Improving the Usability of Speech-Based Interfaces for Blind Users

Ian J. Pitt & Alistair D.N. Edwards

Human-Computer Interaction Group,
Department of Computer Science,
University of York, Heslington,
York, UK YO1 5DD

ianp@minster.york.ac.uk

ABSTRACT

Adaptations using speech synthesis provide a basic level of access to computer systems for blind users, but current systems pose a number of usability problems. A study was carried out in order to assess the impact of certain issues on the usability of a typical speech adaptation. The results suggest that much work needs to be done on the design of speech dialogues.

1. INTRODUCTION

The ever-increasing use of computers in almost every aspect of our daily lives presents both problems and opportunities for people with visual impairments. If computers can be made as easy for blind people to learn and use as they are for sighted people, the steady computerisation of tasks in education, industry and leisure will open up many new opportunities for blind people. If, on the other hand, computers continue to be made such that blind people cannot use them, or can only use them with limited efficiency, blind people will not only miss out on new opportunities but may also find themselves excluded from activities which were open to them in the past.

Blind people are forced to rely on either tactile displays or sound (synthetic speech) in order to receive output from their computers. Tactile displays have a number of advantages but they are expensive and require that the user can read braille. This rules them out for many people because only a very small proportion of blind people are braille readers.

The vast majority of blind computer users are therefore forced to rely on synthetic speech. This avoids the problem of access for non-braille readers and is comparatively cheap, but it also presents other problems. Unlike printed text or braille, speech is transient. A line of text or a row of braille cells can be scanned repeatedly until the message has been absorbed. By comparison, speech has a fixed duration and must be held in memory while a message is being processed. This makes speech slow to use and more demanding of memory than vision or touch.

ASSETS '96, Vancouver, British Columbia Canada

If the efficiency with which blind people can use computers is to be improved, the way in which information is received through speech needs to be considered carefully. In this paper we identify a number of characteristics of speech which are important contributors to the way in which speech is understood, and consider how these characteristics are supported in typical text-to-speech converters for blind users. We identify a number of limitations of existing devices which, if our review of the literature is correct, may impede blind users' understanding of such speech output. We explore this hypothesis in a practical study which examines the way blind and sighted subjects interact with a piece of software presented solely through synthetic speech.

2. SCREEN READERS FOR THE BLIND

A typical speech-based adaptation for a blind computer-user comprises two distinct parts, a screen reader and a speech synthesizer. The screen reader is a piece of software which is loaded into memory when the computer is started-up and runs in the background while the user is running other programs. Its function is to read the screen memory and send any text it finds there to a speech synthesizer. The speech synthesizer can be either another piece of software or, more commonly, a hardware device located on a plug-in card or connected externally via one of the output ports.

The usability of such adaptations depends heavily on the functionality of the screen reader. A screen reader which simply sends *all* the text on the screen to the speech synthesizer is likely to be of little use. The user needs to know when new text appears on the screen and to be able to select previously-output text and hear it again. Therefore most screen readers provide at least the following basic features:

* new lines appearing at the bottom of an upward-scrolling display are automatically spoken out
* the cursor can be temporarily removed from the current insertion point and moved around the screen, selecting any other line of text to be spoken out
* the screen reader can be set-up to monitor particular areas of the screen and read out any new text which appears there (e.g., a dialogue box associated with a particular application)

* the way in which the system deals with numbers, punctuation, special characters, groups of capitals, etc., can be varied to suit the application

* the screen-reader can be configured to recognise different applications software and load pre-selected options as appropriate.

Taken together, these features provide blind users with at least a basic level of access to text-based software. However, we shall now consider how a speech-based adaptation of this type performs in the light of current psychological and linguistic knowledge.

