Speech pattern hearing aids for the profoundly hearing impaired: Speech perception and auditory abilities

Andrew Faulkner, Virginia Ball, and Stuart Rosen
Department of Phonetics and Linguistics, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, England

Brian C. J. Moore
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

Adrian Fourcin
Department of Phonetics and Linguistics, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, England

(Received 19 February 1990; revised 30 July 1991; accepted 3 December 1991)

A family of prototype speech pattern hearing aids for the profoundly hearing impaired has been compared to amplification. These aids are designed to extract acoustic speech patterns that convey essential phonetic contrasts, and to match this information to residual receptive abilities. In the first study, the presentation of voice fundamental frequency information from a wearable SiVo (sinusoidal voice) aid was compared to amplification in 11 profoundly deafened adults. Intonation reception was often better, and never worse, with fundamental frequency information. Four subjects scored more highly in audio–visual consonant identification with fundamental frequency information, five performed better with amplified speech, and two performed similarly under these two conditions. Five of the 11 subjects continued use of the SiVo aid after the tests were complete. A second study examined a laboratory prototype compound speech pattern aid, which encoded voice fundamental frequency, amplitude envelope, and the presence of voiceless excitation. In five profoundly deafened adults, performance was better in consonant identification when additional speech patterns were present than with fundamental frequency alone; the main advantage was derived from amplitude information. In both consonant identification and connected discourse tracking, performance with appropriately matched compound speech pattern signals was better than with amplified speech in three subjects, and similar to performance with amplified speech in the other two. In nine subjects, frequency discrimination, gap detection, and frequency selectivity were measured, and were compared to speech receptive abilities with both amplification and fundamental frequency presentation. The subjects who showed the greatest advantage from fundamental frequency presentation showed the greatest average hearing losses, and the least degree of frequency selectivity. Compound speech pattern aids appear to be more effective for some profoundly hearing-impaired listeners than conventional amplifying aids, and may be a valuable alternative to cochlear implants.

PACS numbers: 43.66.Ts, 43.71.An, 43.72.Ar

INTRODUCTION

There are many who suffer from such profound hearing impairment that they gain little or no benefit from conventional hearing aids even in supplementing lipreading. According to Thornton (1986), approximately 0.05% of the UK population have hearing losses of this degree, that is, an average hearing loss at 0.5, 1, 2, and 4 kHz of 100 dB or more. However, most of this population retain some measurable hearing generally at low frequencies, and an acoustic hearing aid, which made optimal use of such residual capacity could be an innocuous and economical alternative to a cochlear implant. A speech pattern-based approach to the optimal use of residual capacity has been advocated by Fourcin and co-workers (Fourcin, 1977; Fourcin *et al.*, 1977, 1979; Rosen and Fourcin, 1983; Fourcin, 1990), who have proposed the simplification of the complex speech signal to one or more basic auditory pattern elements related to essential phonetic contrasts; such patterns can be matched to the patient's auditory abilities by appropriate mappings of frequency and intensity.

The residual auditory abilities of this population are not thoroughly documented beyond audiometric measures, which typically show, as frequency increases, an increasing threshold and decreasing dynamic range. Since the impairment of frequency selectivity is strongly correlated with audiometric loss (Tyler, 1986), it is likely that this population will show little or no auditory frequency selectivity. Often they do, however, retain the temporal processing abilities required to extract important temporal patterns from speech. In particular, they remain sensitive to the presence of periodicity and changes of periodic rate in the voice funda-

 © 1992 Acoustical Society of America

mental frequency range, to silent gaps of 20 ms or more in duration (Rosen *et al.*, 1990), and to amplitude modulation of low-frequency carriers at rates between 5 and 50 Hz (Faulkner *et al.* 1990a).

One speech pattern of special importance for the lipreader is that of voice fundamental frequency, since it can convey both segmental voicing information and suprasegmental intonation, both of which complement the visual information available to the lipreader (Rosen *et al.*, 1981). For listeners who lack auditory frequency selectivity, the extraction of fundamental frequency information from speech is likely to be difficult because periodicity may be obscured by spectral complexity and by short-term changes in the magnitude and phase spectrum of the speech signal. If, however, speech were greatly simplified to a fixed intensity sinusoid following the voice fundamental frequency pattern, periodicity would be clearly represented.[1] A recoding of this sort also allows the use of transformations that preserve phonological patterning. In the transformation we refer to as MAPITCH (Fourcin *et al.*, 1984), a range of pitch mappings are possible. As used here, fundamental frequency is transposed by a simple subtractive constant so that it maps onto the region of the patient's best residual hearing, while relative frequency changes are at the same time exaggerated to compensate for impaired perception of frequency differences. Additional speech pattern elements can also be considered, with each pattern matched to the auditory abilities of the listener. For example, amplitude envelope and aperiodic stimulation could represent patterns related to the manner of articulation of consonants.

Existing data on the potential receptive advantages of speech pattern presentation are sparse and inconclusive. Rosen and Fourcin (1983) described results from one profoundly deafened subject who obtained little useful information from amplified speech, yet was considerably assisted in the perception of intonation and in audio–visual connected discourse tracking (CDT; DeFilippo and Scott, 1978) when the voice fundamental frequency pattern was presented as a frequency-modulated sinusoid. Grant (1987a,b) made similar comparisons in a group of five profoundly impaired adults, and also found that fundamental frequency information, especially with exaggeration of frequency change, provided similar or superior reception of stress and intonation to amplified speech. Grant (1987a) did not, however, find any advantage of fundamental frequency presentation compared to broadband or low-pass filtered amplified speech in audio–visual CDT. Finally, Rosen *et al.* (1987) compared intervocalic consonant identification in three listeners using lipreading aided by fundamental frequency and lipreading aided by amplified speech, and found little difference between these conditions. There is, then, evidence from only one listener of a receptive advantage for voice fundamental frequency presentation over amplified speech in audio–visual speech perception. It is not known whether such a simplified signal would provide significantly less information than amplified speech for the majority of the profoundly hearing impaired, or whether a more complex speech pattern signal might be more informative. Further, it is not known which audiometric indicators might identify potential users, nor what psychoacoustic abilities underlie their speech perceptual performance. The present study addresses these issues. It is concerned, in part, with results obtained using a wearable acoustic aid that provides voice fundamental frequency information, and also with the potential for extending the speech pattern approach by the incorporation of additional acoustic pattern elements. Because of the clinical nature of aspects of this study, it was not possible to carry out all tests on all of the subjects. An overview of the tests performed and the main summary results are tabulated in the Appendix.

I. THE SiVo AID

A. Functional description of the SiVo aid

A full description of the SiVo (sinusoidal voice) aid is given by Rosen *et al.* (1987). The peak-picking fundamental frequency extractor provides a period-by-period estimate of the speaker's larynx frequency (Howard and Fourcin, 1983; Howard, 1989). When voicing is detected, a sinusoid is synthesized whose period is equal to the estimate of the corresponding larynx period; we call this signal *Sx*. When MAPITCH is selected, the output is at a frequency either 50 or 80 Hz less than the estimated larynx frequency. The intensity of the output is independent of the intensity of the input speech signal, and set to the patient's comfortable listening level for the output frequency. The maximum output of the aid measured using an ear simulator (Bruel and Kjaer type 4157) is approximately 140 dB SPL at all frequencies between 31 and 500 Hz, falling to about 135 dB SPL at the maximum output frequency of 707 Hz. The fundamental frequency extraction process causes a delay of one larynx period; the total delay between the acoustic input and output is one larynx period plus a processing delay of 7 ms.

B. Response of the SiVo aid to speech contrasts

The segmental speech information provided by the SiVo aid can be illustrated by five intervocalic alveolar consonants (see Fig. 1). The voiceless consonants [s] and [t] both give similar outputs, with a silent interval of about 120 ms while voicing is absent. The nasal [n] is continuously voiced. The illustrated voiced fricative [z] shows a devoiced interval of about 70 ms, which is typical of our British English test materials, although some tokens from this speaker are voiced throughout. The illustrated [d] is fully voiced, and the SiVo aid responds throughout the closure; if, however, as is common, the plosive were devoiced, or the speech amplitude during the closure were very low, the voice fundamental frequency extractor would not respond during closure. Since the voice fundamental frequency is encoded, the SiVo aid also conveys information related to intonation and to important aspects of speaker identity (Abberton and Fourcin, 1978).

II. COMPARISON OF THE SiVo AID WITH AMPLIFICATION

A. Selection of patients

Two subjects, S1 and S2, were the first two patients to receive a SiVo aid. S1 had used a desk-top version of the SiVo aid for two years prior to this study, while S2 had used a

