness and text comprehension. In D. A. Balota, G. B. Flores d' Areais, & K. Rayner (Eds.), *Comprehension processes in reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Myers, J. L. & Duffy, S. A. (1991). Causal inferences and text memory. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation* (Vol. 25). Orlando, FL: Academic Press.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26*, 453-465.

Neely, J. H. & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 23, pp 207-248). Orlando, FL: Academic Press.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 1003-1019.

Nicholas, D. W. & Trabasso, T. (1981). Towards a taxanomy of inferences. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), *Information integration by children.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review, 8*, 231-259.

Nolen, S. B. & Wilbur, R. B. (1985). The effects of context on deaf students' comprehension of difficult sentences. *American Annals of the Deaf, 130*, 231-235.

Norman, D., Rumelhart, D., & LNR. (1975). *Explorations in cognition.* San Francisco: Freeman.

O'Brien, E. J. Myers, & J. L. (1985). When comprehension difficulty improves memory for text. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11*, 12-21.

O'Brien, E. J. & Myers, J. L. (1987). The role of causal connections in the retrieval of text. *Memory and Cognition, 15*, 419- 427.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 410-420.

Oakhill, J. & Garnham, A. (1988). *Becoming a skilled reader.* Oxford: Blackwell.

Odom, P., Blanton, R., & McIntyre, C. (1970). Coding medium and word recall by deaf and hearing subjects. *Journal of Speech and Hearing Research, 13*, 54-58.

Olson, G., Duffy, S., & Mack, R. (1984). Thinking- out-loud as a method for studying real-time comprehensive processes. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp 253-286). Hillsdale, NJ: Lawrence Erlbaum Associates.

Omanson, R., Warren, W., & Trabasso, T. (1978). Goals, inferential comprehension, and recall of stories by children. *Discourse Processes, 1*, 337-354.

O'Neill, M. (1973). The receptive language competence of deaf children in the use of the base structure rules of transformational generative grammar. Unpublished doctoral disertation, University of Pittsburgh, PA.

Page, S. (1981). The effect of idiomatic language in passages on the reading comprehension of deaf and hearing students. Unpublished doctoral dissertation, Ball State University

Paris, S. G. & Carter, A. Y. (1973). Semantic and constructive aspects of sentence memory in children. *Developmental Psychology, 9*, 109-113.

Paris, S. G. & Lindauer, B. K. (1976). The role of inference in children's comprehension and memory for sentences.*Cognitive Psychology, 8*, 217-227.

Paris, S. G. & Mahoney, G. (1974). Cognitive integration in children's memory for sentences and pictures. *Child Development, 45*, 633-642.

Paris, S. G. & Upton, L. L. (1976). Children's memory for inferential relationships in prose. *Child Development, 47*, 660-668.

Paul, P. (1984). The comprehension of multimeaning words from selected frequency levels by deaf and hearing subjects. Unpublished doctoral discertation, University of Illinois at Urbana-Champaign.

Payne, J. (1982). A study of the comprehension of verb-particle combinations amomg deaf and hearing subjects. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford.

Perfetti, C. A. (1989). There are generalised abilities and one of them is reading. In L. Resnick (Ed.), *Knowing and learning: Issues for a cognitive science of instruction*. Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (1993). Why inferences might be restricted. *Discourse Processes, 16*, 181-192.

Perfetti, C. A., Beverly, S., Bell, L., Rodgers, K., & Faux, R. (1987). Comprehending newspaper headlines. *Journal of Memory and Language, 26*, 692-713.

Perfetti, C. A. & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlabaum.

Perfetti, C. A. & Roth, S. F. (1981). Some of the interactive processes in reading and their role in reading skills. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading*. Hillsdale, NJ: Erlbaum.

Perrig, W. & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24*, 503-518.

Piaget, J. (1969). *The child's concept of time*. London: Routledge, Keegan, Paul.

Pinhas, J. (1984). Inferential processing abilities of deaf and hearing readers. Unpublished doctoral dissertation, Cornell University.

Polio, M. & Polio, H. (1979). A test of metaphoric comprehension and some preliminary developmental data. *Journal of Child Language*, *6*, 111-120.

Pollatsek, A, Bolozky, S., Well, A. D., & Rayner, K. (1981). Assymetries in the perceptual span for Israeli readers. *Brain and Language*, *14*, 174-180.

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception and Psychophysics*, *40*, 123-130.

Posner, M.·I. & Boies, S. W. (1971). Components of attention. *Psychological Review*, *78*, 391-408.

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Langauge*, *27*, 399-415.

Prinz, P. M. & Prinz, E. A. (1985). If only you could hear what I see: Discourse development in sign language. *Discourse Processes*, *8*, 1-19.

Quigley, S. & Frisina, R. (1961). Institutionalization and psychoeducational development of deaf children, CEC Research Monograph. Washington, DC: Council on Exceptional Children.

Quigley, S, Power, D, & Steinkamp, M. (1977). The language structure of deaf children. *Volta Review*, *79*, 73-84.

Quigley, S, Smith, N, & Wilbur, R. (1974). Comprehension of relativised sentences by deaf students. *Journal of Speech and Hearing Research, 17*, 325-341.

Quigley, S, Wilbur, R, & Montanelli, D. (1974). Question formation in the language of deaf students. *Journal of Speech and Hearing Research, 17*, 699-713.

Quigley, S., Wilbur, R., & Montanelli, D. (1976). Complement structures in the language of deaf students. *Journal of Speech and Hearing Research, 19*, 448-457.

Quigley, S., Wilbur, R., Power, D., Montanelli, D., & Steinkamp, M. (1976). *Syntactic structure in the language of deaf children.* Urbana, IL: University of Illinois, Institute for Child Behaviour and Development.

Quinn, L. (1981). Reading skills of hearing and congenitally deaf children. *Journal of Experimental Child Psychology, 32*, 139-161.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*, 65-81.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85*, 618-660.