2.1 Problems with existing adaptations

2.1.1 Quantity of information

Converting the whole of a piece of text into speech can make huge demands of a listener. For example, using the DIR command on a DOS system to obtain a directory listing produces a large quantity of text which can take a minute or more to read out as speech. Of this, it is unlikely that a listener will be able to recall more than a few items (e.g., Murray, 1966; Engle, 1974; Penney, 1979)

The problem can be overcome to some extent by allowing the user to move around the screen selecting short items of text to be spoken, and also by providing a 'mute' facility to halt the torrent of speech. However, even with these aids a blind user may take considerably longer to extract the required information than a sighted person scanning the list visually.

2.1.2 Order of presentation of information

Research suggests that we are best at remembering the last few items in a string (the recency effect) and that we are more likely to remember these items if they are followed by silence rather than by further items of speech (the suffix effect, described by Conrad, 1960 and Crowder & Morton, 1969). Thus we would expect it to be beneficial if the important information in a speech string is placed towards the end, immediately preceding a period of silence.

In practice, operating systems designed for visual display tend to use the grammatical structures of written language, with the result that important information may be placed almost anywhere within a text string. Even when the important information is placed at the end of a sentence, it may be followed by other sentences which, because of the shortcomings of the speech conversion system, may follow on with little or no pause.

2.1.3 Use of pauses

Psycho-linguistic research suggests that the clause is the fundamental unit of language comprehension (Jarvella, 1971; Abrams & Bever, 1969), and that pauses are an essential aid in delineating clauses. There is also evidence that pauses provide vital processing time during speech comprehension (Reich, 1980).

Pauses can be broadly categorised into two groups: grammatical pauses, which separate one clause from another, and non-grammatical pauses, which occur between the elements of a clause. We normally expect the pauses which occur within clauses to be shorter than those which separate clauses, and where this is not true the speech becomes more difficult to interpret. Reich (1980) found that moving a 330ms pause to an inappropriate place in a stimulus sentence increased the average response time by more than two seconds.

The rules used by most speech synthesis systems to handle pauses do not take account of these factors. They are often designed to handle text which runs over several lines, inserting non-grammatical (long) pauses at the punctuation points and grammatical (short) pauses at line breaks. This makes sense for longer pieces of text but it has unfortunate results when used with the kind of output typical of many command-line operating systems - short statements displayed one-per-line over several lines. In this case the result is all too often completely ungrammatical: long pauses occur in the middle of lines, often triggered by a punctuation character in a filename, while the end of one output line runs with only the briefest of pauses into the next.

2.1.4 Prosody

The natural rise and fall in pitch and volume over the course of a sentence is used, along with pauses, to help delineate clauses and so aid comprehension (Nakatani & Schaffer, 1978; Streeter, 1978). It also helps the listener to determine where they are within the sentence, thus allowing efficient use of resources such as memory.

Many speech synthesis systems provide either no prosodic cues at all or limited prosody which is often inappropriate and so counter-productive. For example, the prosody on many commercial systems is driven by the same rules that are used to determine the placing and length of pauses, so if the pausing is inappropriate (for the reasons discussed earlier) the prosody will be too. This merely compounds the problem for the listener.

2.1.5 Pronunciation and speech quality

Research shows that poor pronunciation and other forms of degradation significantly increase the time and effort involved in assimilating speech (Nusbaum & Pisoni, 1985). In particular, there is evidence that poor speech quality makes additional demands upon the short-term memory (Luce, Feustel & Pisoni, 1983).

Many speech synthesis systems currently in use offer relatively poor speech quality. This is partly because most synthetic speech is generated by rules which manipulate a comparatively limited number of acoustic cues. Some studies have shown significantly lower recognition rates and/or increased recognition times for words presented using synthetic speech compared with those presented using natural speech (e.g., Pisoni, 1981)

2.1.6 Use of non-speech sounds

Non-speech sound is widely used for simple error indications but finds very little other use in current adaptations for blind users. This is a shame, since research suggests that most non-speech sounds do not occupy as much space in short term memory as speech, and that non-speech sound causes less disruption to other cognitive processes (e.g. Crowder & Morton, 1969; Ayres, Jonides, Reitman, Egan & Howard, 1979). In view of this, we might ask whether speech is the best medium to use in all cases or whether we might usefully replace it with non-speech sounds in some cases.