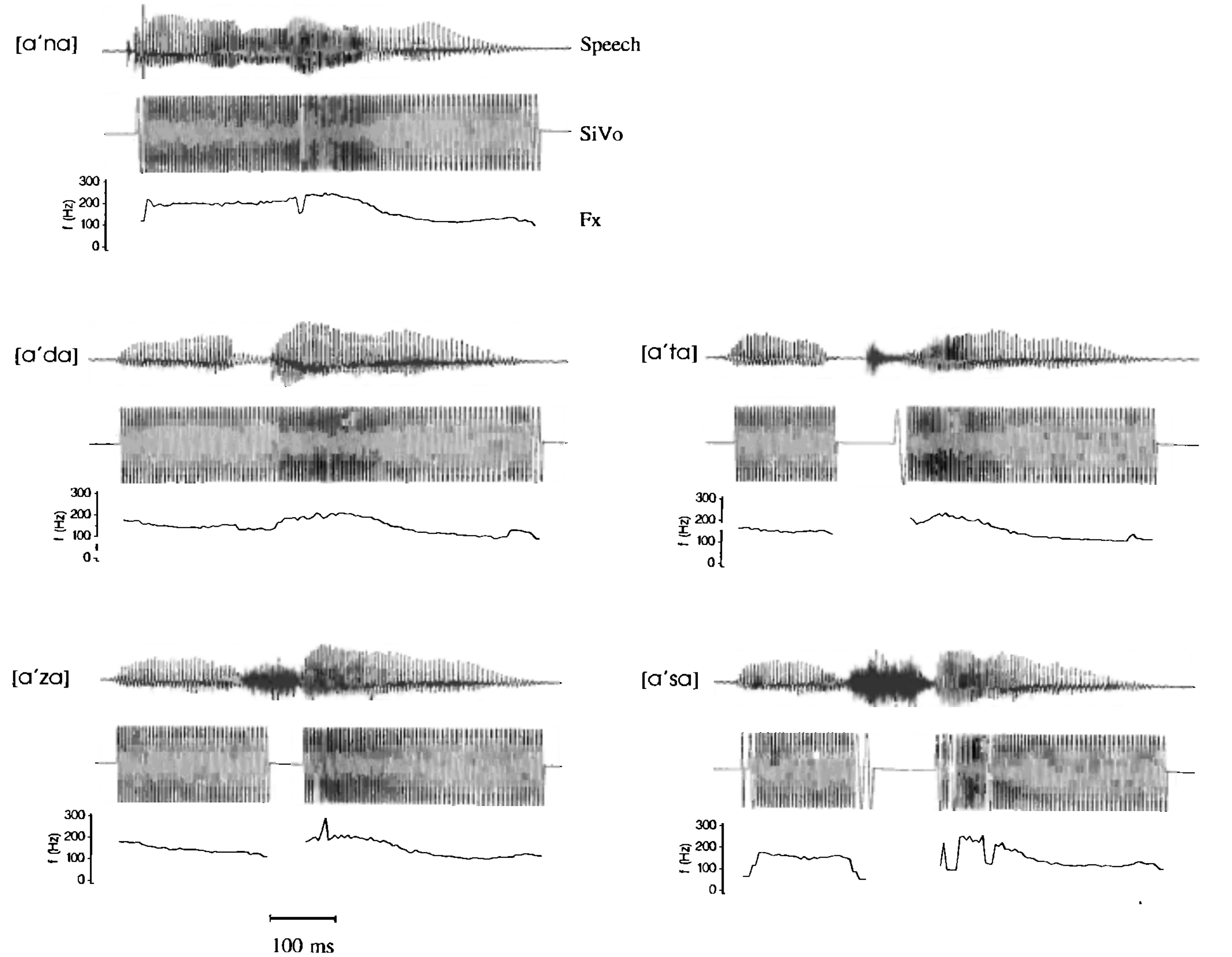

FIG. 1. Speech processing of the SiVo aid for the alveolar intervocalic consonants [ɑˈnɑ], [ɑˈdɑ], [ɑˈzɑ], [ɑˈtɑ], and [ɑˈsɑ], pronounced with the stress on the second syllable. For each consonant, the upper panel shows the speech pressure waveform, the central panel shows the output waveform from a SiVo aid, and the lower panel shows the fundamental frequency pattern that this output signal represents. The irregularities in the fundamental frequency pattern seen in the [ɑˈnɑ], [ɑˈzɑ], and [ɑˈsɑ] tokens represent occasions on which the peak-picking fundamental frequency extractor of the aids either missed one major waveform peak, or erroneously detected a minor peak. The degree of error shown here is typical of the performance of the aid in quiet conditions.

body-worn SiVo aid for about 6 months. Some aspects of these two subjects' speech perceptual and psychoacoustic performance were described by Rosen *et al.* (1987). The rationale for the selection of further patients was to cover a range of audiometric profiles so as to determine predictive criteria for the relative benefit of conventional aids and the SiVo aid. Eleven further subjects were selected using the following criteria: (1) an audiometric loss of 95 dB or more at 500 Hz; (2) a postlingually acquired hearing loss; (3) measurable residual hearing between 125 and 500 Hz; (4) The ability to distinguish a sinusoidal stimulus at 141 Hz from a 100–200 Hz or 71–282 Hz band of noise, at stimulus durations of 200 ms or less (Rosen *et al.*, 1987); and (5) dissatisfaction with conventional hearing aids.

One subject, S11, was included despite having a less profound hearing loss of about 80 dB at 500 Hz, since her auditory area was very restricted above this frequency (see Fig. 2), and she was unable to make use of conventional amplifying aids. All of the subjects had normal or effectively aided hearing during childhood, in contrast to the group studied by Grant (1987a), in whom the hearing loss was either congenital or appeared before two years of age. Apart from S2, the subjects had been referred to Guy's Hospital, London or University College Hospital, London, as potential candidates for cochlear implantation.[2] The age, cause of deafness, and hearing aid use of the subjects is given in Table I.[3]

The audiometric thresholds and discomfort levels of the 13 subjects are shown in Fig. 2. The data were obtained by

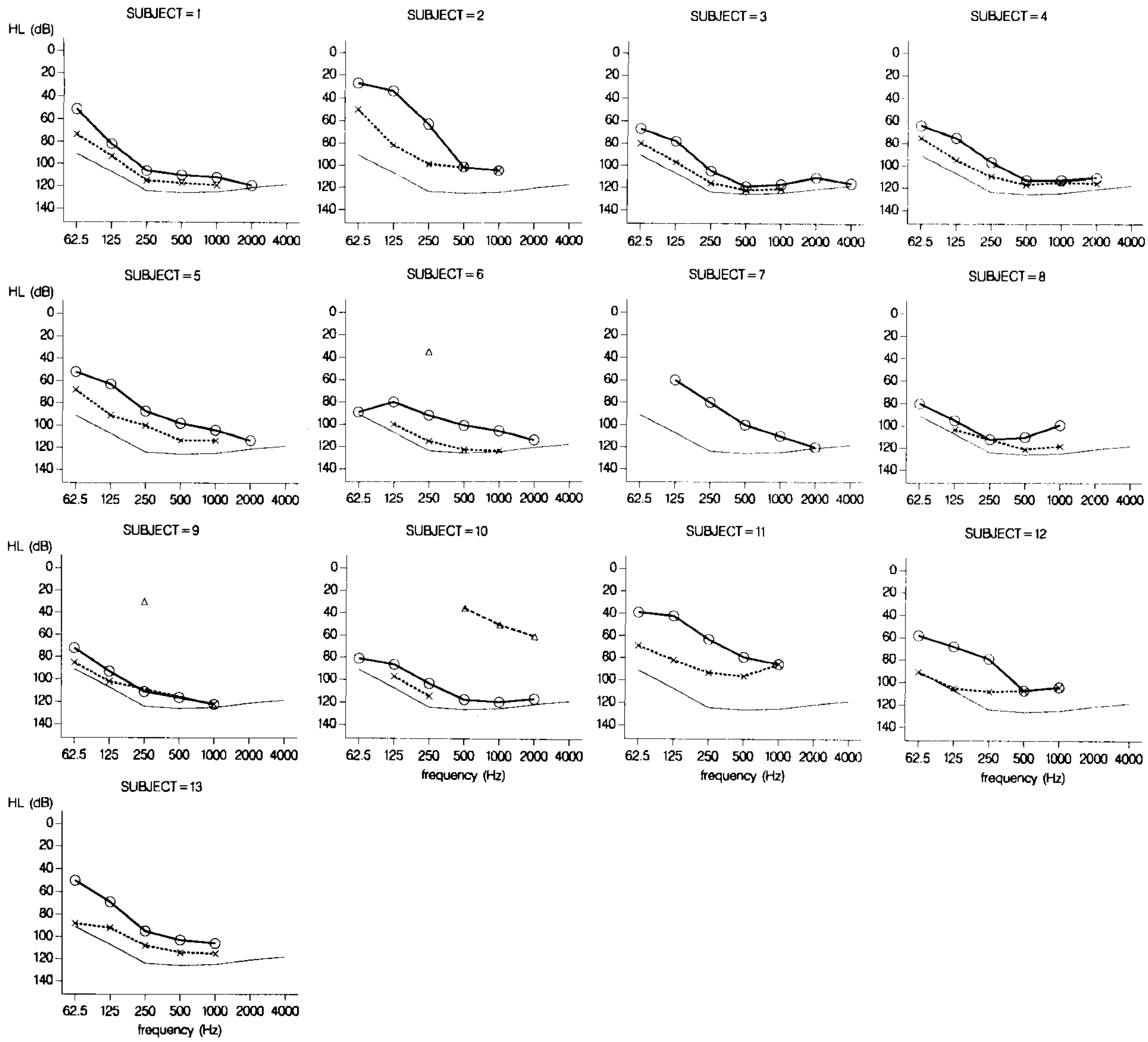

FIG. 2. Audiometric thresholds and discomfort levels. Open circles: audiometric threshold in dB HL (ISO) (manually determined using apparatus capable of at least 130 dB SPL between 125 and 2000 Hz). ×: discomfort level in dB HL; this represents the lowest level at which unpleasant sensations arose from sound or from feeling. Triangles: bone conduction threshold in dB HL (maximum levels testable were; 40 dB HL @ 250 Hz, 50 dB HL @ 500 Hz, 60 dB HL @ 1000–4000 Hz). The fine solid line represents the maximum output of the earphones before clipping. It was not possible to collect discomfort level measures from S7 because his illness limited the time available for testing.

hand, using a procedure similar to the standard audiometric threshold procedure, except that steps as small as 1 dB were used. Discomfort levels were taken to be the lowest levels at which either tactile or auditory discomfort was experienced. Table I includes the low-frequency average (LFA) threshold and dynamic range (125 and 250 Hz), the three-frequency average (500, 1000, and 2000 Hz: 3FA) and four-frequency average (500, 1000, 2000, and 4000 Hz: 4FA) threshold, and a 3FA dynamic range for each subject. Since these subjects typically had no measurable hearing at 4000 Hz, the 3FA loss is more informative than the 4FA loss, which is included for comparison with the population estimates of Thornton (1986). Only S6, S9, and S10 had measurable bone conduction thresholds. In these three subjects, there was therefore a strong indication of a conductive involvement. The presence of a mild-to-moderate conductive component in the remaining subjects could not be determined because the air-bone gap was unmeasurable. Otological examination suggested otosclerosis as a contributory factor in S9 and S10. These two subjects did not take part in the main trial, partly because of the conductive component to their hearing loss, and also because their extreme LFA loss prevented then from adequately hearing with either the SiVo aid or the reference aid.[4]

B. Fitting of the SiVo aid

The fitting procedure for the SiVo aid made use of the aid's built-in audiometer function (Rosen *et al.*, 1987). First, a sinusoidal signal at 125 Hz was set to the most com-

TABLE I. Audiometric data and aetiology of patients taking part. LFAHL: mean hearing loss at 125 and 250 Hz; LFADR: mean dynamic range at 125 and 250 Hz; 3FAHL: three-frequency average hearing loss at 500, 1000, and 2000 Hz; 4FAHL: four-frequency average hearing loss at 500, 1000, 2000, and 4000 Hz; 3FADR: average dynamic range at 500, 1000, and 2000 Hz. Average thresholds were calculated assuming a threshold of 130 dB HL where no threshold could be measured. Average dynamic range values include only frequencies where sound or discomfort was detected. Dynamic range data were not collected from S7.