Rayner, K. (1986). Eye movements and the perceptual span of in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*, 211-236.

Rayner, K. (1995). What the eye tells the brain during reading. Paper presented at Reading 2000: Recent advances in the science of reading. Seminar, University of Glasgow.

Rayner, K. & Bertera, J. H. (1979). Reading without a fovea. *Science, 206*, 468-469.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behaviour*, *22*, 358-374.

Rayner, K. & Pollatsek, A. (1989). *The Psychology of Reading*. London, Prentice-Hall International.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics*, *31*, 537-550.

Reich, P A, & Reich, C M. (1974). A follow-up study of the deaf, Toronto, Ontario, Toronto Board of Educational Research Report.

Richek, M. (1976-77). Reading comprehension of anaphoric forms in varying linguistic contexts. *Reading Research Quarterly*, *12*, 145-165.

Robbins, N & Hatcher, C. (1981). The effects of syntax on the reading comprehension of hearing-impaired children. *Volta Review*, *83*, 105-115.

Sachs, J. (1974). Memory in reading and listening to discourse. *Memory and Cognition*, *2*, 95-100.

Sanford, A. (1990). On the nature of text-driven inference. In D. A. Balota, G. B. Flores d' Arcais, & K. Rayner (Eds.), *Comprehension processes in reading*, pp 515-535. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Sanford, A. & Garrod, S. (1977). Implicit information in comprehending discourse. Paper presented at the 4th International Salzburg Linguistics Meeting.

Sanford, A. & Garrod, S. (1981). *Understanding written language: Explorations in comprehension beyond the sentence.* Chichester, Wiley.

Sanford, A., Garrod, S., Lucas, A., & Henderson, R. (1983). Pronouns without explicit antecedents? *Journal of Semantics, 2*, 303-318.

Sarachan-Deily, A. B. (1984). Written discourse of deaf and hearing students: Semantic analysis. Paper presented at the 9th Annual Boston University Conference on Language Development. Boston, Massachusetts.

Schank, R. C. (1973). Identification of conceptualisations underlying natural language. In R. C. Schank & K. M. Colby (Eds.), *Computer models of thought and language.* San Francisco: Freeman.

Schank, R. C. (1976). The role of memory in language processing. In C. Coffer (Ed.), *The nature of human memory.* San Francisco: W. H. Freeman Press.

Schank, R. C. & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.

Schustack, M., Erlich, S., & Rayner, K. (1987). Local and global sources of contextual facilitation in reading. *Journal of Memory and Language, 26*, 322-340.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Memory and Cognition, 12*, 315-328.

Shankweiler, D. & Crain, S. (1986). Language mechanisms and reading disorder: A modular approach. *Cognition, 24*, 139-168.

Sharkey, N. E. (1986). A model of knowledge-based expectations in text comprehension. In J. A. Galambos, R. P. Abelson, & J. B. Black (Eds.), *Knowledge structures*. Hillsdale, NJ: Erlbaum.

Sharkey, A. J. & Sharkey, N. E. (1992). Weak contextual constraints in text and word priming. *Journal of Memory and Language, 31*, 507-524.

Shatz, C. (1975). The development of social cognition. In E. M. Hetherington (Ed.), *Review of child development research*. Chicago, IL: University of Chicago Press.

Singer, M. (1978). The role of explicit and implicit recall cues. In Symposium: Implication and inference in language comprehension. Paper presented at the annual meeting of the American Psychological Association, Toronto.

Singer, M. (1979). Processes of inference during sentence encoding. *Memory and Cognition, 7*, 192-200.

Singer, M. (1980). The role of case filling inferences in the coherence of brief passages. *Discourse Processes, 3*, 185-201.

Singer, M & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behaviour, 22*, 437-448.

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language, 31*, 543-572.

Slowiaczek, M. L. & Clifton, C. (1980). Subvocalisation and reading for meaning. *Journal of Verbal Learning and Verbal Behaviour, 19*, 573-582.

Smith, P. T. & Groat, A. (1979). Spelling patterns, letter cancellation and the processing of text. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds). *Processing of visible language*, Vol 1, New York, Plenum.

Southgate, V. (1962). *Group reading test* (test 2, form A- sentence completion). Sevenoaks: Hodder & Stoughton.

Spache, G. D. (1953). A new readability formula for primary grade reading materials. *Elementary School Journal, 53*, 410-413.

Spiro , R. (1975). Inferential reconstruction in memory for connected discourse. (Technical Report No. 2). Urbana, Illinois: University of Illinois, Laboratory for Cognitive Studies in Education.

Spiro, R. (1977). Remembering information from text: Teoretical and empirical issues concerning the "State of Schema" reconstruction hypothesis. In R. Anderson, R. Spiro, & W. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp 137-165). Hillsdale, NJ: Lawrence Erlbaum Associates.

Stein, N. (1978). How children understand stories: A developmental analysis. (Technical Report No. 69),. Urbana Illinois: University of Illinois, Centre for the Study of Reading. (ERIC Document Reproduction Service No. ED 153 205).

Stein, N. & Glenn, C. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing: Advances in discourse processes* (Vol. 2), (pp 53-120). Norwood, NJ: Ablex.

Stern, J. A. (1978). Eye movements, reading, and cognition. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and higher psychological functions*. Hillsdale: Erlbaum.

Stoefen, J. (1981). Hearing impaired children's reading comprehension of anaphoric relationships. Unpublished doctoral thesis, University of Nebraska.

Strassman, B. (1992). Deaf adolescents' metacognitive knowledge about school related reading. *American Annals of the Deaf, 137*, 326-330.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Stuckless, E. & Birch, J. (1966). The influence of early mannual communication on the linguistic development of deaf children. *American Annals of the Deaf, 111*, 452-460, 499-504.

Stuckless, E. & Pollard, G. (1977). Processing of fingerspelling and print by deaf students. *American Annals of the Deaf, 122*, 475-479.