3. EVALUATION

A study was carried out in order to assess the extent to which the above issues affect the usability of speech adaptations. A piece of software designed for use by sighted people was presented to a group of blind and sighted subjects using a commercial screen-reader and speech-synthesizer. The subjects' performance at the task and their responses to questions about its usability were recorded and analysed in order to identify any problems introduced by the use of the speech adaptation.

3.1 The evaluation system

The software used was a version of the spelling game 'Hangman'. The program chooses a word at random from a standard dictionary and then invites the player to guess the word, one letter at a time. Players have eleven 'lives', one of which is lost each time the player enters an incorrect letter. The object is to identify the word before all the lives are lost.

The program uses keyboard input and presents its output in the form of text on the screen. The output was converted into speech by a Hal screen-reader and an Apollo speech synthesizer. This combination of screen-reader and speech synthesizer is widely used by blind people.

The game produces the following response each time a letter is entered:

> There are <number> occurrences of the letter <letter>
>
> You have <number> lives left
>
> The letters you have found are: <letter 1> <letter 2> ... <letter n>
>
> Please enter a letter

This response presents no problems when read from the screen by a sighted user. However, in the light of the research findings discussed earlier we would expect it to be difficult to assimilate when rendered into speech for a number of reasons.

* It could be argued that the only important items in this dialogue are the number of lives left and the current arrangement of letters and spaces in the partly-guessed word, and even this information may not be required after every move. Thus there is a considerable amount of extraneous speech.

* The Hal/Apollo combination reproduced the text with negligible pauses between lines but with a long pause after the colon in the third line. The output was thus heard as:

  There are <number> occurrences of the letter <letter> you have <number> lives left the letters you have found are colon

  (pause)

  <letter 1> <letter 2> ... <letter n> please enter a letter

  This destroys the grammatical sense of the speech.

  In addition, not only was the pause omitted between "you have <number> lives left" and "the letters you have found are", but the words 'left' and 'the' were run together so tightly that the 'th' sound at the beginning of 'the' was all but lost.

* The placing of the pauses and the presence of redundant speech resulted in each of the important items of information - the recitation of the letters and the number of lives - being followed immediately by distracting material.

* The speech output lacked the natural rise and fall in pitch and volume which normally helps listeners to disentangle the components of a sentence.

* The Hal/Apollo combination reproduced the word 'lives' so that it rhymed with 'gives' rather than 'hives'. There were no other serious mis-pronunciations, but a few words were slightly garbled.

3.2 Method

The study was carried out at the RNIB Vocational College at Loughborough, UK. The college caters only for visually disabled students but runs many of its courses in conjunction with the local Technical College. This made it possible to select a group of blind and sighted subjects who were closely matched for age, educational background and prior experience of computers. The group eventually chosen comprised 10 blind and 7 sighted subjects.

The subjects sat at a table in front of a keyboard and the speech-synthesizer, with the computer screen present but hidden from their view. A video camera recorded the actions of the subjects and experimenter while a second camera recorded the details of the game as displayed on the computer screen. Thus a substantial amount of information was recorded for later analysis.

The subjects were divided into two groups:

> **Group One** consisted of four blind and four sighted subjects who undertook the task in pairs. Each pair consisted of one blind subject and one sighted subject, and they were instructed to cooperate in order to complete the task. The intention was that, free from intervention by the experimenter, they would discuss the game openly. Thus their thoughts, their feelings about the interface and their understanding of the system state at various stages of the process would be reflected in their discussions and recorded for later analysis.

Group Two consisted of six blind and three sighted subjects who undertook the task individually. The experimenter questioned each of them about their understanding of the system state after each letter had been entered, allowing their understanding of the system at each stage to be recorded. The intention was that this would yield less wide-ranging but more consistent information than might be revealed through the discussions of the cooperating pairs.