S	Age	LFAHL	LFADR	3FAHL	4FAHL	3FADR	Aetiology
1	59	94	10	114	118	8	Head injury
2	48	49	42	112	117	0	Progressive/unknown
3	63	92	15	116	116	5	Progressive
4	56	86	17	112	117	4	Progressive/unknown
5	53	75	21	105	111	10	Progressive/possibly mumps
6	43	86	22	107	113	17	Progressive/unknown
7	60	70	...	110	115	...	Progressive/unknown
8	55	104	8	113	117	15	Unknown
9	50	102	9	123	125	4	Progressive/otosclerosis
10	58	95	11	117	121	8	Progressive/otosclerosis
11	64	53	35	98	106	9	Progressive
12	44	74	33	114	118	0	Pituitary tumor/Hypophysectomy
13	58	82	18	113	117	10	Progressive/unknown

fortable level; sinusoids were then presented at the half-octave-spaced frequencies between 31 and 707 Hz and adjusted firstly to a comfortable level, and, as far as possible, made similar in loudness to the 125-Hz tone.

C. Selection and fitting of reference aids

Six subjects (S3–S8) were provided in balanced sequence with the SiVo aid and a whole speech reference amplifying aid. The reference aid was chosen on the basis of high output, automatic gain control (agc), and a reasonable degree of adjustment of frequency response. The selected aid was a Danavox A/S body-worn model 107-2, modified by the manufacturer to extend its low-frequency response to ensure that speech information was delivered with sufficient intensity, and free from distortion, in the frequency region where the subjects had their most useful hearing. The free-field frequency responses of this and the other hearing aids used in the speech receptive tests described below are shown in Fig. 3. The maximum output of the modified 107-2 aid with a type PP earpiece at 125 Hz was 138 dB SPL; the largest harmonic distortion product was at least 20 dB below the level of the test signal. For the flatter response type W earpiece, the comparable maximum output was 131.5 dB. At output levels of 3 dB or more below these maxima, all distortion components from a sinusoidal input were at least 40 dB below the level of the test signal. The output of the reference aid was thus amplitude compressed speech with a bandwidth of about 4 kHz, as was that from the Phonak PP-C-L-A post-aural aid used by S1; the speech signal from the reference and the Phonak aids will be described as "whole speech." [5]

Conventional hearing aid fitting based on selective amplification was not considered appropriate since the aim was to provide useful information where hearing was least impaired. Fitting was guided by the first two authors' and the subjects' subjective appraisal of the information provided by the aid, with particular weight given to segmental voicing and intonation contrasts. Free-field-aided audiometry was performed to assist in the choice of earpiece and the setting of the agc. The aid was always set with the minimum agc required by the subject, in order to maximize the transmission of amplitude information. S3, S5, S6, and S7 used the extended low-frequency response 107-2 aid with the PP earpiece because of its high output level. For S4, whose dynamic range above 500 Hz was particularly small, the frequency response peak of the PP earpiece above 1 kHz limited the overall gain that could be used, so for this subject, the flatter-response W earpiece was used with the extended low-frequency response 107-2 aid. S8, whose LFA loss was relatively high, was fitted with an unmodified Danavox 107-2 aid with a PP earpiece.

For S2, S11, S12, and S13, of whom all but S13 showed minimal dynamic range above 500 Hz, the whole speech signal from the reference aid was not tolerable, but all four were

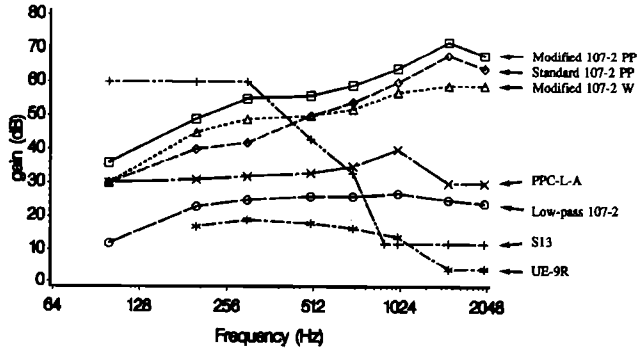

FIG. 3. Free-field frequency responses of the amplifying hearing aids used in this study; the frequency response of the headphone-based amplification used for S13 is also shown. The curves are displaced for clarity, and do not represent the true gains of the aids, which were set to suit each individual patient. The responses from the various Danavox versions were measured with an IEC ear simulator. The Phonak and Unitron curves are from the manufacturers' published ANSI frequency responses.

able to tolerate low-pass filtered speech, using cutoff frequencies of 1000 Hz or below. It was possible to fit low-pass amplifying aids to S2, S11, and S12, but for S13, no available low-pass aid had sufficient gain. S2 and S11, who both had a moderate low-frequency hearing loss, were fitted with a commercial low-pass post-aural aid (Unitron UE-9R). S12 had a greater low-frequency hearing loss, and was fitted with a Danavox 107-2 aid modified to have a mild low-pass characteristic rolling off from 1000 Hz. Test results from S2, S11, and S12 used speech presentation from these aids; the frequency responses of these aids are included in Fig. 3. S13 was tested using low-pass filtered speech presented from a headphone (Beyer DT 48); the frequency response of this amplification system closely followed the comfortable loudness characteristic used in his SiVo aid. The low-pass frequency response of the equipment used to test S13 is included in Fig. 3.

D. Training

Prior to each test session, each subject received speech receptive training making use of a real-time electrolaryngograph-based visual display of voice fundamental frequency (Fourcin and Abberton, 1971; Ball, 1991) which corresponded to the auditory information provided by the SiVo aid. This ensured that the subjects were familiar with the test materials and allowed the experimenters to verify that the aids used were functioning correctly and optimally adjusted.

III. SPEECH PERCEPTUAL ASSESSMENT: SPEECH AMPLIFICATION VERSUS THE SiVo AID

Three speech perceptual assessments were used to give an overall comparison of the relative benefits of the two approaches and to provide an analytic assessment of the degree to which particular speech factors were conveyed in each case. A consonant lipreading task was used to assess the reception of segmental voicing, manner, and place information. A question/statement test was used to assess the perception of intonation; here, performance with speech is unlikely to exceed that with the SiVo signal, since the crucial information is carried by the voice fundamental frequency pattern, and this is presented in near optimal form by the SiVo aid. In subjects whose performance in the first two tests showed an advantage, or the lack of a marked disadvantage, for the SiVo aid, CDT was used to assess the extent to which these auditory segmental and intonational cues could be combined in the perception of continuous speech. CDT was preferred to a sentence test because the available sentence materials have simple syntactic structures, and are thus likely to be rather insensitive to the contribution of prosodic factors.

A. Intervocalic consonant test

The 12-intervocalic-consonant test (Rosen *et al.*, 1979), which uses the consonants /m b p v f n d t s z k g/ in /aˈCa/ syllables, was administered under conditions of lipreading alone, lipreading with the SiVo aid, and lipreading with speech amplification. The stimuli were produced by a female speaker of Southern English Received Pronunciation. Each test list comprised four occurrences of each of the 12 consonants in random order. Five such lists of 48 consonants were used; each of the total of 240 test items was a different token. The test provided a measure of overall proportion correct consonant identification, and information transmission scores for individual phonetic features (Miller and Nicely, 1955).

1. Apparatus

Tests were carried out in a carpeted, sound-isolated, ventilated room, approximately 3×2.5 m. The test material was taken from U-matic video recordings, replayed on a Sony VO 2631 recorder, and 14 in. color monitor (Microvitec 1431). The speech soundtrack was presented through an electrostatic loudspeaker (QUAD PRO-63).[6] The acoustic speech signal was presented at an average SPL of 75 dB (85-dB SPL peak). The microphone of the reference aid was about 1 m from the loudspeaker. The SiVo aid was held in a clamp with its microphone 30 cm from the center of the grille of the loudspeaker.[7]

2. Procedure

Unaided consonant lipreading data, which are little affected by practice (Rosen *et al.*, 1985; Faulkner *et al.*, 1989), were collected early in the study while the subjects were still gaining experience with the alternative aids For 8 of the 11 subjects, the aided lipreading results presented here were collected after the subjects had received at least 6 weeks experience of both aids. This was not possible for three subjects, S7, S8, and S13. S7 and S8 were unavailable for later testing due to illness, and in these instances, aided lipreading data collected during earlier sessions were used. S13 was not able to gain extensive experience of low-pass filtered speech because no suitable wearable aid could be found, and the aided lipreading data both with the SiVo aid and with amplification were collected after six weeks experience of the SiVo aid. Except where noted for individual subjects, the MAPITCH facility of the SiVo aid was not used. Since no variability metric is known for information transmission scores, comparisons between performance in different conditions are made by the χ^2 statistic computed from the proportion correct scores.

3. Results: SiVo aid versus whole speech amplification

The results are shown in Table II. All of the subjects showed significantly higher scores in the aided conditions than when unaided, except for S1 when using whole speech amplification, and, for S7 when using the SiVo aid. S1 and S4 showed higher overall scores with the SiVo signal than with whole speech, but this difference was only statistically significant for S1. For S5, S6, S7, and S8, the overall score was lower with the SiVo signal than with whole speech. This difference reached significance for S5, S7, and S8. For S3, the overall scores with the two aids were within 2%.

Two subjects, S1 and S4, received significantly more voicing information from the SiVo signal than from whole

TABLE II. Consonant identification results; SiVo aid versus whole speech. For each patient, the table shows the percentage correct score overall, and the percentage correct and information transfer percentages for voicing, place of articulation, manner combined, and for nasal and fricative manners separately. Information transfer was calculated from the cumulated confusion matrices. The number of trials is also shown (n). The rightmost three columns indicate the outcome of χ^2 tests comparing the percentage correct scores for SiVo versus whole speech aid (Sx–Sp), SiVo versus unaided lipreading (Sx–LA) and amplifying aid versus unaided lipreading (Sp–LA). The sign of the difference is indicated by " + " or " – ," the significance level by one ($p<0.05$) or two ($p<0.01$) symbols.

