Proceedings of The Institute of Acoustics

A VISUAL LETTER-IDENTIFICATION TASK WHICH CORRELATES WITH SPEECHREADING ABILITY.

Graham A. Day, J.S. Finnigan and Stuart Gatehouse

MRC Institute of Hearing Research (Scottish Section),
Glasgow Royal Infirmary, Glasgow, UK.

ABSTRACT

Correlations have been reported between speechreading scores and many visual attributes, including pattern recognition and latency of visual evoked potentials. Many such attributes must combine to determine an individual's ability as a speechreader, so any comprehensive assessment of speechreading aptitude should reflect at least some of these attributes. A fundamental model of speechreading has been used to design a test which aims to reflect the overall potential of an individual to perform speechreading tasks. The test involves speed of perception and pattern recognition and is contained in a microcomputer program; it is quick and easy to perform.
Results are reported from 18 severely hearing impaired subjects. This homogeneous group was not selected to maximise the variance in speechreading skills, but to reflect the constraints in a relevant target population. Nevertheless a significant correlation (-0.73) was found between the test score and speechreading as assessed with a video-based sentence test. After the removal of subjects' age, hearing level, visual acuity and I.Q. as co-variates, the correlation remains.

INTRODUCTION

The need to formulate an assessment of speechreading ability has been the aim of many workers with the deaf over the last century. The requirement is especially apparent in the young deaf person when speech is not fully developed and the benefit of tuition needs to be gauged. These individuals could not be assessed with the audiovisual techniques (e.g. Utley Sentence Test [1]) that are suitable in post-lingual adults. Early researchers [2,3] believed that speechreading was a skill which, by definition, could be tutored to enhance the abilities of an individual. However, it is often observed by teachers that good speechreaders are 'born and not made'.

Jeffers and Barley [4] detail a literature review encompassing the possible involvement of some psychophysical factors in speechreading. An early investigation by Simmons [5] reported a large number of factors with significant correlations with speechreading ability, as measured on the Utley Test. The significant factors are listed in Table 1. These results contradicted the reports of O'Neil and Davidson [6] who used the same speechreading test. They presented non-significant correlations between speechreading ability and intelligence, digit memory span and reading comprehension. However an individuals' ability to perform a concept-formation task correlated significantly with speechreading ($p<0.05$). Hardick and colleagues [7] derived a score for visual ability from a large number of factors, for example acuity, width of vision, blink rate, accommodation, in a group of normally hearing and seeing students. A significant correlation ($p<0.05$) was found which they attributed to deviations from perfect visual acuity, not necessarily requiring correction.

Table 1. Results reported by Simmons [5]

Factor	Test	Corr. Coefficient
Hearing	Duration of loss	0.43 *
Reading	Sentence meaning	0.44 *
	Key words	0.58 +
Intelligence	Picture arrangement	0.48 *
	Digit symbol	0.50 +
	Block design	0.43 *
Synthetic Ability	Fragmentary sentences	0.40 *
Visual memory	Object picture span	0.57 +

+ Significant at 1% level
* Significant at 5% level

Erber [8] simulated varying acuity in normal subjects and assessed speechreading with a sentence identification task, similar to the Utley Test. He reported that speechreading scores have a non-linear relationship with visual acuity, there being a steep decrease in performance with vision-alone speechreading. However audiovisual speechreading scores had a more linear relationship with visual acuity. Recently, Ronnberg [9] measured audiovisual memory performance. His subjects were shown ten pictures of objects and told what they were. Immediately afterwards, the subject was required to re-identify the pictures. The difference in memory performance between the severely hearing impaired individuals and the expected performance measured in normals, correlated significantly with speechreading ability ($p<0.01$). He concluded that those individuals with severe hearing losses developed advanced short-term memory processing abilities.
Shepherd and colleagues [10] investigated visual evoked-response and reported a high correlation (-0.9) between the latency of response and speechreading ability. Samur [11] reported the correlation at a much lower level of significance (-0.58) but found that a factor calculated from various components of the response, reflected speechreading ability at a higher level of significance (0.84). Shepherd re-investigated [12] and reported correlations of -0.82 -0.66 and -0.86 over three repeats. It is probable that the visual evoked potential measures an aspect of the physiological information-processing capacity of the nervous system, perhaps relating to perceptual speed. In their text, Jeffers and Barley [4] conclude that good speechreaders have the ability to focus rapidly, perceive fine detail at a rapid rate, maintain sharp focusing over relatively long periods of time, have good memory, interest in detail and good peripheral vision to see all speechreading detail. The observation that good speechreaders do not constantly need to look unhindered at lip-movement in a one-to-one conversation suggests that peripheral vision maybe an important factor. The eye can only sharp focus one half of a degree so that at 1m., apart from fast involuntary eye movements, the 'window' of vision is about 1cm. diameter. Very little detail is seen outside a diameter of 3.5cm. The ring between these regions is best described as the semi-periphery. The ability to see fine detail therefore seems of unlikely importance. Visual memory and pattern perception are central phenomena, therefore the pre-mentioned

literature point towards the importance of central factors in speechreading perception.