Taylor, I. & Taylor, M. M. (1983). *The psychology of reading*. New York, Academic Press.

Taylor, S. E. (1965). Eye movements while reading: Facts and fallacies. *American Educational Research Journal, 2*, 187-202.

Till, R. E., Mross, E. F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory and Cognition, 16*, 283-298.

Tinker, M. A. (1963). *Legibility of print*. Ames, Iowa: Iowa State University Press.

Tinker, M. A. (1965). *Bases for effective reading*. Minneapolis: University of Minnesota Press.

Trabasso, T, Nicholas, D. W., Omanson, R. C., & Johnson, L. (1977). Inferences in story comprehension. Paper presented at the Symposium on the Development of Discourse Processing Skills: Society for Research in Child Development, New Orleans.

Trabasso, T. & Sperry, L. L. (1985). Causal relatedness and the importance of story events. *Journal of Memory and Language*, *24*, 595-611.

Trabasso, T. & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes*, *16*, 3-34.

Trabasso, T. & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*, 612-630.

Trevor-Roper, P. D. (1955). *Ophthalmology: A textbook for diploma students*. Chicago: Year Book Publishers.

Trieman, R. A., Baron, J., & Luk, K. (1981). Speech recoding in silent reading: A comparison of Chinese and English. *Journal of Chinese Linguistics*, *9*, 116-124.

Trieman, R. A., Freyd, J. J., & Baron, J. (1983). Phonological recoding and use of spelling sound rules in reading of sentences. *Journal of Verbal Learning and Verbal Behaviour*, *22*, 682-700.

Trieman, R. A. & Hirsh-Pasek, K. (1983). Silent reading: Insights from second-generation deaf readers. *Cognitive Psychology*, *15*, 39-65.

Trueswell, J. C., Tanenhaus, M. K., & Garney, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285-318.

Underwood, N. R. & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20*, 153-162.

van den Broek, P. (1990a). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d' Arcais, & K. Rayner (Eds.), *Comprehension processes in reading*. Hillsdale, NJ: Erlbaum.

van den Broek, P. (1990b). Causal inferences and the comprehension of narrative text. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation: Inferences and text comprehension.* New York: Academic.

van den Broek, P. & Lorch, R. F. (1993). Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. *Discourse Processes, 16*, 75-98.

van den Broek, P. & Trabasso, T. (1986). Causal networks versus goal hierarchies in summarizing text. *Discourse Processes, 9*, 1-15.

van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

van Meter, P. & Pressley, M. (1994). Encoding of instruments when 10- to 14-year-olds process isolated instrument-implicit sentences: More evidence of improved encoding during childhood resulting from elaborative instructions. *Journal of Educational Psychology, 86*, 402-412.

Vernon, M. & Koh, S. (1970). Early manual communication and deaf children's achievement. *American Annals of the Deaf, 115*, 527-536.

Wallace, G. & Corballis, M. (1973). Short-term memory and coding strategies of the deaf. *Journal of Experimental Psychology*, *99*, 334-348.

Webster, A, Wood, D, & Griffiths, A. (1981). Reading retardation or linguistic deficit? I: Interpreting reading test performance of hearing impaired adolescents. *Journal of Research in Reading*, *4*, 136-147.

Webster, A. (1986). *Deafness, development and literacy.* London: Methuen.

Wender, K. (1989). Mental models for spatial knowledge. In A. F. Bennet & K. M. McConkey (Eds.), *Cognition in individual and social contexts.* Amsterdam: North-Holland.

West, R. F. & Stanovich, K. E. (1982). Automatic contextual facilitation in readers of three ages. *Child development*, *49*, 717-727.

White, A & Stevenson, V. (1975). The effects of total communication, manual communication, oral communication and reading on the learning of factual information in residential school deaf children. *American Annals of the Deaf*, *120*, 48-57.

Whitney, P. (1986). Processing category terms in context: Instantiations as inferences. *Memory and Cognition*, *14*, 39-48.

Whitney, P. & Kellas, G. (1984). Processing category terms in context: Instantiation and structure of semantic categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *10*, 95-103.

Wilbur, R, Fraser, J, & Fruchter, A. (1981). Comprehension of idioms by hearing-impaired students. Paper presented at the American Speech-Language-Hearing Association Convention, Los Angeles.

Wilson, K, Karchmer, M, & Jensema, C. (1978). Literal versus inferential item analysis of reading acheivement in hearing-impaired students. In H Reynolds C Williams (Eds), Proceedings of the Gallaudet conference on reading in relation to deaf.

Wilson, K. (1979). Inference and language processing in hearing and deaf children. Unpublished doctoral dissertation, Boston University.

Winer, B. J. (1971). *Statistical principles in experimental design.* New York: McGraw Hill.

Wolk, S & Schildroth, A. (1984). Consistency of an associational strategy used on reading comprehension tests by hearing-impaired students. *Journal of Research in Reading*, *7*, 135-142.

Wolverton,. G. S. & Zola, D. (1983). The temporal characteristics of visual information extraction during reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes.* New York: Academic Press.

Wood, D. (1980). A developmental psychologist looks - with teachers - at the education of deaf children. *Educational Analysis*, *2*, 61-83.

Wood, D. (1984). The assessment of linguistic and intellectual abilities of hearing impaired school children. *Journal of the Association of Educational Psychologists*, *6*, 31-39.

Wood, D, Griffiths, A, & Webster, A. (1981). Reading retardation or linguistic deficit? II: test-answering strategies in hearing and hearing-impaired school children. *Journal of Research in Reading*, *4*, 148-157.

Wood, D., Wood, H., Griffiths, A., & Howarth, I. (1986). *Teaching and talking with deaf children.* Chichester: Wiley.

Wrightstone, J, Aronow, M, & Moskowitz, S. (1963). Developing reading test norms for deaf children. *American Annals of the Deaf, 108*, 311-316.