Aid	SiVo (Sx)		Speech (Sp)		Unaided (LA)				
Score	$p(c)$	% inf	$p(c)$	% inf	$p(c)$	% inf	Sx–Sp	Sx–LA	Sp–LA
S1									
Overall	69.8		49.0		39.6		++	++	
Voicing	87.5	45.3	67.7	10.6	56.2	1.3	++	++	
Place	99.0	97.5	94.8	89.6	93.8	86.2			
Manner	81.3	48.2	76.1	38.6	68.8	28.0			
Nasality	90.6	34.0	85.4	9.5	83.3	6.1			
Frication	90.6	52.6	90.6	51.6	85.4	37.4			
n	96		96		96				
S3									
Overall	80.2		82.3		43.2			++	++
Voicing	97.4	82.2	94.8	70.5	54.7	1.3		++	++
Place	94.8	90.8	96.4	91.4	90.6	83.4			+
Manner	85.4	53.7	89.6	69.0	80.2	45.2			+
Nasality	94.8	54.6	100	100	85.9	10.7	– –	++	++
Frication	90.6	52.2	89.6	47.5	94.3	65.6			
n	192		192		192				
S4									
Overall	62.0		54.7		43.8			++	+
Voicing	82.8	32.5	72.4	13.8	56.8	1.8	+	++	++
Place	92.2	82.1	95.3	89.2	92.7	84.8			
Manner	77.6	42.8	71.9	31.2	73.4	37.6			
Nasality	86.5	34.7	85.9	26.4	83.9	18.7			
Frication	90.1	50.1	84.9	34.4	89.6	47.8			
n	192		192		192				
S5									
Overall	80.2		91.7		40.3		–	++	++
Voicing	97.9	86.7	97.9	87.7	50.0	1.1		++	++
Place	90.6	87.0	96.9	93.4	91.0	87.3			
Manner	87.5	63.8	96.9	88.2	70.8	27.7	– –	++	++
Nasality	99.0	91.2	100	100	84.0	4.9		++	++
Frication	88.5	44.4	96.9	80.0	86.8	39.0			++
n	96		96		144				
S6									
Overall	80.2		86.5		42.2			++	++
Voicing	94.8	70.5	95.8	75.8	60.4	5.0		++	++
Place	100	100	99.0	97.0	93.2	88.1		+	
Manner	85.4	56.7	90.6	68.0	71.9	34.5		+	++
Nasality	95.8	68.9	95.8	62.8	82.3	8.7		++	++
Frication	89.6	49.8	94.8	71.3	89.6	48.1			
n	96		96		192				
S7									
Overall	54.3		67.6		50.3		–		++
Voicing	71.3	12.8	75.4	18.0	61.4	3.7		++	++
Place	97.2	91.9	99.3	97.1	96.3	89.9			
Manner	75.9	43.8	88.4	64.7	78.8	47.9	– –		++
Nasality	83.7	21.5	97.5	83.6	86.2	31.0			
Frication	92.2	57.4	90.8	53.4	92.6	58.8			
n	288		192		192				
S8									
Overall	53.6		65.6		34.0		–	++	++
Voicing	76.0	20.0	85.4	40.0	50.0	0.0	–	++	++
Place	93.2	83.0	91.7	87.8	90.3	82.3			
Manner	71.9	28.1	75.5	34.5	67.4	24.7			
Nasality	85.4	25.6	84.4	11.3	81.2	3.1			
Frication	82.8	30.6	90.1	49.3	84.7	34.5			
n	192		192		144				

speech. S3, S5, S6, and S7 showed essentially identical voicing scores with both aids, while one subject, S8, received less voicing information from the SiVo signal than from whole speech. In three of the subjects, whole speech provided significantly more manner information than the SiVo signal. For S3 and S7, this was due to the improved reception of the voiced plosive/nasal contrast, while for S5, the manner advantage was largely due to the voiceless fricative/voiceless plosive contrast.

4. Results: SiVo aid versus amplification of low-pass-filtered speech

Table III shows the results for subjects tested with low-pass filtered speech. For S2 and S13, overall performance with the SiVo signal was marginally better than with low-pass speech. For S11 and S12, the overall score was significantly lower with the SiVo signal than with low-pass speech. In all cases, both acoustic signals provided a significant advantage over unaided lipreading.

S2 was able to extract significantly more voicing information from the SiVo signal than from low-pass filtered speech. S11 and S13 showed essentially identical voicing scores with both signals. Only S12 received significantly more voicing information from low-pass speech than from the SiVo signal.[8] For both S11 and S12, significantly more nasality information was received from low-pass filtered speech than from the SiVo signal. It is not clear whether this arose from amplitude envelope information or from low-frequency spectral cues. There was no difference in the transmission of manner information between the SiVo signal and low-pass speech for S2 or S13. As would be expected from the spectral content of the low-pass speech, this signal provided no more frication information than the SiVo signal or unaided lipreading.

B. Question/statement test

1. Procedure

Two lists of 32 sentences were used. The female speaker was the same as used for the intervocalic–consonant test material. The lists were based upon eight sentences, selected so that the final contour contained voiceless as well as voiced consonants (e.g., "They're playing in the garden" and "She's reading a newspaper?"). Each sentence occurred four times in each list, twice as a statement, and twice as a question. Voice fundamental frequency at the end of a statement typically fell from 270 to 150 Hz over about 200 ms. At the end of a question it rose from 135 to 230 Hz, again over about 200 ms. The materials were presented without lipreading. The apparatus was the same as used for the consonant lipreading tests, except that the video monitor was not used. The subject was asked to indicate whether a sentence was heard as a question or a statement. S7 was unwilling to at-

TABLE III. Consonant identification results; SiVo aid versus low-pass filtered speech. The table entries follow those of Table II. Speech was presented from a low-pass hearing aid except for S13, with whom headphone presentation was used.

Aid	SiVo(*Sx*)		Speech(*Sp*)		Unaided(*LA*)				
Score	*p*(*c*)	% inf	*p*(*c*)	% inf	*p*(*c*)	% inf	*Sx*–*Sp*	*Sx*–*LA*	*Sp*–*LA*
S2									
Overall	66.7		57.3		36.5			++	++
Voicing	92.7	69.7	79.2	26.9	53.1	1.1	++	++	++
Place	90.6	84.3	92.7	88.7	92.7	88.8			
Manner	74.0	34.4	72.9	40.1	75.0	36.0			
Nasality	83.3	0.0	87.5	17.6	85.4	8.5			
Frication	88.5	54.5	84.4	51.8	89.6	52.0			
n	192		192		96				
S11									
Overall	79.2		89.6		44.6		− −	++	++
Voicing	97.9	85.3	100	100	57.8	1.6		++	++
Place	95.8	88.6	96.2	89.8	91.6	83.8			+
Manner	83.3	50.6	92.7	78.4	76.0	35.7	− −		++
Nasality	95.8	66.4	100	100	85.4	11.1	− −	++	++
Frication	87.0	39.7	92.7	62.6	90.2	49.9			
n	192		288		287				
S12									
Overall	60.4		78.1		50.0		− −		++
Voicing	74.5	17.0	92.2	62.4	65.0	7.1	− −	+	++
Place	98.4	95.9	99.0	97.0	95.4	87.3			
Manner	78.6	40.6	85.4	56.4	73.8	33.5			++
Nasality	91.1	39.4	97.9	78.9	85.4	17.3	− −	+	++
Frication	87.0	42.1	87.5	41.1	87.9	42.9			
n	192		192		240				
S13									
Overall	83.3		81.3		44.3			++	++
Voicing	96.9	80.1	95.3	73.0	65.1	7.0		++	++
Place	99.0	97.0	97.9	93.8	96.9	91.4			
Manner	85.4	55.4	87.0	58.6	70.3	31.8		++	++
Nasality	94.8	56.0	97.4	74.1	83.3	6.1		+	++
Frication	90.6	54.2	89.6	42.6	87.0	43.6			
n	96		192		192				

tempt the task, although he had been able to perform well in practice.

2. Results and discussion

The results from ten subjects are shown in Table IV. Results from S1 have been published previously (Rosen and Fourcin, 1983; Rosen *et al.*, 1987). Four subjects (S1, S4, S6, S8) performed significantly better with the SiVo signal than with whole speech. For S1 and S4, the MAPITCH option was particularly valuable. S2, S12, and S13, who were tested with low-pass speech, each showed marginally better performance with the SiVo signal, but the differences did not reach significance; it may well be that low-pass speech has similar advantages to the SiVo signal for the reception of intonation contrasts. Three subjects (S3, S5, S11) obtained almost perfect scores with both aids. Over the group, performance using the SiVo aid was significantly better than with speech amplification according to a sign test ($p = 0.008$).

TABLE IV. Results of question/statement test using the SiVo aid and amplified speech. S1, S4, and S13 were tested with the SiVo aid both with fundamental frequency mapped down by 50 Hz (*Sx*–50) and without mapping. The rightmost two columns refer to the significance of χ^2 tests between scores with the SiVo aid and with speech. (Results from S1 were previously reported by Rosen *et al.*, 1987.) S7 refused to complete the task, although he was able to perform well in a practice session where feedback was available.

Aid	Speech (*Sp*)	SiVo (*Sx*)	SiVo–50 Hz (*Sx*–50)	*Sx* versus *Sp*	*Sx*–50 versus *Sp*
S1	56	76	82	++	++
S2	88	100			
S3	92	95			
S4	47	50	64		++
S5	98	98			
S6	61	81		++	
S8	47	66		+	
S11	100	100			
S12	68	75			
S13	78	86	83		

with low-pass speech, each showed marginally better performance with the SiVo signal, but the differences did not reach significance; it may well be that low-pass speech has similar advantages to the SiVo signal for the reception of intonation contrasts. Three subjects (S3, S5, S11) obtained almost perfect scores with both aids. Over the group, performance using the SiVo aid was significantly better than with speech amplification according to a sign test ($p = 0.008$).

C. Connected discourse tracking (CDT)

Six subjects (S1, S2, S3, S4, S11, and S12) took part in CDT to assess the extent to which the segmental and prosodic information available from the SiVo aid can contribute to the perception of continuous speech. In the consonant lipreading tests, these subjects had either performed similarly with the two aids or had shown greater benefit from the SiVo aid.

1. Procedure

The text was taken from graded reading material for students of English as a second language (Milne, 1977). A female Southern English Received Pronunciation speaker (author VB) administered the tests. Presentation varied randomly across 5 min sessions between unaided lipreading, lipreading with whole speech (S1, S3, S4), or low-pass speech (S2, S11, S12), and lipreading with the SiVo signal. Phrases were produced repeatedly by the talker until the subject had correctly repeated back each word. If the subject was not successful after three repetitions, the talker paraphrased twice or spelled the words, and if that was also unsuccessful, finally wrote out the words. For most of the testing, the SiVo aid was used in the voice fundamental frequency condition. To avoid possible errors in fundamental frequency extraction, some later tests used error-free voice fundamental frequency extraction based on an electrolaryngograph; the sinusoidal signal was generated as before by the subject's own SiVo aid, and presented through Beyer DT 48 earphones.

2. Results and discussion

The results are shown in Table V. According to t tests, none of the individual subjects showed significant differences between tracking rate with whole speech and with the SiVo signal. S1, who had previously shown a substantial advantage in this task for the SiVo signal as compared to whole speech (Rosen *et al.*, 1987), showed only a 7-words/min advantage here. In the four subjects for whom comparisons were made between aided and unaided lipreading, two (S3, S11) performed significantly better with the SiVo signal than unaided. These two subjects and S12 also performed significantly better with amplified speech than unaided. S2 also appeared to track more slowly unaided than in the two aided conditions, but insufficient data was collected to establish a significant difference. This subject showed similar tracking rates in both aided conditions. For S4, unaided tracking performance was substantially faster than that of the other subjects, and was probably close to ceiling levels for the text. There was, therefore, little room for improvement in tracking rate when auditory information was added.