METHOD

Our test is termed the Adaptive Visual Attribute Test (ADAPVAT). It aims to test those information-processing factors in vision, previously thought to distinguish poor from good speechreaders, using brief visual stimuli on a TV monitor. For the visual message to occur in that region of vision between sharp focus and no detail, the stimulus is presented at a random position along a ring of radius 17cm., about the centre of the monitor. A variable stimulus duration time ensured that aspects of neural conduction in central processing is stressed. By displaying a four-choice confusion of the visual message on the monitor and asking the individual to compare and select the correct item on a four-choice response box, the effects of long-term memory and stimulus familiarity are reduced. Short stimulus duration prevent eye movement prior to pattern recognition and stress speed of perception. Therefore the patterns needed to be easily recognisable, so alphabetic characters are used. The sets of confusions are:-

A	H	M	N
C	G	O	Q
K	S	X	Z
B	D	P	R

Prior to each stimulus, a small white square is displayed in the centre of the monitor for fixation. The four choice confusion is displayed at the bottom of the monitor throughout each stimulus presentation. After the presentation the four confusion letters have identifying numerals displayed beneath each. The individual is told that "the white square will disappear and one second later, a letter will flash somewhere on the screen. The letter will be one of the four shown. Press the button which corresponds to the letter you saw. If you are uncertain, press that button which corresponds to a letter similar to that which you saw."

After a practice with long stimulus durations (330ms), the test proceeds with the stimulus duration varying according to the individuals' performance. Higher performance leads to increased stimulus difficulty (shorter presentation) and vice versa. The up-down procedure uses a 70.7% correct criterion (Levitt, 1971). Two correct resposes lead to increased difficulty and one incorrect response results in decreased difficulty. The step size is halved from 80ms at the start, to 40, 20 and 10ms at reversals 1, 4 and 7 respectively. Before changes in step size are implemented, the stimulus duration reverts to the value of the mean of the last two reversals. The median of the last five reversals, when the step size is 10ms., is taken as the threshold. For such a procedure to be repeatable it is necessary that each of the four stimuli in the confusion have monotonic psychometric response functions. Each set of confusions should also lead to the stimulus items being of equal difficulty. Correcting offset values for each stimulus were obtained from the psychometric response function of 10 normal, student volunteers, see Figure 1. This involved the subjects responding to 10 repetitions of stimuli consisting of 24 random letters, at five random stimulus durations of 10, 60, 110, 160 and 210ms. The offset factor was determined from the difference between the stimulus duration necessary to predict 70.7% correct responses and a median value, taken as 150ms. for each letter (Table 2). Many functions so obtained were unsuitable (e.g. non-monotonic). If more than two stimuli of a confusion were unsuitable, the corresponding confusion was dropped from the procedure. In confusions where one or two stimuli are unsuitable, they are presented but not evaluated (dummy stimuli) in order to maintain the structure of that confusion.

Figure 1. The psychometric functions for each stimulus in the confusion set : C G O Q.

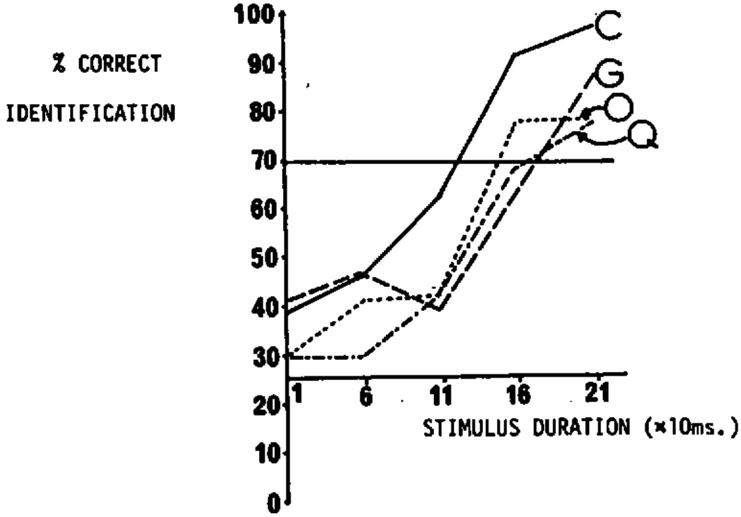


Table 2. Offset values determined for the confusions used.