Each subject or pair of subjects was informed of the purpose of the experiments, after which they received a brief demonstration of the system and were told the rules of the game. They were then asked to play the game as many times as they could within a period of 20 minutes. When they had finished playing they were asked a number of questions:

First, subjects were asked if they felt that alternative or additional pauses in the speech output would have made the task easier, and if so, where the pauses should have been placed. A tape recording of a typical speech segment from the program was played and subjects were given a switch which allowed them to pause the recording whenever they wished. They were invited to listen to the recording as many times as they wished and to try the effect of different pause positions before making a final choice.

Subjects were asked whether they considered all the spoken text to have been useful or if some of it could have been usefully discarded. If they felt that some of the speech was redundant, they were asked to indicate which items they felt were useful and which should have been left out.

Subjects were asked if they found any of the speech items harder to remember as a result of their position within an utterance.

Finally, subjects were asked whether they felt that any of the speech might usefully have been replaced with non-speech sounds.

4. RESULTS

The subjects' responses to the questions asked at the end of the experiments are summarised in Table 1. It can be seen that there is a sharp difference of opinion between the blind and sighted subjects: the overwhelming majority of the sighted subjects were unhappy with all aspects of the interface whereas only around half the blind subjects expressed dissatisfaction.

Item	Blind Ss	Sighted Ss
Wanted less Speech	55%	100%
Preferred a different order	55%	100%
Wanted more pauses	40%	80%
Wanted non-speech sounds	0%	0%

Table 1. Summary of subjects' agreement with suggestions on how to improve the interface.

The blind subjects generally described the interface as difficult to use, but most said it was no worse than others they had worked with and all said that they had enjoyed playing the game. Many asked for copies of the software. All the blind subjects, both those working on their own and those working in pairs, were happy to continue playing the game for the whole of the allotted time.

By comparison, the sighted subjects found the interface extremely difficult to use and reported less enjoyment of the game. The sighted subjects who worked in pairs with blind subjects expressed less dissatisfaction than did the sighted subjects who worked alone. None of the sighted subjects who worked alone said that they had enjoyed the game, and two out of the three abandoned the task before the end of the allotted time.

4.1 Errors in the recall of information

Both blind and sighted subjects had difficulty remembering the sequence of letters and spaces. Only one (blind) subject clearly recalled the sequence correctly at all times. Three other blind subjects and one sighted subject made no audible errors, but since they worked in pairs their recollection of the sequence was not checked at every stage of the game. The remaining six blind and six sighted subjects all made errors on at least one occasion when attempting to recall a sequence of letters and spaces, and most made many errors.

Subjects also had difficulties recalling how many lives they had left. On a few occasions, subjects confused a number which formed part of the sequence (e.g., 'three spaces') with the number of lives remaining. More commonly they simply forgot how many lives remained. On one occasion, a member of a mixed pair asked the other how many lives remained, and the other replied "I don't know, I was too busy concentrating on the letters and spaces". Similar exchanges took place on a number of other occasions.

Subjects also had considerable difficulty remembering what letters they had already entered. Most of the subjects guessed at letters more than once, and a considerable amount of the conversation between the mixed pairs centred upon whether a particular letter had been tried or not. Subjects frequently guessed at letters they had tried only a few attempts earlier, and on two occasions the same letter was entered twice in succession. Subjects also forgot letters even when they had been successfully identified as part of the word.

4.2 Quantity of speech

As Table 1 shows, all the sighted subjects and over half the blind subjects said they would have preferred less speech. However, while the blind subjects claimed to be less troubled by the quantity of speech when asked directly, it was observed that four of them tried to mute sections of speech on numerous occasions, even though they had been asked not to use the mute facility on the screen reader.