Since none of these subjects performed significantly more poorly with the SiVo signal than with speech, it seems that fundamental frequency information remains useful to these listeners in a task where the information rate is substantially higher than in the identification of isolated intervocalic consonants. Where subjects were tested using both the laryngograph and the SiVo aid's acoustic fundamental frequency extractor to provide fundamental frequency information, there was no indication that errors arising from the SiVo aid led to significant difficulties in the test environment.

TABLE V. Results of connected discourse tracking. The table shows the average score in words per minute, the number of 5-min sessions, and the standard deviation where there were more than three sessions. The unaided lipreading data for S1 (*) were taken from a later experiment (Sec. IV C). The second set of data from S3 (#) was collected using a more complex text than the first.

Aid	SiVo	Speech	Unaided	Source of voice F_0 information
S1	44.0 (12:5.0)	37.0 (12:11.8)	26.7 (6:5.2)*	SiVo
S2	68.6 (1)	66.2 (1)	55.4 (1)	SiVo
S2	76.5 (3)	83.5 (3)		SiVo
S2	72.6 (2)	77.8 (2)		Laryngograph
S3	79.0 (5:6.0)	87.3 (5:4.0)	57.3 (4:5.3)	SiVo
S3#	69.1 (4:1.9)	65.7 (4:1.5)		Laryngograph
S4	68.4 (4:4.4)	70.9 (4:8.9)	73.3 (3)	SiVo
S11	68.0 (4:3.7)	70.1 (4:5.0)	31.6 (2)	Laryngograph
S12	52.3 (4:3.2)	60.7 (4:4.1)	43.6 (4:4.6)	Laryngograph

D. Patients' use of the SiVo aid

Four subjects (S1, S2, S3, S4) continue to make regular use of the SiVo aid, while a fifth (S13), makes occasional use of the aid. S1, S4, and S13 no longer make use of amplifying aids, while S2 and S3 continue to use amplifying aids when they need to hear nonspeech sounds, and when the size of the SiVo aid makes its use inconvenient. The remaining six subjects (S5, S6, S7, S8, S11, and S12) prefer to use the amplifying aids used in the above tests rather than the SiVo aid or the amplifying aids which they had previously used.

E. Overview of results with the basic SiVo aid

The provision of simple voice fundamental frequency information by the SiVo aid clearly provides important help in speech communication for the five subjects who continue to use the aid in the field. These subjects showed consonant identification scores with the SiVo aid which were either higher than, or very similar to, those obtained with amplification. For the group as a whole, the reception of intonation with the SiVo aid was enhanced compared to amplification. It is, perhaps, surprising that the subjects who use the SiVo aid regularly did not also show an advantage compared to amplification from its use in CDT, where intonation and segmental voicing information might be expected to be dominant auditory factors in aiding lipreading. That such an advantage was not found here does not seem to be due to errors in the fundamental frequency extraction of the SiVo aid, and may in part be due to the ability of some of the subjects to extract consonantal manner information from speech, or to a degree of sensitivity to vowel quality information based on first formant information (Rosen *et al.*, 1987; Faulkner *et al.*, 1990c). One factor that may contribute to some subjects' preference for the SiVo aid is the clear feedback which it provides regarding voice regularity and pitch control (Ball *et al.*, 1990).

It was not anticipated that a number of these subjects would be able to extract such a range of manner of articulation information from amplified speech. More nasality information was often available from lipreading with amplified speech than from lipreading with the SiVo aid, even for subjects such as S2, S3, and S13, whose overall scores in consonant identification were no better with amplified speech than with the SiVo signal. Furthermore, one subject (S5) showed better reception of frication information when lipreading with speech than when using the SiVo aid. These findings point to important residual abilities, and it may be hypothesized that these are likely to be better utilized by compound speech pattern aids which encode and present aspects of manner information in addition to voice fundamental frequency. This hypothesis was tested by the studies described in Sec. IV.

IV. COMPOUND SPEECH PATTERN PRESENTATION

A. Acoustic patterns related to manner information

The auditory abilities that enabled some of these subjects to extract manner of articulation information from amplified speech are unlikely to be optimally utilized with conventional hearing aids. One important acoustic pattern in speech is related to speech amplitude, which in voiced speech is strongly correlated with the extent of oral closure and the presence of nasal airflow, and hence can contrast plosive, nasal and semivowel manners (Summerfield, 1983). The amplitude compression required to fit an amplifying aid to a limited dynamic range is likely to reduce the salience of such cues. Since, in commercial hearing aids, compression is rarely frequency dependent, amplitude variation will be poorly matched to dynamic range when, as is common in this population, the audiogram is steeply sloping.

A second major factor of consonantal manner, the aperiodic energy that represents fricative and other turbulent speech components, is largely present in frequency regions beyond the receptive range of these patients, and cannot be made available at any setting of an amplifying aid. We have encoded voiceless excitation as a low-frequency random noise over a frequency band chosen to match the subject's auditory area. In the present study, this band is from 50 to between 400 and 1000 Hz, depending upon the subject. At the present time, voiceless excitation is not coded during the mixed excitation found in voiced fricatives.

In the compound speech pattern signals used here, voiceless speech components are represented by a low-frequency noise, while amplitude information is encoded by amplitude modulation around the most comfortable listening level for the noise or for the sinusoidal voice fundamental frequency signal. Figure 4 shows speech pressure waveforms of the same consonants as shown in Fig. 1, together with these compound speech pattern signals. The voiceless excitation pattern is present whenever the speech signal shows no voicing and has a predominantly high-frequency energy content; its presence is signaled here by a 50- to 700-Hz band of noise. The amplitude pattern is derived from full-wave rectification of the wide-band speech signal, followed by low-pass filtering at 20 Hz and 24 dB/oct. The envelope so derived is used to amplitude modulate the *Sx* and noise signals.

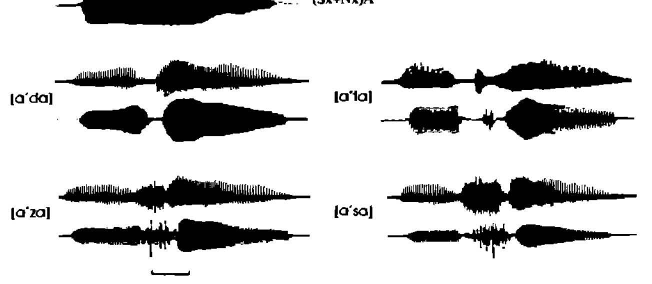

FIG. 4. Compound speech pattern processing of alveolar intervocalic consonants [ɑˈnɑ], [ɑˈdɑ], [ɑˈzɑ], [ɑˈtɑ], and [ɑˈsɑ]. The upper panel shows the speech pressure waveform, and the lower panel shows a compound speech pattern signal representing voice fundamental frequency, voiceless excitation, and the overall amplitude envelope, labelled as "$(Sx+Nx)A$."

The amplitude patterning provides a prominent cue to the release of the voiced plosive [d], which is thus distinguished from the voiced nasal [n]; this contrast can only be made from the simple voice fundamental frequency pattern (see Fig. 1) when the voiced plosive is devoiced. The voiceless excitation patterning contrasts the two voiceless consonants [s] and [t] through the gross temporal patterning of voiced and voiceless excitation; the *Sx* signal, since it represents only voiced excitation, does not do so. The voiceless excitation pattern also indicates the devoicing seen in this token of the voiced fricative [z]. Consonants having different places of articulation will be similarly contrasted by these patterns. For example, the bilabials /m/, /b/, and /p/ will show analogous patterns to the corresponding alveolars /n/, /d/, and /t/, although the amplitude of the /p/ release burst will generally be lower than that of /t/.

Studies of aided lipreading in normal listeners have shown that the addition of amplitude information to voice fundamental frequency information can lead to faster CDT performance (Grant *et al.*, 1985), and can provide additional useful cues to nasal and plosive consonantal contrasts (Breeuwer and Plomp, 1985). In profoundly hearing-impaired listeners, Grant (1987a) examined the effect of adding amplitude information to fundamental frequency information in audio–visual CDT; he reported an advantage for three listeners, but a disadvantage for a fourth. A voiceless pattern similar to that proposed here has been shown to assist lipreading in normal listeners and two profoundly hearing-impaired subjects by distinguishing voiceless plosives and affricates from voiceless fricatives, and by enhancing the voiced/voiceless contrast in plosives (Faulkner *et al.*, 1989). The potential of a combination of these three factors in the profoundly hearing impaired has not been examined previously.

B. Speech receptive assessment

1. Subjects

Five subjects (S1, S2, S3, S11, and S12), who represented a range of receptive abilities and degrees of benefit from the SiVo aid, were used in these tests. S2, S3, and S11 had shown high performance in discriminating periodic and aperiodic stimuli 50 ms or less in duration, while S1 and S12 performed relatively poorly in this task (Faulkner *et al.*, 1990b).

2. Methods

a. Speech materials. Intervocalic consonant and connected discourse tracking assessments were again used. The speakers were as described earlier.

b. Speech pattern extraction and coding. Voice fundamental frequency information was derived from electrolaryngograph measurements of larynx period which were recorded on one audio channel of the video-tape for the consonant stimuli, and taken live from the speaker in CDT. Voiceless excitation was detected by a spectral-balance comparator operating on the speech signal, and based upon a 450-Hz low-pass filter and a 3- to 10-kHz bandpass filter. The comparator's output was overridden if voicing was detected by the electrolaryngograph. Voice fundamental frequency was presented as a sinusoidal signal (Sx) derived from an SiVo aid programmed to match the subject's comfortable listening levels. Voiceless excitation was represented by a band of noise with a lower cutoff frequency of 50 Hz, and an upper cutoff set to suit individual patients; this was 400 Hz for S1, 500 Hz for S12, 700 Hz for S2 and S11, and 1000 Hz for S3. The upper limit was chosen so that the range of noise intensities employed when amplitude coding was in use did not exceed the subject's dynamic range. The comfortable listening level for the noise was set independently of the level of the sinusoid. Amplitude envelope information was extracted by passing the wideband speech signal through a full-wave rectifier and a 20-Hz, 24-dB/oct low-pass filter. The resulting envelope signal was digitized (8 bits at 10 kHz) and reconstructed by a multiplying digital-to-analog converter for which the reference voltage was the Sx signal from the SiVo aid or the voiceless noise signal. The signals shown in Fig. 4 were produced by this processing. The low-pass filter used for amplitude envelope extraction led to a delay of about 30 ms in the envelope relative to the original speech signal. Average presentation levels were set to be comfortable and clearly audible during normal speech. The maximum levels presented to S1 and S11 were kept within the listener's dynamic range by digitally clipping the envelope to a 12-dB range; for the other subjects the amplitude range was determined directly by the speech amplitude envelope.

c. Procedure. The presentation conditions are designated as follows: $L+Sp$: lipreading with amplified speech, $L+Sx$: lipreading with Sx, $L+Sx+Nx$: lipreading with Sx and voiceless information (Nx), $L+(Sx)A$: lipreading with Sx and amplitude information (A), $L+(Sx+Nx)A$: lipreading with Sx, voiceless, and amplitude information.