M	-30ms.	H	20ms.	N	0ms.	A	unused
B	-20ms.	D	-30ms.	P	unused	R	unused
C	-30ms.	G	20ms.	O	-10ms.	Q	10ms.
K	-90ms.	X	unused	S	-20ms.	Z	-40ms.

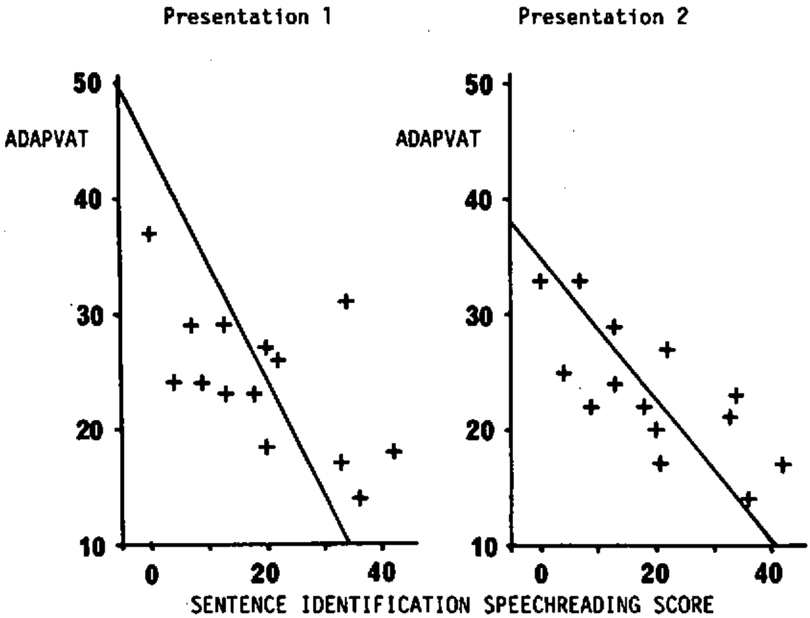
The measure taken of speechreading ability was obtained from a visual presentation of sentence lists, recorded on video. (BKB short sensible sentences of restricted vocabulary and banal content, scored from keywords [13]). All individuals tested were made familiar with the type of sentences used. All lists were assessed for ease of speechreading using normal student volunteers. Weighting factors were calculated and used to normalize the lists. Even though the speakers are different, the factors obtained were remarkably similar to those calculated by Rosen and collegues [14]. The speechreading score was obtained from one list presentation, the order of lists being rotated between subjects. Correlation of the adjusted scores between lists is high; a correlation coefficient of 0.87 was obtained for the mean scores of two different lists.

ADAPVAT was administered to 18 individuals with mean hearing levels of 76dB HL (four frequency average in the better ear). The mean age was 63 with a range from 40 to 77 years. The average time of hearing aid use was 21 years (S.D. 8.7 years). Also performed at the same test visit as two consecutive ADAPVAT and a speechreading assessment test were the Snellen and Arden visual acuity tests and a Raven Coloured Progressive Matrices Test [15]. Aswell as being a test of visual acuity, the Arden Test assesses contrast sensitivity. It is a more relevant assessment of peripheral visual resolution than the Snellen Test in individuals with vision better than 20/30.

RESULTS

Figure 2 shows plots of the median stimulus duration necessary to produced 70.7% correct identification on ADAPVAT versus speechreading assessment score, for both presentations of the test. Corresponding correlation coefficients are -0.55 ($p < 0.02$) and -0.72 ($p < 0.01$), respectively. The correlation between presentations is 0.78 ($p < 0.01$).

Figure 2. ADAPVAT versus sentence-identification speechreading score.



To investigate how the other measured variables influence speechreading in relation to the contribution of ADAPVAT score, their partial correlations with speechreading were first determined. A step-wise regression was performed inserting the most important of these variables into the regression sequentially prior to ADAPVAT score. The remaining correlations between speechreading and the ADAPVAT scores were then deduced (Table 3).