When asked which parts of the speech could usefully have been removed, most subjects were unable to make any suggestions, even though many of them had previously stated that there was too much speech. The only phrase which was identified by a significant number of subjects was the instruction "Please enter a letter". Most felt that some audible prompt was needed at this point but that this phrase was unnecessarily long and distracting.

The only other phrase which was mentioned with any frequency was the first line, "There are <number> occurrences of the letter <letter>". Several subjects commented that the word 'occurrences' was unnecessarily long and awkward and suggested that the phrase should be shortened to "There are <number> <letter>", e.g., "There are 2 As". Others suggested shortening it still further to just "<number> <letter>", e.g., "2 As". It was also suggested that the word 'colon' need not have been spoken out.

A number of the subjects spoke (either to their partner or, in the case of those working alone, to themselves) over parts of the speech, which suggests that they did not consider it worth listening to. In most of these cases, subjects listened only until the number in the first line of text was spoken out ('there are <number> ...'). They then talked over the word 'occurrences' and continued to talk through the phrase 'the letters you have found are colon', falling silent again to hear the sequence of letters and spaces spoken out. They frequently talked through the phrase 'Please enter a letter' and sometimes through the announcement of the number of remaining lives, although on occasion it appears that the latter was unintentional. The words 'there are <number>' and the sequence of letters and spaces were the only items of the speech which were listened to on every occasion.

4.3 Ordering of the information

A majority of subjects felt that the interface would have been easier to use had the information been presented in a different order, but there was very little agreement as to what a better order might be. The only suggestion made with any frequency was to omit everything except the number of lives left and the sequence of letters and spaces, and to present those two items of information in that order.

During discussion, some also said they felt that it was important to receive an indication of success or failure each time they entered a letter, and that this information should precede everything else. They wanted to know first of all whether their guess had been successful, and if so, what the effect on the sequence of letters and spaces was. Being told how many times the new letter occurred *before* hearing the changed sequence of letters and spaces in some way made it easier to absorb the sequence when it came. Thus the phrase "There are x occurrences of the letter y" may have been appropriately placed even if it was unnecessarily long.

Several subjects had strong views on the placing of the phrase "Please enter a letter". A number said that it was distracting when placed after the sequence of letters and spaces and should be placed elsewhere. One subject said that she was conscious of having to 'filter out' this phrase in order to concentrate on the structure of the word. Another commented that he always felt 'rushed' when listening to the last few letters and spaces, perhaps because he was conscious of the approaching distraction. One subject said that she did not find this particular phrase a problem, but felt that if it had been replaced with one that contained important information (for example, the number of lives left) it would have hindered her recollection of the sequence of letters and spaces.

A number of subjects said that the greatest problem with the interface was not the quantity or ordering of the speech but the fact that it was fixed. They felt that different information was important at different points in the game, and that ideally the presentation should vary to reflect this. For example, some subjects felt that it was not necessary to know how many lives were left if the number had not changed, so this piece of information should only be presented when a life was lost. One subject felt that the number of lives left was irrelevant early on in the game and only became important when the number became very small.

A number of subjects felt that the information should be available 'on demand', either as a replacement for a fixed-order dialogue or as a complement to it. The most favoured suggestion was that 'hot keys' should be added which would allow them to hear the sequence of letters and spaces or the number of lives left at any time they chose. Opinion varied as to what information should then be provided when a letter was entered. Some felt that the sequence of letters and spaces should be spoken out, others wanted an indication of success or failure followed by the sequence, and others wanted to know only whether the guess had been successful or not, leaving them free to decide what other information to listen to.

4.4 Use of pauses

A number of the subjects agreed readily with the suggestion that more pauses would have made the speech easier to comprehend. Others said they felt it would have made little or no difference, but some of these changed their minds when pauses were mentioned indirectly during later questions. One blind subject commented initially that the speech was "... a bit fast" but later decided that retaining the same reading speed and introducing pauses would be more helpful than simply slowing the speech down. The figures given in Table 1 represent all those who eventually agreed with the proposition that more pauses would have made the speech easier to comprehend.