Speech pattern signals were presented through headphones (Beyer DT 48), except for S1, who complained of vibration from the headphone body. For him the headphone was replaced by an Oticon CP100 hearing aid receiver and earmould as used in his SiVo aid. Speech was presented through the same hearing aids as described earlier except to S1, for whom speech was directly presented to the same hearing aid receiver as used for the speech pattern signals. While all of the subjects were experienced with speech amplification and the SiVo aid, they had received little experience of the other speech patterns; S11 and S12 received about 30 min training prior to testing. S1, S2, and S3 had less than 10 h experience with the combined $Sx+Nx$ patterns. All of the subjects were given interactive training with the consonant materials using a voiced/voiceless visual speech display (Bootle, 1986). All presentation conditions were used in the consonant identification task, while only selected conditions were employed in the more time-consuming CDT task.

3. Results and discussion

a. Intervocalic consonant test. Figure 5 shows the results. These data were all newly collected, and differ in detail from comparable entries in Tables II and III. All five listeners could make use of additional speech pattern information in combination with Sx. Group scores in different conditions

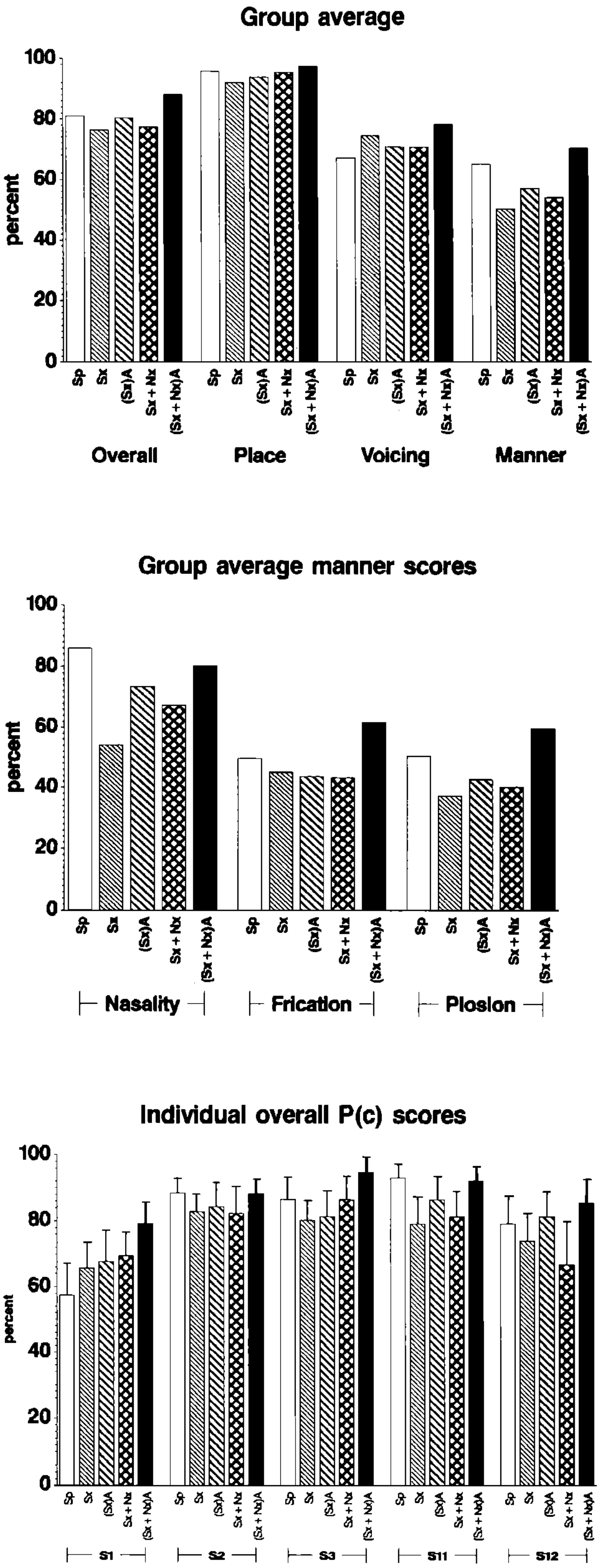

FIG. 5. Scores in the intervocalic consonant test obtained in lipreading with speech and with compound speech pattern presentation. The upper panel shows the overall percentage correct score and the information transmission scores for voicing, manner, and place information averaged across the five subjects. The middle panel shows the group average information transmission for the three manner features of nasality, plosion, and frication. The lower panel shows the overall percentage correct scores for the individual subjects; the error bars represent upper 95% confidence limits.

were compared by Tukey HSD tests ($\alpha = 0.95$) based on a general linear model analysis performed by the GLM procedure of SAS (SAS Institute Inc, 1987), in which the condition by subject interaction was used as the error term. The overall correct score and the transmission of manner information were both significantly higher in condition $L+(Sx+Nx)A$ than in $L+Sx$ and $L+(Sx+Nx)$; the transmission of voicing and place information did not differ between conditions. The transmission of information for the three manner features which vary in the consonant set used here is shown in the central panel of Fig. 5. The transmission of nasality was highest with speech, and was significantly higher in all conditions where amplitude information was present than in condition $L+Sx$. The transmission of plosion information was significantly higher in condition $L+(Sx+Nx)A$ than in conditions $L+Sx$ and $L+(Sx+Nx)$, while the transmission of frication information did not differ significantly between conditions. The Nx signal did not lead to improved transmission of frication or plosion information in the absence of amplitude information; this is probably related to the role of amplitude envelope in contrasting gradual fricative onsets from the more abrupt onsets of the burst and friction noise of voiceless plosives.

In individual subjects, 95% confidence limits based on the binomial distribution (see bottom panel of Fig. 5) showed that the $(Sx+Nx)A$ signal produced significantly better performance than amplified speech for two subjects (S1, S3). Performance with the $(Sx+Nx)A$ signal was also marginally better than with speech for S12. For S2 and S11, there was only a very small overall performance difference between $(Sx+Nx)A$ and speech. Except for S1, the simple Sx signal provided less overall information than that available from speech through the subjects' usual hearing aids. In contrast to the results obtained from S12 with the wearable SiVo aid, this subject's performance with Sx was not markedly poorer than her performance with low-pass filtered speech.

b. Connected discourse tracking. CDT results are shown in Fig. 6. These data were newly collected, and do not correspond to those in Table V. Data from each individual were analyzed using the SAS GLM procedure (SAS Institute Inc, 1987) and the Tukey HSD test ($\alpha = 0.95$), using the variation over sessions as the error term. When unaided lipreading was included (S1, S11, S12), scores in the aided conditions were significantly better than unaided scores in all cases except for S1 when aided by amplified speech. There were significant differences between the aided conditions for three of the five subjects. S1 tracked significantly faster by a margin of 12 words/min in condition $L+Sx$ than in condition $L+Sp$ (34.1 words/min), and significantly faster in condition $L+(Sx)A$ than with speech by 17 words/min, that is, a 50% proportional increase. S1 tracked marginally slower in condition $L+(Sx+Nx)A$ than in condition $L+Sx$. For S3, tracking rate in condition $L+(Sx+Nx)A$ was significantly faster by 9.1 words/min than in condition $L+Sp$ (68 words/min), while performance in conditions $L+Sp$ and $L+Sx$ did not differ significantly. For S12, tracking in condition $L+Sx(A)$ was significantly faster by 12 words/min than in condition $L+Sp$ (56.3 words/min), and condition $L+(Sx+Nx)A$ was, for this subject, significantly slower by 12 words/min than condition $L+(Sx)A$. S2 and S11, whose intervocalic consonant scores were similar with speech and the $(Sx+Nx)A$ signals, both achieved relatively fast tracking rates (about 72 words/min in condition $L+Sp$) and showed no significant differences between the aided conditions.

C. Conclusions

These laboratory studies have shown that in lipreading both consonants and continuous prose, profoundly hearing-impaired listeners can make significant use of matched speech pattern information beyond that provided by fundamental frequency. For three of the five subjects, at least one of the compound speech pattern conditions produced better aided lipreading performance than that obtained with speech presentation, while the other two subjects performed similarly in both tasks with amplified speech and in the more effective speech pattern conditions. In CDT, the addition of amplitude information to voice fundamental frequency information appears to provide a general advantage in the present group of profoundly hearing-impaired listeners, although Grant (1987a,c) has shown that this may not always be true. In consonant identification, voiceless speech information generally provided an additional advantage. In CDT, however, a further advantage was not seen. There are a number of possible reasons why a consistent advantage of voiceless information was not found. First, this may be because phonetic detail is more significant in consonant identification than in CDT. Second, it may be that there was an imperfect match of the noise signal to the listener's limited dynamic range given the much greater amplitude variations in the speech materials used in CDT than in the consonant materials. Third, it may be that there is a difficulty in processing the additional voiceless information in connected speech, where speech rates are higher; this difficulty would be likely to be overcome with further experience of the compound speech pattern signal. In any case, in a practical aid using this coding strategy, the range of amplitude modulation for both voiceless and voiced speech components would need to be dynamically matched to the subject's dynamic range for the carrier signal.

V. AUDITORY ABILITIES

To discover the basic auditory abilities that underlie the speech perceptual performance of this group and the individual variations between subjects, three psychoacoustic measurements were performed. These were: (1) pure-tone frequency difference limens at 125 and 250 Hz to establish pitch discriminability in the voice fundamental frequency region; (2) a highly simplified psychoacoustic tuning curve to establish a measure of frequency selectivity at low frequencies; and (3) gap detection thresholds to give an indication of temporal acuity.