Yuill, N. M. & Joscelyne, T. (1988). Effects of organisational cues and strategies on good and poor comprehenders' story understanding. *Journal of Educational Psychology,*

Yuill, N. M. & Oakhill, J. V. (1988). Effects of inference awareness training on poor reading comprehension. *Applied Cognitive Psychology,*

Zwaan, R. A. (1993). *Aspects of literary comprehension: A cognitive approach.* Philadelphia, PA: John Benjamins.

Zwaan, R. A. & van Oostendorp, H. (1993). Do readers construct spatial representations in naturalistic story comprehension? *Discourse Processes, 16*, 125-144.

Appendix A1

Textual material used in first study

Booklet 1:

(i) The man was running down the road.
The man tripped on the pavement.
The man broke his leg.
Could the man walk?
<u>No. The man could not walk.</u>

(ii) The lady was hungry.
The lady was eating a sandwich.
Then the lady had some tea.
Did the lady have the tea before eating the sandwiches?
<u>No. She ate the sandwiches first.</u>

(iii) The ball was under the chair.
Frank put the bat beside the ball.
Then Frank went out to play.
Was the bat under the chair?
....

(iv) Joe was drinking a cup of milk.
Joe put the cup on the table.
The cup was beside the plate.
Was the plate on the table?
....

(1) Mother cooked the dinner. (T)

Mother laid the table.

Then mother read a book.

Did mother read the book before laying the table?

....

(2) The cake was on the table. (S)

The cake was on a plate.

Mary cut the cake with a knife.

Was the plate on the table?

....

(3) The children were playing with a ball. (C)

John threw the ball to Susan.

The ball broke a window.

Did John break the window?

....

(4) The girl took the cake. (C)

Her mother was angry.

The cake was for tea.

Did mother eat the cake?

....

(5) The milkman dropped the bottle. (T)

The milkman got a brush and bucket.

Then the milkman cleaned up the mess.

Did the milkman clean up the mess before he dropped the bottle?

....

(6) The shirt was under the bed. (S)

The shoes were on the shirt.

The shirt was very dirty.

Were the shoes under the bed?

....

(7) The doll was in the bag. (S)

Jenny was playing with the toy cat.

Jenny put the toy cat in beside the doll.

Was the toy cat in the bag?

....

(8) The car crashed into the wall. (T)

The car was badly smashed.

Later the policeman came.

Did the car crash before the policeman came?

....

(9) The boy opened the door. (C)

The cat ran out.

His father was cross.

Did father let the cat out?

....

Booklet 3:

(10) The bread was in the cupboard. (S)

The butter was beside the bread.

The cheese was on the table.

Was the butter on the table?

....

(11) Kate is very thirsty. (C)

Kate pours out a glass of lemonade.

Kate washes her glass when she finishes.

Did Kate drink the lemonade?

....

(12) James was eating the apple. (T)

James finished the apple.

Then James put his book down.

Did James put his book down before eating the apple?

....

(13) The tree fell on the road. (C)

The car had to stop.

This made James late.

Was James driving the car?

....

(14) Mary and Tom were playing in the garden. (S)

Mary was looking for Tom.

Tom ran behind a wall.

Could Mary see Tom?

....

(15) The boy was playing football. (T)

The boy got a goal.

Then the boy went home.

Did the boy get the goal before going home?

....

(16) Mother had some money. (C)

All the money fell out of mother's purse.

Mother wanted to get some fruit.

Could mother pay for the fruit?

....

(17) Ann was walking along the street. (T)

Ann slipped on some ice.

Ann fell down.

Did Ann fall down before she slipped?

....

(18) Frank was cold. (S)

Frank put on his jacket.

Then Frank put on his coat.

Was Frank's coat over his jacket?

....

Booklet 2:

This contained the first line of each of the texts in Booklet 1, followed by blank lines for subjects to complete the rest of the story, eg

(1) Mother cooked the dinner.

....

....

etc

Booklet 4:

This was similar to booklet 2, only containing the first lines of those texts in Booklet 3, eg

(10) The bread was in the cupboard.

....

....

etc

Appendix A2

Instructions to teachers for the first study

INSTRUCTIONS TO THE TEACHER

The directions below are to be read to the group of children (or signed if the children are deaf and normally use TC).

The examples on pages 1-4 of the first booklet should be worked through with the pupils to ensure that they understand what is required. Any difficulties or misunderstandings should be explained at this point. No help, feedback or encouragement should be given except on the sample items. The pupils should be encouraged to write a fuller answer rather than simply giving a "yes" or a "no". The answer does not have to be grammatical so long as the meaning is clear,

eg. on item (3) ..... "The ball is under the chair"
etc.

ANSWER: Yes the bat was under the chair.
OR: Yes. Bat under chair.

Allow one minute for each page before telling the children to turn to each new item.

DIRECTIONS TO PUPILS:

Have the children complete the identification information on the front of each booklet. (There are 4 booklets in all).

"OPEN YOUR BOOK AT THE FIRST PAGE.
YOU WILL SEE SOME SENTENCES AND A BLANK PAGE FOR YOU TO ANSWER THE QUESTION."

Read the first paragraph aloud (or sign), then say:

"THE MAN TRIPPED AND BROKE HIS LEG. SO THE MAN COULD NOT WALK BECAUSE HE BROKE HIS LEG. YOU SHOULD WRITE YOUR ANSWER LIKE THIS."

Point to the answer given in the booklet.
Coach children if there are any misunderstandings as to what is required.
When everyone understands say:

"NOW TURN TO THE NEXT PAGE."

Read the paragraph aloud then say:

"NOW TURN TO THE NEXT PAGE."

Read the paragraph aloud then say:

"THE LADY WAS EATING THE SANDWICH, THEN THE LADY HAD SOME TEA. SO THE LADY DID NOT HAVE THE TEA BEFORE EATING THE SANDWICH. YOU SHOULD WRITE YOUR ANSWER LIKE THIS."

Point to the answer in the booklet.
When everyone has understood say:

"TURN TO THE NEXT PAGE. TRY THIS ONE BY YOURSELF."