Table 3. Ordered step-wise regression of speechreading with various variables.

variable	PRESENTATION 1		PRESENTATION 2	
	F to delete	final correlation	F to delete	final correlation
Arden	2.85	-0.54	0.25	0.24
age	0.63	-0.29	0.45	-0.32
I.Q.	0.35	-0.21	0.05	0.10
HL	0.9	-0.34	3.32	-0.67
Snellen	0.0	0.0	0.32	0.27
ADAPVAT	2.0	-0.47	7.13	-0.80 *

* significant at $p<0.025$

Partial correlations between ADAPVAT (Presentation 2) and speechreading when the effect of each of the above factors is removed are listed:-

-	age	I.Q.	Arden	Snellen	HL
-0.73 ($p<0.01$)	-0.63 ($p<0.02$)	-0.73 ($p<0.01$)	-0.61 ($p<0.05$)	-0.65 ($p<0.02$)	-0.73 ($p<0.01$)

Age and visual acuity (resolution and contrast sensitivity) have an effect on Presentation 2 correlations, but in each case it is small.

CONCLUSIONS

ADAPVAT correlates significantly with speechreading as measured using a sentence identification task. Visual acuity (resolution and contrast sensitivity) and age as partial variables have a slight contribution in this relationship, contrast sensitivity having the greatest effect. However, the correlation between ADAPVAT and speechreading is still highly significant when the influence of age, I.Q., visual acuity and hearing level is removed.

The mean hearing level of individuals used in this investigation was 76dB HL. Most had used hearing aids for many years (mean time = 21 years: minimum = 8 years). Other studies show that speechreading ability correlates with duration of hearing loss, so that these individuals should have approached their optimal speechreading ability. It is therefore hypothesized that ADAPVAT reflects an individuals' potential for speechreading, not just their ability at the time of testing. Hence ADAPVAT measures those perceptual attributes which influence speechreading but does so at a level low enough not to be greatly influenced by learning. Validation of this hypothesis would involve testing individuals before and after speechreading tuition.

In summary, ADAPVAT performs well in predicting the ability of individuals to speechread and has considerable appeal on the grounds of shortness of test and repeatability.

REFERENCES

[1] J. Utley, 'A test of lipreading ability', J. Speech Dis., Vol. 11, 109-116, (1946).

[2] H.D. Kitson, 'Psychological tests for lip-reading ability', The Volta Review, Vol. 17, 471-476, (1915).

[3] E.B. Nitchie, 'Tests for determining skill in lip-reading', The Volta Review, Vol. 19, 222-223, (1917).

[4] J. Jeffers and M. Barley, 'Speechreading', Pub: Thomas, Springfield, Illinois, (1971).

[5] A.A. Simmons, 'Factors relating to lipreading', J. Speech & Hearing Res., Vol. 2, 340-352, (1957).

[6] J.J. O'Neil and J.L. Davidson, 'Relationship between lipreading ability and five psychological factors', J. Speech & Hearing Disord., Vol. 21, no.4, 478-481, (1956).

[7] E.J. Hardick, H.J. Oyer & P.E.Irion, 'Lipreading performance as related to measurements of vision', J. Speech & Hearing Res., Vol. 13, 92-100, (1970).

[8] N.P. Erber, 'Auditory-visual perception of speech with reduced optical clarity', J. Speech & Hearing Res., Vol. 22, 212-223, (1979).

[9] J. Ronnberg, G. Ohngren & L.G. Nilsson, 'Hearing deficiency, speechreading and memory functions', Scand. Audiol., Vol. 11, 261-268, (1982).

[10] D.C. Shepherd, R.W. DeLavergne, F.X. Frueh & C. Clobridge, 'Visual-Neural Correlate of speechreading ability in normal-hearing adults', J. Speech & Hearing Res., Vol. 20, 752-765, (1977).

[11] V.J. Samur and D.G. Sims, 'Visual evoked-response correlates of speechreading performance in normal hearing adults...', J. Speech & Hearing Res., Vol. 26, 2-9, March (1983).

[12] D.C. Shepherd, 'Visual-neural correlate of speechreading ability in normal hearing adults', J. Speech & Hearing Res., Vol. 25, 521-527, Dec (1982).

[13] J. Bench and J.Bamford, 'Speech-hearing tests and the spoken language of hearing-impaired children', Pub: Academic Press, (1979).

[14] S.M. Rosen and T. Corcoran, 'A video-recorded test of lipreading for British english', Brit. J. Audiol., Vol. 16, 245-254, (1982).

[15] J.C. Raven, 'The Coloured Progressive Matrices', Pub: Lewis & Co. Ltd., London (1956).