SPEECH UNDERSTANDING WITH LOW FREQUENCY HEARING A CASE STUDY

Geoff Plant
National Acoustic Laboratories

ABSTRACT - A subject with a long standing high frequency loss was tested using a variety of speech materials. The results indicate much useful information may be gained from even limited amounts of low frequency hearing.

INTRODUCTION

A number of studies (Grant, Ardell, Kuhl and Sparks, 1985; Plant, Macrae, Dillon and Pentecost 1984a; Risberg, 1974; Risberg and Agelfors 1978; Risberg and Lubker, 1978; Rosen, Fourcin and Moore, 1981) have highlighted the great potential of auditorily presented low frequency speech information as a supplement to lipreading. The results of these studies have prompted the development of aids for hearing-impaired persons which present fundamental frequency (F_0) cues either auditorily (Rosen and Fourcin, 1983; Walliker, Rosen and Fourcin, 1986) or tactually (Boothroyd and Hnath 1986; Plant, 1987; Plant, Macrae, Dillon and Pentecost, 1984b, Plant and Risberg, 1983; Rothenberg and Molitor, 1979). The results obtained with these aids have been encouraging but are based on limited training and exposure times by experimental subjects. As a result it is difficult to determine the true potential of the experimental devices. The subject of this paper offers some insights into the long-term potential of presenting low frequency speech information as a supplement to lipreading.

SUBJECT

The subject of this study is a 50 year old female with a "ski-slope" sensori-neural loss of unknown origin. The subject reports that she first noticed a slight hearing loss when she was in her mid 20s. Audiometric testing at the commencement of the study revealed hearing thresholds within normal limits at 125 and 250 Hz and then a rapid drop to severe-profound levels at 500 Hz and above (Figure 1.)

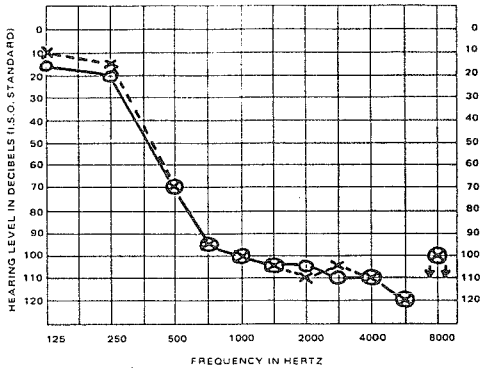

Figure 1. The subject's audiometric thresholds.

TEST MATERIALS AND PRESENTATION

The test materials used to evaluate the subject's performance were:

1. Consonant identification. The consonants /p,t,k,b,d,g,f,v,s,z,m,n/ were presented live-voice in an [aCa] frame. Lists were prepared presenting each consonant five times in a random order. Two lists were presented in both the auditory alone and the auditory-visual conditions by a male speaker. A female speaker presented one list auditory alone and one list auditory-visually.

2. Vowel identification. The vowels /i,I,ɛ,æ,a,ʌ,ɒ,ɔ,ʊ,ɜ,u/ were presented live-voice in a [hVd] frame. The lists used presented each vowel five times in a random order.

3. Sentence Identification. The materials used to evaluate sentence perception were: live-voice presentations of Kallikow, Stevens and Elliott's (1977) S.P.I.N. Lists, audio recordings of two 25 item lists of the high predictability sentences of the S.P.I.N. Test and video recordings of Tonisson's (1976) Australian adaptation of the C.I.D. Everyday Sentence Test, the Helen Test (Plant, Phillips and Tsembis, 1982) and the NAL Lipreading Test (Plant and Macrae, 1981).

4. Connected Discourse Tracking (C.D.T.) One 10 minute session of De Filippo and Scott's (1978) C.D.T. procedure was presented live-voice. Two live voice sessions were also presented using Plant's (1987) adaptation of C.D.T.
The presentation conditions for the test materials are presented in Table I

	Auditory alone	Visual alone	Auditory Visual
Consonants	√		√
Vowels	√		√
Sentences			
Helen	√	√	√
C.I.D. Sentences		√	√
S.P.I.N.	√		√
S.P.I.N.High Pred.	√		
NAL Lipreading		√	√
C.D.T.	√		

Table I. Presentation conditions for test materials.

Testing was conducted in a quiet non-sound-treated room. The auditory signal was set at the subject's preferred listening level. Live voice materials were presented using normal vocal effort. The subject did not use any form of amplification during testing.

RESULTS.

The subject's performance with the consonant and vowel materials showed a clear superiority for the auditory-visual condition. The responses for the male and female speakers were pooled and confusion matrices drawn up. The results of these feature analyses are presented in Tables II and III. These show that considerable information is available to the subject via audition alone. The consonant features voicing, sibilance, nasality and the stop/continuant contrast were all identified with a high degree of proficiency in the unisensory condition. Vowel length and height were also available to the subject via audition alone.

Table II. The subject's scores for the consonant identification tasks

	Auditory alone	Auditory-Visual
Overall	46.7%	94.4%
Voicing	100%	99.4%
Manner of Articulation		
Overall	81.1%	96.1%
Stops	90%	94.4%
Fricatives	23.5%	100%
Sibilants	93.5%	93.3%
Nasals	100%	100%
Place of articulation		
Overall	48.9%	97.2%
Bilabials	40%	100%
Labio-dentals	23.5%	100%
Alveolars	76%	94.7%
Velars	20%	96.7%

Table III. The subject's scores for the vowel identification tasks.

	Auditory alone	Auditory-visual
Overall	56.4%	78.2%
Vowel length	100%	91.8%
Vowel height		
Overall	84.5%	90%
High	95%	97.5%
Mid	72.5%	85%
Low	86.7%	86.7%
Vowel place		
Overall	61.8%	84.5%
Front	47.5%	97.5%
Mid	77.5%	72.5%
Back	60%	83.3%

The subject's sentence identification scores are presented in Table IV.