Use the examples on pages 3 and 4 in a similar manner to that above.
After the example on page 4, say:

"THERE ARE SOME MORE SENTENCES LIKE THESE FOR YOU TO TRY ON YOUR OWN. AFTER A FEW PAGES YOU WILL BE ASKED TO REMEMBER WHAT YOU HAVE READ. TURN TO THE NEXT PAGE. ONCE YOU HAVE FINISHED, WAIT UNTIL I TELL YOU TO GO ON."

Allow the pupils to tackle each page, allowing one minute for completion. After completion of page 13, collect in all the booklets, then distribute the second booklet containing pages a-k. Then say:

"HERE IS THE FIRST SENTENCE OF SOME OF THE PAGES YOU HAVE JUST READ. TRY TO WRITE THE REST OF THE STORY."

Illustrate by using the two examples given on pages a and b.
Again allow <u>one minute</u> for completion of each page.
Once the second booklet has been completed and removed, give the third booklet containing pages 14 to 22. Then take this booklet away and give the fourth booklet, containing pages l to k, in a manner to that above.

Make sure that each child has written his or her name on each of the four booklets.

Appendix A3

Textual materials used in the on-line and signing studies

IN THE SHOP *(Practice item)*

Tony was hungry.

He went into the shop.

He bought some sweets.

Then he went home.

Did Tony Buy milk?

THE ACCIDENT *(Practice item)*

There was ice on the road.

Mary fell down.

She cut her knee.

She was crying.

Was Mary hurt?

IN THE PARK *(Practice item)*

It was a sunny day.

Fred went for a walk.

He went to the park.

He saw lots of flowers.

Did Fred go to the park?

BEDTIME *(Practice item)*

William was tired.

It was getting late.

He went to bed.

Soon he was asleep.

Did William stay up?

THE CAR *(Practice item)*

The car came round the corner.

It was going very fast.

It could not stop and crashed into the wall.

People stood and watched.

Did the car hit the wall?

GOING TO SCHOOL

John was on his way to school.

He went by bus.

He was very late getting to class. *Explicit*

He had to wait a long time. *Implicit*

His teacher was cross.

Did John arrive early at school?

IN CHURCH

It was Sunday.

Ian went to church.

After coming out, he saw his sister. *Explicit*

Three hours later, he saw his sister. *Implicit*

She was wearing a large hat.

Did Ian see his sister after leaving church?

THE FUN FAIR

Carol was at the fun fair.

She went on the big wheel.

After coming off, she smoked a cigarette. *Explicit*

An hour later, she smoked a cigarette. *Implicit*

She was having fun.

Did Carol smoke her cigarette on the big wheel?

SAILING TO AMERICA

Mr Jones was sailing to America.

He went onto the boat.

As the boat sailed away, he had a drink. *Explicit*

About three hours later, he had a drink. *Implicit*

He was sad to be leaving home.

Did Mr Jones have a drink before the boat sailed away?

IN THE KITCHEN

The dog was in the kitchen.

It was very hungry.

It ate the meat on the table. *Explicit*

It saw some meat on the table. *implicit*

Mother was angry.

Was there any meat left?

BATHTIME

It was Sunday night.

William was having a bath.

After his bath he started to sing loudly. *Explicit*

Two hours later he started to sing loudly. *Implicit*

It was his favourite song.

Did William sing after he had his bath?

WAITING AT THE BUS STOP

Jill was waiting for a bus.

She saw it coming.

It was full. *Explicit*

The bus ran past. *Implicit*

She was very annoyed.

Did the bus stop?

PLAYING GAMES

Frank was at his friend's house.

They played games.

After he went home, he drank some lemonade. *Explicit*

About five hours later, he drank some lemonade. *Implicit*

He was very thirsty.

Did Frank drink the lemonade at his friend's house?

HORSE RIDING

A noise frightened the horse.

Andrew fell from the horse.

He was very dirty. *Explicit*

He landed in mud. *Implicit*

The horse ran away.

Was Andrew clean?

THE TELEVISION

The television broke down.

The repair man came to the house.

After he left, mother made a cup of tea. *Explicit*

Three hours later, mother made a cup of tea. *Implicit*

There were biscuits and cakes.

Did the repair man have a cup of tea?

IN THE KITCHEN

Mary was cutting bread.

She was making sandwiches.

She cut herself badly. *Explicit*

The knife slipped suddenly. *Implicit*

She managed to finish the job.

Did Mary get hurt?

FOOTBALL

James went to the football match.

It was a good game.

After the game he saw his sister.

Three hours later he saw his sister.

He kissed her.

Did James kiss his sister after the game?

IN LONDON

Joe went to London.

He travelled all day.

The car stopped outside the hotel. *Explicit*

He parked just outside the hotel. *Implicit*

It was getting dark.

Was the car far away?

AT THE SEASIDE

Michael was on holiday.

He went for a swim in the sea.

After his swim, he shouted to his friends. *Explicit*

Two hours later, he shouted to his friends. *Implicit*

They waved to him.

Did Michael shout to his friends before going into the sea?

THE TRAIN

John was coming back from school.

He took the train home.

Getting off the train, it got very warm. *Explicit*

About two hours later, it got very warm. *Implicit*

He bought some ice-cream.

Did John buy ice-cream after getting off the train?

THE DANCE

Alice went to a dance.

It was great fun.

After the dance, she met her friend. *Explicit*

Five hours later, she met her friend. *Implicit*

They talked about school.

Did Alice meet her friend after the dance?

A WINTER'S DAY

David was skating on the frozen pond.

The ice was very thin.

The ice cracked and he fell into the water. *Explicit*

He was far too heavy for the thin ice. *Implicit*

He shouted for help.

Did the ice break?

THE DRESS

The purse was hanging out of mother's bag.

She was looking at some clothes.

A thief took her purse. *Explicit*

A thief stood behind her. *Implicit*

She wanted to get a dress.

Was mother's purse stolen?