A. Common apparatus

Stimuli were always presented monaurally, and, except in the gap detection task, through Connevans CE-8 head-

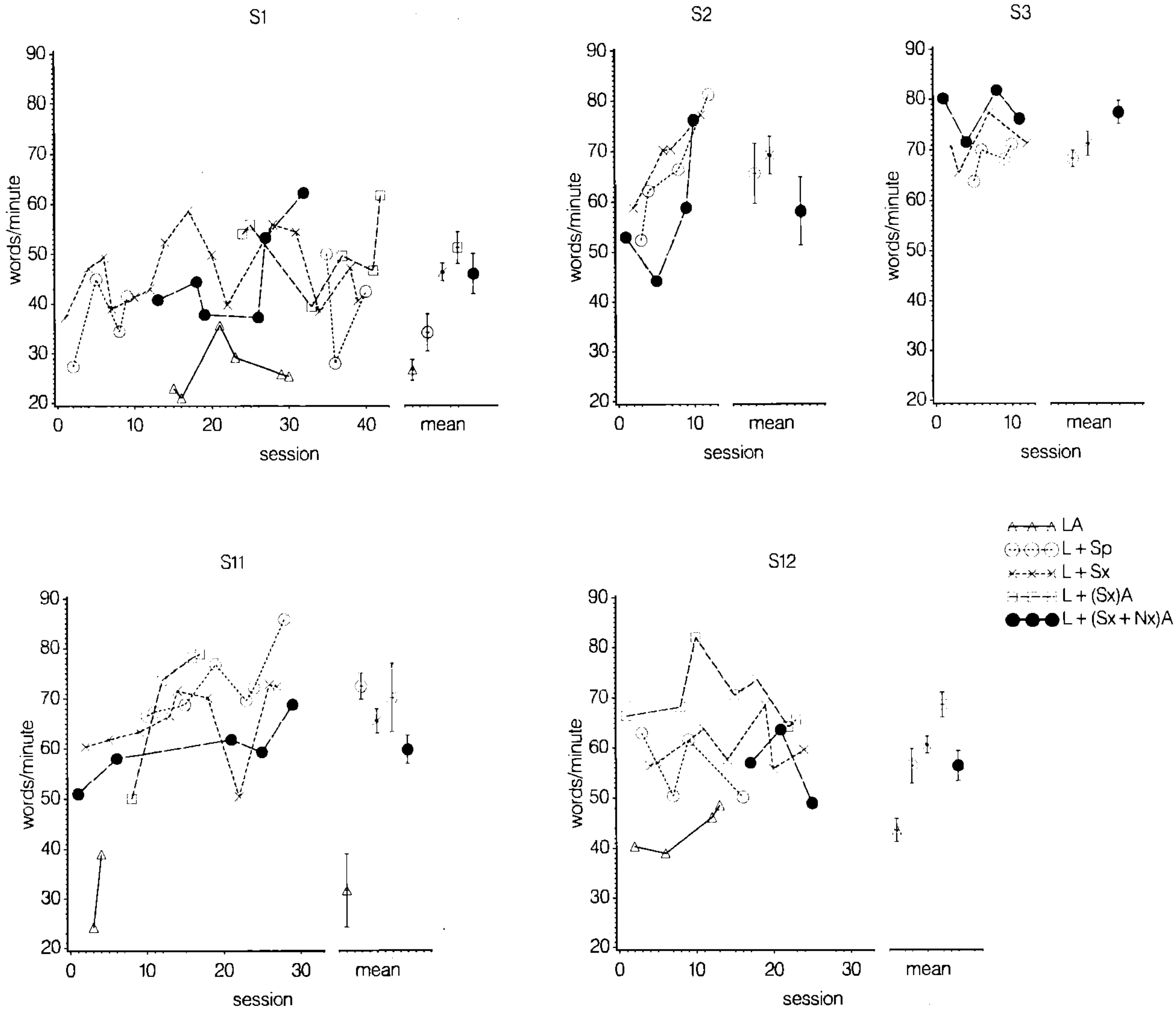

FIG. 6. Connected discourse tracking rates over a series of 5-min test sessions. At the right of each panel appear mean tracking rates and error bars representing $\pm$ 1 standard error. Data are shown for unaided lipreading (*LA*; open triangles), $L+Sp$ (open circles), $L+Sx$($\times$), $L+(Sx)A$ (open squares) and $L+(Sx+Nx)A$ (closed circles).

phones. These were chosen for their extended low-frequency response and high-output capability (up to 138 dB SPL at 125 Hz). For gap detection, Beyer DT 48 headphones were used, as the stimulus was a noise band from 50 to 1000 Hz, and the Beyer headphones have a flatter frequency response above 500 Hz. The headphones were driven by a Yamaha P200 amplifier. The headphone output was monitored using a microphone mounted on the grid of the headphone and a real-time spectrum analyzer (Rosen and Nevard, 1987).

B. Low-frequency difference limen

1. Procedure

The sinusoidal stimuli were 400 ms long, inclusive of 50-ms raised-cosine onset and offset ramps. A 2I-2AFC task was used, with an interstimulus interval of 400 ms. Stimuli were digitally synthesized, played out at a 10-kHz sampling rate, and low-pass filtered at 4 kHz (Kemo VBF/8). The stimulus level was made comfortable, and extreme frequency differences of 25% were presented to establish the degree of loudness change with frequency over that range. Equalization was employed where consistent loudness changes were noted. In addition, stimulus levels were varied randomly over a 3-dB range to minimize the usefulness of residual frequency-related loudness differences. In the region of the frequency difference thresholds typically observed, between about 4% and 12%, loudness differences with frequency were estimated to be smaller than those produced by the level variation. Subjects received interactive practice, with the experimenter gradually reducing the frequency difference from an extreme value until errors were made, and increasing it again to a value where the subject could comfortably perform the discrimination. A one-up, three-down adaptive staircase was used, with a multiplicative step size of 1.414. Each threshold was based on the geometric mean of at least 20 reversals of the staircase.

2. Results and discussion

The frequency difference limens (DLs) obtained are shown in Table VI. The DLs shown by the better performing subjects were quite similar to those from normal listeners, where, for example, at 200 Hz and at comparable sensation levels, Wier *et al.* (1977) found DLs (based on a 71% correct threshold) of 1.5% at 10 dB SL, and 0.7% at 20 dB SL in highly practiced subjects. Some subjects, however, showed performance an order of magnitude poorer than normal listeners. It is interesting that, for seven of these nine subjects, the relative DL increased with frequency, which is opposite to the effect found for normal listeners, where the relative DL at 250 Hz is smaller than at 125 Hz (Nordmark, 1968; Stock and Rosen, 1986). Stock and Rosen (1986) reported this inverted effect of frequency for a different profoundly hearing-impaired listener, and similar trends have also been reported by Grant (1987c).

C. Simplified psychoacoustic tuning curve (PTC)

The most practical measure of frequency selectivity was considered to be a highly simplified PTC based on a narrowband noise (NBN) masker. An 80-Hz wide NBN was chosen since this bandwidth is narrower than the filter bandwidths expected in these listeners, yet not so narrow as to give rise to audible fluctuations in level. Listeners with moderate hearing loss show bandwidths two to three times larger than normal (Moore and Glasberg, 1986) and normal auditory-filter bandwidths are around 40 Hz at 125 Hz and 60 Hz at 250 Hz (see for example, Moore and Glasberg, 1983). Hence, we expected the filter bandwidths of these profoundly hearing-impaired listeners to be considerably greater than 80 Hz. Although measures of bandwidth and filter-shape based upon this method can be affected by off-frequency listening (Patterson and Moore, 1986), the gross indication of the extent of frequency selectivity required here is unlikely to be affected by this factor.

TABLE VI. Frequency DLs for sinusoids. In addition to the DL, the table shows the stimulus frequency (f), the sound-pressure level (SPL) and sensation level (SL) at which the test was performed. DLs for S1 and S2 are taken from Rosen *et al.* (1987).

	f (Hz)	SPL	SL	DL (%)
S1	100	120	10	4.3
	240	126	3	5.7
S2	100	72	10	3.2
	200	95	28	3.5
S3	125	122	21	3.8
	250	124	13	6.7
S4	125	112	10	7.5
	250	121	10	11.9
S5	125	106	31	5.9
	250	112	24	9.9
S6	125	116	13	8.2
	250	127	23	10.2
S11	125	101	32	5.5
	250	99	24	2.2
S12	125	120	25	9.0
	250	125	30	18.3
S13	125	107	10	41.9
	250	118	10	7.5

1. Procedure

A fixed level sinusoidal probe at 10 dB SL was presented with a variable level 80-Hz wide NBN masker. The NBN was generated by the multiplication method (Weber, 1977); the noise spectrum was about 90 dB down 40 Hz below and above the passband edges. The probe frequencies and masker center frequencies were either 125 or 250 Hz: All four combinations of probe and masker frequency were presented. The probe was 200 ms in duration, and presented 150 ms after the onset of a 500-ms masker. Both probe and masker had 20-ms raised-cosine onset and offset ramps. The noise level required to mask the probe was estimated in a 2I-2AFC staircase converging on the 71% correct point and using a 1-dB step size. A minimum of 20 reversals was used for each threshold estimate. One normally hearing listener was also tested to provide a point of comparison. For this listener, the probe level was either 20 or 40 dB SL. At the higher probe level, the SPLs for the off-frequency masker were similar to those for the impaired listeners.

2. Results and discussion

The simplified PTCs are shown in Fig. 7, together with absolute thresholds at the probe frequencies, determined in a 2IFC task (note 8). If selectivity is absent, the PTCs should be parallel to each other and also parallel to the absolute threshold function. On this basis, the PTCs observed from S1 and S4 show no evidence of frequency selectivity. For the other seven subjects, the two PTCs deviate from each other, and except for S13, the masker level for the 250-Hz probe is greater for the 125-Hz noise than for the 250-Hz noise. The PTCs for the 125-Hz probe do not in general show a clear deviation from the slope of the absolute threshold function.

The average masking level difference (MLD) between off-probe and on-probe maskers was taken to be an overall measure of frequency selectivity. The MLDs for each subject are shown in Table VII. The normal listener showed an MLD of 39 dB. S2, the impaired listener showing the greatest selectivity, showed an MLD of 14 dB, about one-third of that for the normal listener. More detailed PTC measurements from most of these subjects are described by Faulkner *et al.* (1990c). Those data indicated that when the audiometric loss at the probe frequency was 95 dB or less, selectivity was present. Where auditory filter bandwidths could be estimated, these were between two and three times wider than in normal listeners.