Table IV. The subject's performance for the sentence level tests.

	Visual alone	Auditory alone	Auditory-visual
Helen Test	75%	35%	100%
C.I.D. Sentences	30%	-	80%
S.P.I.N. Sentences	2%	14%	70%
S.P.I.N. High Predict.	-	24%	-
NAL Lipreading Test	54%	-	98%

These again show a clear superiority for the auditory-visual condition. The score obtained for the NAL Lipreading Test via vision alone (54%) indicates that although the subject is an above average lipreader her performance is not exceptional (score is at the 26th centile level). The contribution of audition to the subject's face-to-face communication skills is highlighted by the score obtained (98%) when the test was presented auditory-visually.
The subject's main tracking rate for the C.D.T. session was 24.5 words-per-minute. The mean score was obtained for the two sessions using the modified tracking task was 90.5% correct.

DISCUSSION

The results obtained in this study highlight the great potential of even limited auditory information for speech understanding. It should be emphasised that the subject has more than F_0 information available via audition. It would appear that she receives information from a speech band which is essentially low-pass-filtered at approximately 500 Hz. This provides much useful information on consonant voicing and manner of articulation and vowel length and height. When this information is combined with vision the subject's consonant recognition scores rise above 90% for all contrasts examined. Vowel recognition scores also rise significantly in the auditory-visual conditions although the overall correct score of 78.2% indicates that the lack of higher formant cues results in some vowel confusions.

The subject's scores auditory alone for both the sentence materials and the connected discourse task show the value of contextual cues in understanding speech. The subject's score of 14% for the S.P.I.N. sentences for example resulted from correct recognition of six high predictability items and one low predictability sentence. Similarly the Helen Test utilises a very predictable question format which gives cues to the nature of the key words in the sentences. The connected discourse tests allowed the subject to use skills which more closely approximate everyday communication. In these tasks the subject revealed an ability to understand connected discourse via audition alone with a high degree of proficiency, albeit at a reduced rate.

The lipreading-alone scores indicate that although the subject is a proficient lipreader, vision alone does not account for her high performance in the auditory-visual conditions. The subject's scores for all of the sentence tasks are significantly higher in the auditory-visual condition. The auditory information available to the subject obviously provides much information which is not available via lipreading alone.

Further research needs to be undertaken to determine whether the subject's performance can be improved by the use of speech recoding aids. A simple tactile cue to high frequency speech information may, for example, lead to improved speech understanding. More complex approaches such as frequency transposition should also be investigated.

REFERENCES

Boothroyd, A. and Hnath, T. (1986) "Lipreading with tactile supplements", J. of Rehab. Res. and Devel, 23, 139-146.

De Filippo, C.L. and Scott, B.L. (1978) "A method for training and evaluating the reception of ongoing speech" J. Acoust. Soc. Am. 63, 1186-1192.

Grant, K.W., Ardell, L.A., Kuhl, P.K. and Sparks, D.W. (1985) "The contribution of fundamental frequency, amplitude envelope and voicing duration cues to speech reading in normal hearing subjects." J.Acoust.Soc.Am. 77, 671-677.

Kalikow, D.N., Stevens, K.N. and Elliott, L.L. (1977) "Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J.A.S.A. 61, 1337-1351.

Plant, G. (1987) "A single-transducer vibrotactile aid to lipreading." Speech Communication, 6, 335-342.

Plant, G.L. and Macrae, J.H. (1981) "The NAL Lipreading Test. Development, standardisation and validation." Aust. J. Audiol. 3, 49-57.

Plant, G., Macrae, J., Dillon, H. and Pentecost, F. (1984a). "Lipreading with minimal auditory cues". Aust. J.Audiol. 6, 65-72.

Plant, G., Macrae, J., Dillon, H. and Pentecost, F. (1984b) "A single-channel vibrotactile aid to lipreading: Preliminary results with an experienced subject". Aust.J.Audiol. 6, 55-64.

Plant, G.L., Phillips, D. and Tsembis, J. (1982) "An auditory-visual speech test for the elderly hearing impaired". Aust. J.Audiol. 4, 62-68.

Plant, G. and Risberg, A. (1983). "The transmission of fundamental frequency variations via a single channel vibrotactile aid. STL-QPSR 2-3/1983, 61-84.

Risberg, A. (1974) "The importance of prosodic speech elements for the lipreader". Scand. Audiol. Suppl. 4., 153-164.

Risberg, A. and Agelfors, E. (1978). "On the identification of intonation contours by hearing-impaired listeners". STL-QPSR, 2-3, 51-61.

Risberg, A. and Lubker, J. (1978) "Prosody and speechreading" STL-QPSR, 4, 1-16.

Rosen, S. and Fourcin, A.J. (1983 "When less is more - further work". Speech, Hearing and Language. Work in Progress. University College, London. 3-27.

Rosen, S.M., Fourcin, A.J. and Moore, B.C.J. (1981). "Voice pitch as an aid to lipreading". Nature, 291, 150-152.

Rothenberg, M. and Molitor, R.D. (1979) "Encoding voice fundamental frequency into vibrotactile frequency". J.Acoust.Soc.Am. 66, 1029-1038.

Tonisson, W. (1976) "Australian standardisation of CID Everyday Sentence Test". Proceedings of the 2nd National Conference of the Audiological Society of Australia. Melbourne. A.S.A.

Walliker, J., Rosen, S. and Fourcin, A. (1986.) "Speech pattern prostheses for the profoundly and totally deaf". Speech, Hearing and Language. Work in Progress. University College, London . 223-236.