THE SHOP WINDOW

Alan walked along the street.

He looked in a shop window.

After he went home, a car crashed into the shop.

About thirty minutes later, a car crashed into the shop.

The car was badly damaged.

Did Alan see the crash?

IN A FIELD

Alan crossed the field.

It had been raining.

When he came home, his shoes were wet and dirty. *Explicit*

When he came home, he left dirty marks on the carpet. *Implicit*

He went straight to his room.

Were Alan's shoes covered in mud?

THE RAILWAY STATION

Margaret was in the railway station.

She got on the train.

As it moved away, she read a book.

Half an hour later, she read a book.

She was going to London.

Was the train moving when Margaret read her book?

THE LETTER

Catherine sat down at the table.

She was writing to her uncle.

She finished the letter and had a cup of tea. *Explicit*

About an hour later she had a cup of tea. *Implicit*

The tea was hot.

Did Catherine have tea before finishing the letter?

THE SWIMMING BATHS

Jim was on holiday.

He went to the swimming baths.

He dived from the edge into the water. *Explicit*

He dived from the edge of the pool. *Implicit*

He was enjoying himself.

Did Jim get wet?

THE CINEMA

Jane went to the cinema.

She saw a good film.

After going home, she ate some crisps. *Explicit*

Five hours later, she ate some crisps. *Implicit*

They were very salty.

Did Jane eat the crisps in the cinema?

DINNER

Ann was hungry.

She had her dinner.

She had some chicken. *Explicit*

The chicken was tasty. *Implicit*

She washed the dishes.

Did Ann eat the chicken?

A DAY IN THE COUNTRY

Tom was driving in the country.

The car stopped suddenly.

He noticed that the petrol tank was empty. *Explicit*

He took an empty can to the garage. *Implicit*

He was vross about all this.

Did the car run out of petrol?

CLEANING THE WINDOWS

Denis fell off the ladder.

His mother phoned for help.

The ambulance rushed off to hospital. *Explicit*

The ambulance rushed along the road. *Implicit*

Mother followed in her car.

Did Denis go to hospital?

GOING HOME

Martin left to catch his bus.

He ran down the street.

He just missed the bus. *Explicit*

He was just too late. *Implicit*

It started to rain.

Did Martin get on the bus?

THE CAFE

Mary went into the cafe.

She had a cup of tea.

She left the cafe and smoked a cigarette. *Explicit*

About an hour later she smoked a cigarette. *Implicit*

She was wearimg a red coat.

Did Mary smoke the cigarette in the cafe?

SHOPPING

Frank came out of the shop.

It started to rain.

He did not have an umbrella and got all wet. *Explicit*

He did not have an umbrella and had to run. *Implicit*

He hurried all the way home.

Did Frank stay dry?

Appendix A4

VCR Program for Signing Study

```
10REM FILE "VCRPROG"
20REM © F.Shoveller.
30REM      FLOW CHART
40REM      SETUP & LOAD PREVIOUS DATA
50REM      WAIT FOR TONE FROM V.C.R.
60REM      DETECT TONE. GIVE AUDIO
70REM      SET TIME=0  START CLOCK
80REM      WAIT FOR KEY PRESS
90REM      IF KEY NOT PRESSED (15s) LOOP
100REM     TIMEOUT. GIVE AUDIO. LOOP 50
110REM     STORE TIME TO KEY PRESS
120REM     GIVE AUDIO. DISPLAY ON MONITOR
130REM     CHECK ANSWER
140REM     FLAG ANSWER IF WRONG "-TIME"
150REM     STORE TIME (& ANSWER)
160REM     QUESTION NO. IN DATA STATEMENT
170REM     DO AVERAGES ON CORRECT & STORE
180REM     LOOP50
190REM     REMEMBER TO BACKUP REGULARLY.
200:
210REM START
220MODE7
230VDU23,1,0;0;0;0;
240PROCinit:PROCinit2
250PROCignore
260PROCget
270PROCreset:PROCinit2
280END
290:
300DEFPROCtest
310REPEAT:T=TIME/100:PRINTTAB(31,13);T-.03;"
    ":X=INKEY(1)-48:UNTILX>0 OR T-.02>15
320ENDPROC
330:
340DEFPROCtone
```

```
350 PRINT'H$
360PRINT'"Waiting on cue tone from V.C.R."

370REM ADVAL(1)>2000 ? NOISE INHIBIT
380REM TEST
390REM IFGET:TIME=0:GOTO410
400 REPEAT:UNTILADVAL(1)>2000:TIME=0
410VDU7:PRINT'"B82Cue Detected"
420PRINT'"Clock zeroed and running."
430PRINT'"Waiting on answer key press."
440PRINT'"Elapsed time to answer (Secs) "
450REM TEST
460REM PROCtest:GOTO480
470REPEAT:T=TIME/100:X=(ADVAL(0) AND
    3):PRINTTAB(31,13);T-.03;"  ":UNTIL X>0 OR T-.02>15
480IFT-.02>15 PRINT'"B86TIMEOUT":GOTO510
490VDU7:PRINT'"B81Key Detected"
500REM X=1=YES 2=NO T>15s=timeout
510ENDPROC
520:
530DEFPROCinit
540N%=0:X%=0:Y%=0:X=0:T=0
550REM SOUND OFF
560*FX210,1
570@%=&20209
580DIMA(40,30)
590REM GROUPS OF 15, 40 QUESTIONS
600REM WRONG ANSWERS.. -VALUES
610REM OPEN CHANNEL 0 ONLY
620CLOSE#0
630*FX16,1
640PRINT''"WHICH TAPE? 1 or 2":W%=GET
650REM R%=Data set EDIT IF REN.
660REM W%=WHICH?
670IFW%=49 THEN D$="TAPE1" ELSE D$="TAPE2"
680PRINT';D$;" SELECTED"
690PRINT'"Loading data for ";D$; "....WAIT."
700X=OPENIN D$
710FOR Y%=0 TO 30:FOR X%=0 TO 40
720INPUT#X,A(X%,Y%)
730NEXT:NEXT
740CLOSE#X
```

```
750PRINT'"Data loaded."
760ENDPROC
770:

780DEFPROCinit2
790PRINT'"WHICH PUPIL? (D)eaf or (C)ontrol ?
     ";:P%=GET:PRINT''"FINDING NEXT RECORD"
800REM SELECT CORRECT HALF OF ARRAY
810REM i.e. Deaf or Control
820IFP%=68 FORN%=1 TO 15:IF A(1,N%)=99
     Y%=N%:H$="DeafB83"+STR$(Y%)+"using"+D$:N%=15
830IFP%=68 NEXT
840IFP%=67 FORN%=16 TO 30:IF A(1,N%)=99
     Y%=N%:H$="ControlB83"+STR$(Y%-15)+"using
     "+D$:N%=30
850IFP%=67 NEXT
860PRINT'H$
870PRINT'"Start V.C.R., then press any key.":IFGET
880ENDPROC
890:
900DEFPROCignore
910REM IGNORE 1st. 5
920FOR N%=1 TO 5:CLS
930PRINT'"Ignoring preliminary question";N%
940PROCtone
950PROCdelay
960NEXT
970ENDPROC
980:
990DEFPROCget
1000IF D$="TAPE1" THEN RESTORE1290 ELSE RESTORE1310
1010FOR N%=6 TO 45:CLS
1020READA%
1030IFA%=999:N%=45:NEXT:PROCsave:ENDPROC
1040X%=N%-5:PRINT'"Logging data on question";N%
1050PROCtone
1060IF T>15 A(X%,Y%)=0:GOTO1100
1070T=T-.03:REM"B83CORRECTION FACTOR
1080IFA%=X THEN PRINT'"Correct";:A(X%,Y%)=T
     ELSEPRINT'"Wrong";:A(X%,Y%)=-T
1090PRINT"answer given in";T;"Secs."
1100PROCdelay:NEXT:ENDPROC
```

```
1110:
1120DEFPROCdelay
1130T=TIME:REPEAT:UNTILTIME-T>300:ENDPROC
1140:
1150DEFPROCsave:CLS

1160PRINT'"Saving data to disc."
1170PRINT'"Do NOT touch keyboard!!"
1180X=OPENOUT D$
1190FOR Y%=0 TO 30:FOR X%=0 TO 40
1200PRINT#X,A(X%,Y%)
1210NEXT:NEXT:CLOSE#X:PRINT'"Data
    saved.....BACKUP?":PROCdelay:ENDPROC
1220:
1230DEFPROCreset
1240CLS:PRINT'"IF YOU ONLY REQUIRE ANOTHER RUN,
    REWIND TAPE AND PRESS ANY KEY WHEN READY.":IFGET
1250ENDPROC
1260:
1270REM 1=YES 2=NO
1280REM TAPE 1
1290DATA
```