There was general agreement among those who wanted more pauses as to where they should be placed. Almost all of them felt that pauses should be introduced to divide the speech into four distinct statements, viz:

> *There are <number> occurrences of the letter <letter>*
>
> *You have <number> lives left*
>
> *The letters you have found are: <letter 1> <letter 2> ... <letter n>*
>
> *Please enter a letter*

There was some disagreement, however, regarding the pause after the colon in the third line. Some felt it was useful while others found it awkward and unhelpful. Some of those who disliked this pause suggested that it might become tolerable or even helpful when balanced by a longer pause at the end of the line.

4.5 Quality of the speech

Only one blind subject commented upon the quality of the speech, but all three of the sighted subjects who worked alone were unhappy with it. One thought that the speech was all but incomprehensible, and said it had felt as though he was listening to a foreign language. Another commented upon the lack of flow in the speech and said it was like listening to someone speak without ever drawing breath. The third sighted subject said he had found the pronunciation poor and this had made the task more difficult. The blind subject who commented upon the speech compared the Hal/Apollo combination with some of the newer systems he had tried, and said that in some respects he had found the speech a little garbled.

4.6 Use of Non-Speech Sounds

As Table I shows, when asked directly none of the subjects felt that substituting non-speech sounds for any of the speech would have made the interface easier to use. However, several recommended the use of non-speech sound when suggesting ways of improving the interface. One subject suggested that success or failure should be indicated by different musical tones, or perhaps by a tone for success and silence for failure. Other subjects also suggested the use a tone in place of "Please enter a letter".

5. DISCUSSION

All the subjects clearly had problems processing the information presented through the speech. The fact that more blind subjects than sighted subjects correctly recalled the sequence at each stage suggests that the blind subjects are better at this task, although the nature of the data collected makes it difficult to put a figure on the difference.

The subjects' comments on the use of pauses, the quantity of information and the ordering of the information are broadly in line with what might be expected from the literature reviewed earlier. Pauses were considered helpful, and there was general agreement as to where they should be placed. The quantity of information was too great, and most subjects would have preferred an interface which reproduced only the bare minimum of essential speech. The ordering of the information was poor, and most subjects would have preferred a different ordering.

It is more difficult to identify any clear consensus as to what was essential and what was not and how the ordering of the information might have been improved. Clearly the sequence of letters and spaces was regarded as essential, and there was also some agreement that the user should be told whether a guess had been successful or not. All other information, including the number of lives, was deemed either irrelevant or necessary on some occasions only. There was also clear agreement that all statements should be kept as short as possible.

As to the ordering of the information, the responses suggest a preference for shorter dialogues made available in a more flexible manner. Specific suggestions included using different, shorter dialogues depending on the success or failure of a key entry, and making some information available on 'hot keys'. There was also a suggestion that complex information was easier to digest if the listener were suitably primed to receive it. Thus in this case, the subjects felt that the sequences of letters and spaces would have been more difficult to recall had they not previously been told whether the guessed letter was present or not. There was also a clear recognition that extraneous speech placed after important material was distracting, and should be avoided wherever possible.

Finally, there does seem to be some recognition that non-speech sounds might be more useful than speech in some cases. Although none of the subjects wanted more non-speech sound when asked directly, their suggested improvements clearly recognise the value of non-speech sound over speech in certain situations. The apparent advantage non-speech sounds were seen to have in such cases was that they could convey simple information in a quicker and less distracting way than might be possible through speech.

6. CONCLUSIONS

The following conclusions can be drawn from this study:

* Too much spoken material presents problems for listeners. If such problems are to be avoided in the design of practical interfaces, it is essential that ways are found of gauging the complexity of a piece of dialogue in ways which relate to the likelihood of its being held in memory without corruption.

* Unnecessary speech poses problems by impeding the retention of important information. This is particularly true of speech which follows important information.

* The placing and length of pauses needs to be considered carefully when preparing text to be reproduced through synthetic speech.