Appendix A 5

Textual materials for the verb study

The first four passages are practice items.

The experimental passages are presented in the following form:

First sentence.

Second sentence: *Instrument stated / unstated form*

Third sentence: *Noun form / pronoun form*

Question: *Affirmative response / negative response*

Susan went to her desk.
She looked at the book. *Practice item*
The book was interesting.
Was there a book on the desk?

Simon was fixing the roof.
He went up a ladder. *Practice item*
The ladder was against the wall.
Did Simon stand on a box?

Peter was a good artist.
He was painting a new picture. *Practice item*
The brush was not very good.
Was Peter using a pencil?

Marie was now a big girl.

She was the same height as her brother. *Practice item*

She was six feet tall.

Was Marie tall?

David was in the pool.

He moved through the water. / He swam to the edge.

The water was deep. / It was very deep.

Was David in the water / on the ground?

Jill was very hungry.

She took all the food. / She ate every last crumb.

The food was tasty. / It was very tasty.

Did Jill eat food / nothing?

William sat down.

He did up his lace. / He tied his shoe.

The lace had come undone. / It had come undone again.

Had the lace been loose / tight?

Peter was in the garden.

He was using a spade. / He was digging up some flowers.

The spade was heavy. It was very heavy.

Was Peter using a spade / hammer?

Susan was in the country.

She pointed the gun at some rabbits. / She tried to shoot some rabbits.

The gun was old. / It was very old.

Did Susan use a gun / knife?

Simon got out of bed.

He used his toothbrush. / He cleaned his teeth.

The toothbrush was green. / It was bright green.

Did Simon use a toothbrush / cloth?

Tom was going to London.

He was taking his car. / He was driving there.

The car had been fixed. / It had just been fixed.

Did Tom go to London by car / train?

Jim went to the hairdressers.

The lady used her scissors. / The lady cut his hair.

The scissors were sharp. They were very sharp.

Did the lady use her scissors / a razor?

Frank was out for a walk.

The pond was covered with ice. / The pond had frozen over.

The ice was thin. / It was very thin.

Was the pond covered by ice?

Andrew was on holiday.

He used his new camera. / He photographed his friends.

The camera was a present. / It was a new present.

Did Andrew have a camera / telescope?

Martin was on his way to school.

He put some letters in the postbox. / He posted some letters to his friends.

The postbox was painted red. / It was painted bright red.

Were the letters in the postbox / litterbin?

Sally went to the stables.

She got on a horse. / She was going to ride home.

The horse was young. / It was very young.

Was Sally on a horse / bike?

Mother was in the kitchen.

She put some cakes in the oven. / She was baking some cakes.

The oven was hot. / It was very hot.

Were the cakes in the oven / cupboard?

Jack was enjoying himself.

He liked it on the boat. / He liked to go sailing.

The boat was big. / It was very big.

Was Jack on a boat / bus?

Mary was going shopping.

She put the clothes on the baby. / She dressed the baby.

The clothes were made of blue wool. / They were made of soft blue wool.

Did the baby have blue clothes?