* Information is easier to recall when it is presented as a series of short, terse dialogues whose content reflects the importance of the information according to the current system state. The long dialogue used in the game, which presented the same information regardless of whether a guess had been successful or not, was difficult to use for this reason

* Complex items of information are easier to process when preceded by a brief speech saying what they contain. With the sequence of letters and spaces in the game, for example, it was helpful to know what, if anything, had changed since last hearing before hearing it again.

* Providing control over what is heard and when makes an interface easier to use. Short dialogues, the ability to mute dialogues, and the ability to obtain specific speech output quickly would have improved the usability of the interface described here.

* Blind subjects appear to be better at processing spoken information than sighted subjects. It is not clear whether this is because they are generally more adept at memorization tasks or because they have developed strategies to cope with the problems presented by speech interfaces.

In summary, the results of this study suggest that most of the research findings presented earlier are of importance in the design of a speech-based interface. However, it is clear that much more work needs to be done before we can arrive at a model or a set of guidelines which adequately accounts for all the important factors in the design of a speech-based interface.

7. REFERENCES

Abrams, K. & Bever, T.G. (1969) Syntactic Structure Modifies Attention during Speech Perception and Recognition, *Quarterly Journal of Experimental Psychology*, 21, 280-290

Ayres, T.J., Jonides, J., Reitman, J.S., Egan, J.C. & Howard, D.A. (1979) Differing Suffix Effects for the Same Physical Suffix, *Journal of Experimental Psychology: Human Learning and Memory*, 5, 315-321

Conrad, R. (1960) Very Brief Delay of Immediate Recall, *Quarterly Journal of Experimental Psychology*, 12, 45-47

Crowder R.G. & Morton, J. (1969) Precategorical Acoustic Storage (PAS), *Perception & Psychophysics*, 5, 365-373

Engle, R.W. (1974) The Modality Effect: Is Precategorical Audio Storage Responsible?, *Journal of Experimental Psychology*, 102, 824-829

Jarvella, R.J. (1971) Syntactic Processing of Connected Speech, *Journal of Verbal Learning & Verbal Behaviour*, 10, 409-416

Luce, P.A., Feustel, T.C. & Pisoni, D.B. (1983) Capacity Demands in Short-term memory for Synthetic and Natural Speech, *Human Factors*, 25 (1), 17-32

Murray, D.J. (1966) Vocalization at Presentation and Immediate Recall with Varying Recall Methods, *Quarterly Journal of Experimental Psychology*, 18, 9-18

Nakatani, L.H. & Schaffer, J. (1978) Hearing Words without Words: Prosodic Cues for Word perception, *Journal of the Acoustical Society of America*, 63, 234-244

Nusbaum, H.C. & Pisoni, D.B. (1985) Constraints on the Perception of Synthetic Speech generated by Rule, *Behaviour Research Methods, Instruments & Computers*, 17 (2), 235-242

Penney, C.G. (1979) Interactions of Suffix Effects with Suffix Delay and Recall Modality in Serial Recall, *Journal of Experimental Psychology*, 5, 507-521

Pisoni, D.B. (1981) Speeded Classification of Natural and Synthetic Speech in a Lexical Decision task, *Journal of the Acoustical Society of America*, 70, s98

Reich, S. (1980) Significance of Pauses for Speech Perception, *Journal of Psycholinguistic Research*, 9 (4), 379-389

Streeter, L. (1978) Acoustic Determinants of Phrase Boundary Perception, *Journal of the Acoustical Society of America*, 64, 1582-1592.

ACKNOWLEDGEMENTS

Thanks are due to the RNIB Vocational College, Loughborough, UK, for their assistance in carrying out this study.

This work was supported by the EPSRC under grant number 92567838.

The Hal screen reader and Apollo speech-synthesizer are products of Dolphin Systems for People with Disabilities, P.O. Box 83, Worcester, UK, WR3 8TU