Anne was in the cafe.

She had a cigarette. / She was smoking.

The cigarette was in the ashtray. / It was lying in the ashtray.

Did Anne have a cigarette / pipe?

Elizabeth went to the cinema.

She paid her money. / She bought a ticket.

The ticket was for the balcony. It was for a balcony seat.

Did the ticket cost money?

Ian went home after school.

He put his keys in the door. / He unlocked the door.

His keys got stuck. / But they got stuck.

Did Ian use his keys? / Did Ian ring the bell?

Jane went to America.

She went by areoplane last week. / She flew there last week.

The plane was fast. / It was very fast.

Did Jane go to America by plane / boat?

Margaret sat down at the table.

She used her spoon. / She stirred the tea.

The spoon was shiny. / It was very shiny.

Did Margaret use a spoon?

Appendix A6
Calibration of the Purkinje Dual Image Generation 5.5 Eyetracker

Tracker Position and Beam Alignment:

Turn **Power** and **Servo** on. It is worth leaving the Servo on for 5-10 minutes to warm up. Before turning the main Power on, make sure that the servo switch is off. Similarly, after tracking always ensure that Servo is off before switching power off.

When the subject is in position, still, and fixating on the fixation point, the task is to centre the **Infra red beam** in the **pupil plane** of the subject's eye. When it is in that plane, turning **Set Up/Track** on should result in a lock. The beam locks onto two of the four **purkinje images** that are formed as the light source passes through the eye. The first purkinje image is a virtual image formed by the cornea. It is up to two hundred times more intense than the fourth image which is formed by the posterior surface of the lens (the second and third purkinje images are not needed for tracking). It is by monitoring the relative positions of the first and fourth images that eyemovements can be tracked accurately.

Aligning the beam is achieved by moving the tracker housing. It is set up on motors and these are controlled via three switches on the front panel: **In/out, up/down,** and **left/right.** This task is made easier by viewing the video picture of the subject's eye (via a small infra red camera situated on the housing). The video picture is split in two. The left half gives a side view of the right eye, and the right half a front view (only movements of the right eye are tracked).

First, align the in/out dimension. Monitoring the left hand side of the screen, flick the in/out switch until the surface of the eye is level with the **alignment mark** on the screen. Altering the in/out dimension will move the position of the **dlachroic mirror** to about 3/4 of an inch from the subject's eye. This mirror reflects the beam onto the eye at an angle of 30 degrees: it is transparent, so subjects are not aware of its presence while focusing on the fixation point and subsequent text.

When the surface is level with the alignment mark, use the up/down and left/right switches to centre the beam on those dimensions. The right half of the video screen is most helpful at this stage. The beam will appear as a green spot over the subject's eye. When the three dimensions are centred, both the first and fourth purkinje images should be visible as small specks of light over the pupil. Now put the Set Up/Track on. If the beam has been accurately centred, a lock should be obtained immediately. **System lock** is signalled by a bank of lights on the front panel changing from red to green. If there is no immediate lock, adjust on the three dimensions in the usual manner (usually a small change to the up/down or left/right is all that is needed).

When a lock is obtained, each of the three dimensions has to be fine-tuned. Fine-tune the in/out first. Turn the dial in the centre of the control panel to **A 1.** Now the LED read-out refers to this dimension, and it has to be set to zero (it will usually vary between about -5 and +5 in the first instance). Flicking the in/out switch will decrease or increase the level. Any value between -1 and +1 is acceptable. When this dimension is zeroed, turn the **Focus Centre** switch on.

Zero the up/down and left/right dimensions in the same manner. The corresponding dial positions are **V 1** (vertical) and **H 1** (horizontal). When this is complete, an automatic alignment algorithm is used th centre the position of the fourth detector. Turn the dial to ∂**H** and pull the **4th detector auto position** switch. Release the switch when the whirring noise discontinues. Perform the same procedure with the dial at ∂**V**. Both these parameters should now be zeroed (anything less than 0.5 is acceptable). Their values are also displayed on the LED read out. Finally, check that the light level on the fourth image is set. Select the **4LL** dial position and use the rotary pot to settle the LED read-out at (approximately) 5.0.

Although beam alignment involves a number of sub-procedures, in practice all of these stages can be performed quite quickly (providing the subject is still and trackable). It is worth learning to align the beam quite fast since during this stage subjects are left looking at a single fixation point and are more prone to extended blinks.

Calibration:

The purpose of the calibration is to test the accuracy of the lock obtained. When all the beam alignment is complete, hit any key on the keyboard and the subject will be cued to start the calibration. A single flashing cross will appear on the main screen. When they are looking directly at this cross, they press the left-hand pace (response) key, and within a short space of time the cross moves to a different screen position. When they have re-fixated it, pressing the key again has the same effect.

There are about a dozen successive crosses to be fixated. After the final cross, the slave screen (ie the screen linked to the computer which is running the experiment) displays two banks of x and y coordinates: the left-hand bank gives the **tracker readout** of eye-position as the pace key was hit on each occasion; the right-hand bank gives the **actual on-screen position** of the crosses. Providing the tracker is aligned accurately and the subject is hitting the key while fixating directly on the crosses, these banks of coordinates should be almost identical. If there are marked discrepancies, check to see if they occur in a particular part of the visual field (eg all low x coordinate inaccuracies) or ask the subject to perform the calibration again stressing the importance of looking directly at the cross as they hit the key.

There is no time constraint in calibration. However, some subjects seem to move from cross to cross as fast as they can. If subjects calibrate poorly, ask them to repeat the procedure with less haste. If a second calibration is inaccurate, it suggests a deeper problem in beam alignment. Perhaps the subject has moved unintentionally, or there is some interference from eyelashes.

When the subject is calibrated, press 'c' to commence the reading phase. If anything goes wrong during the reading and the procedure has to be stopped, always re-calibrate after setting the subject up again. Similarly, after each break in the reading, re-align and re-calibrate in the usual fashion.