D. Gap detection

1. Procedure

The stimulus was a 400-ms burst of noise, which was bandpass filtered between 50 and 1000 Hz after gating. The gap was centered in the noise. The noise was presented monaurally at a comfortable listening level. Thresholds were determined in a 2I-2AFC task, using an adaptive staircase converging on the 79% correct point. Each staircase run

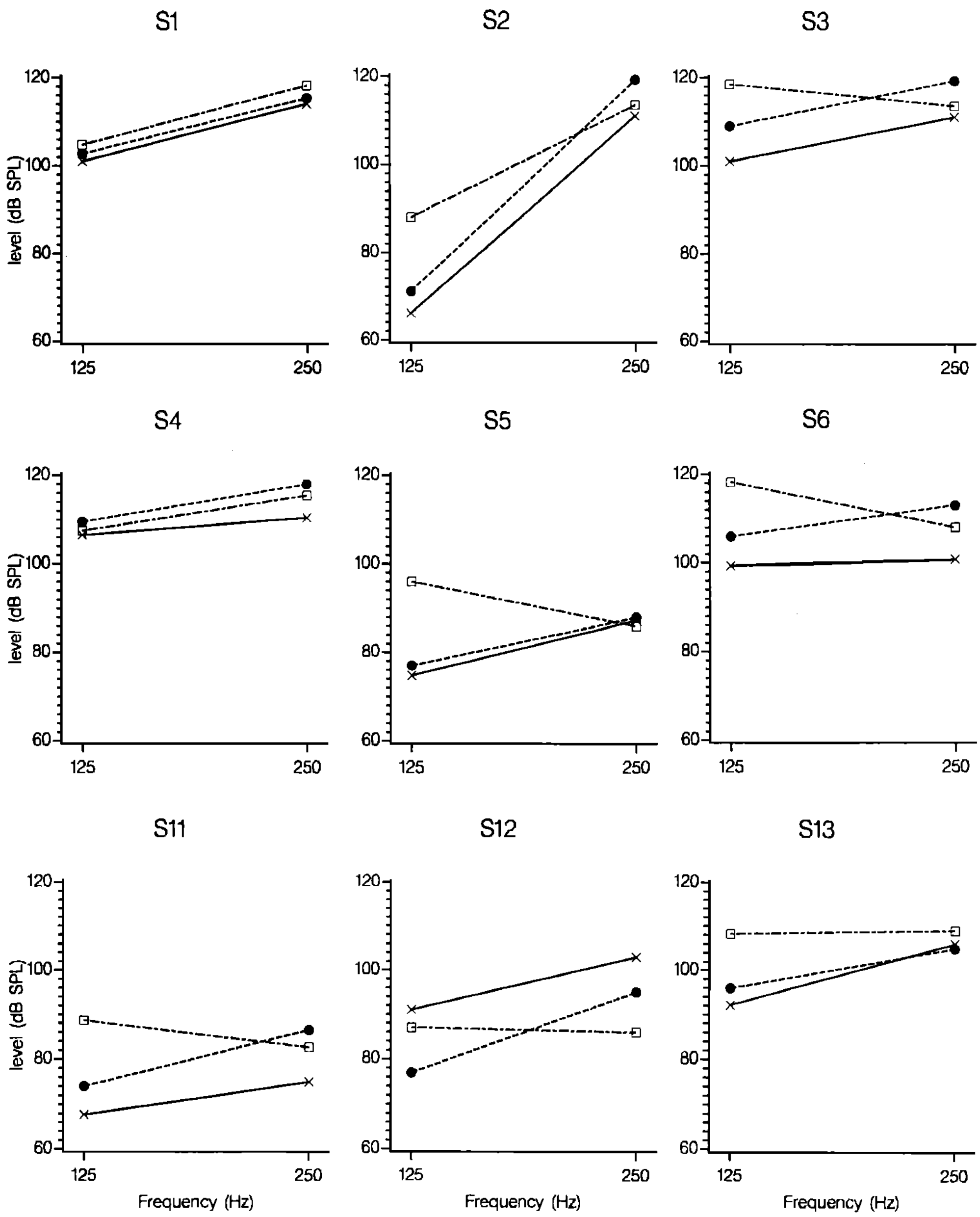

FIG. 7. Simplified psychoacoustic tuning curves. The figure shows the 2IFC absolute threshold in dB SPL at 125 and 250 Hz (× and solid line), and the level in dB SPL of an 80-Hz wide masker centered on 125 and 250 Hz which masked a 10-dB SL probe at 125 Hz (solid circle and dashed line) or 250 Hz (open square and dashed line).

comprised ten reversals; thresholds were based upon at least two staircase runs. The interstimulus interval was 400 ms.

2. Results and discussion

Gap thresholds were obtained from nine of the subjects, and are shown in Table VIII. Seven subjects showed thresholds between 20 and 26 ms, while S1 and S13 gave thresholds around 40 ms. A normal listener would show a threshold in the region of 8 ms (Shailer and Moore, 1983).

E. Auditory abilities, speech perceptual results, and the benefit of the SiVo aid

1. Interrelationships between audiometric and psychoacoustic measures

Table IX shows the correlations between the audiometric and psychoacoustic measures collected from 11 subjects. As would be expected, the hearing loss measures are highly intercorrelated. LFA, 3FA, and 4FA hearing loss are also significantly negatively correlated with low-frequency dynamic range, but not with 3FA dynamic range. Frequency difference limens showed no significant correlation with other measures. The MLD measure of frequency selectivity was significantly negatively correlated with 4FA and LFA hearing loss, and positively correlated with low-frequency dynamic range. Low-frequency hearing loss, dynamic range, and frequency selectivity, are presumably all related to the number of hair-cell and auditory-nerve processes that remain functional. Finally, the gap threshold measure was significantly correlated only with the MLD measure, which suggests that both tasks have some common component which is unrelated to hearing loss, dynamic range, or frequency discrimination. In order to examine the relationships between auditory abilities and speech measures, a Principal

TABLE VII. Results of simplified PTC procedure with probe frequencies of 125 and 250 Hz. The table shows probe levels and masker levels in dB SPL at masked threshold, and the masking level difference (MLD) between the off-probe and on-probe masker levels for each probe frequency and averaged over both probe frequencies.

	Probe frequency	Probe SPL	125-Hz masker level	250-Hz masker level	MLD	Average MLD
S1	125	111	103	115	12.5	− 1.0
	250	124	105	118	− 14.4	
S2	125	77	71	101	30	14.0
	250	97	88	90	− 2	
S3	125	111	109	120	11	8.0
	250	121	119	114	5	
S4	125	115	109	118	9	0.5
	250	120	108	116	− 8	
S5	125	85	77	88	11	10.5
	250	98	96	86	10	
S6	125	113	107	115	8	7.5
	250	114	118	111	7	
S11	125	80	72	86	14	10.5
	250	85	91	82	9	
S12	125	101	77	87	10	9.5
	250	113	95	86	9	
S13	125	102	96	108	12	4.0
	250	116	105	109	− 4	
Normal listener						
	125	61	66	94	28	39.0
	125	81	86	114	28	
	250	51	82	30	52	
	250	71	95	47	48	

Components analysis (SAS Institute Inc., 1987), was applied to the auditory measures. This analysis produced three components having eigenvalues of greater than 1 (see Table X) which were used in subsequent analyses.

2. Auditory abilities and speech perception

Table XI displays the correlations between the first three principal components derived from the auditory measures, and measures of speech perceptual performance using the SiVo aid and amplified speech. Results from CDT and from the tests using compound speech pattern stimuli were not included since these were from only five subjects.

TABLE VIII. Gap detection thresholds in ms. The noise level was measured as the peak spectrum level of the noise, at about 350 Hz.

Subject	Noise level (dB/Hz)	Gap threshold (ms)
S1	91	41.4
S2	77	20.0
S3	92	21.4
S4	95	24.1
S5	75	21.3
S6	93	23.0
S11	52	23.6
S12	89	26.1
S13	89	38.9

Of the 91 correlations that were computed, there were nine significant correlations ($p < 0.05$) with the principal components. Since between four and five significant correlations would be expected by chance, it is likely that there are genuine relationships between auditory measures and the speech data, although some of these nine significant correlations may be spurious.

The first auditory principal component was significantly negatively correlated with performance in the question/statement task both with the SiVo aid and with amplified speech, and, in addition, was significantly positively correlated with the difference between the extraction of consonantal manner information with the SiVo aid and with amplified speech. This first component, which accounted for over 50% of the overall variance among the auditory measures, was highly weighted on hearing loss and loss of frequency selectivity; subjects with low scores for this component, that is, with less extreme auditory impairment, were better able to identify intonation contours and to extract manner information from amplified speech.

The second principal component was significantly negatively correlated with overall performance in consonant identification using the SiVo aid, with the extraction of consonantal voicing information using the SiVo aid, and also with the advantage of lipreading with the SiVo aid over unaided lipreading for these same two scores. The second component, which accounted for 18% of the variance among the auditory measures, was related to poorer 250-Hz frequency discrimination, and smaller 3FA dynamic range. The 250-Hz frequency DL, which relates to the fundamental frequency range of the female speech used in testing, is presumably based upon temporal processing ability, which must necessarily be the perceptual basis of the extraction of voicing contrasts using the SiVo aid.

The third principal component was correlated only with the extraction of consonantal voicing and place information in unaided lipreading, a relationship that can have no direct auditory basis.

3. Aid preferences

The two subjects who gained most substantially from the SiVo aid, and who now use it to the exclusion of amplifying aids, showed 3FA losses of 114 dB (S1) and 112 dB (S4). S2, S3, and S13, who also continue use of the SiVo aid, had similar 3FA losses of 112, 116, and 113 dB, respectively. S12, who had a similar 3FA loss of 114 dB, rejected the SiVo aid on cosmetic grounds, but in the later tests (Sec. IV C), she was performing at least as well with the same information over headphones as with her preferred low-pass hearing aid. Subjects with 3FA losses of 110 dB or less preferred amplifying aids to the SiVo aid. Those subjects with LFA losses of 95 dB or more (S8, S9, and S10) were unable to effectively use the SiVo aid.

4. Summary

Three indicators are likely to be useful in predicting which subjects will benefit from simplified voice fundamental frequency presentation as opposed to whole speech presentation.

TABLE IX. Correlation matrix for audiometric and psychoacoustic data. Correlations significant at $p < 0.05$ ($n = 11$) are underlined. LFAHL: average hearing loss at 125 and 250 Hz; 3FAHL: average hearing loss at 500, 1000, and 2000 Hz; 4FAHL: average hearing loss at 500, 1000, 2000, and 4000 Hz; LFADR: average dynamic range at 125 and 250 Hz; 3FADR: average dynamic range at 500, 1000, and 2000 Hz; DLFL: frequency DL at 125 Hz (100 Hz for S1); DLFH: frequency DL at 250 Hz (240 Hz for S1, 200 Hz for S2); MLD: masking level difference (Table VIII); GAP: gap threshold.

	3FAHL	4FAHL	LFADR	3FADR	DLFL	DLFH	MLD	GAP
LFAHL	0.782	0.755	− 0.907	0.406	0.111	0.296	− 0.721	0.417
3FAHL		0.945	− 0.693	− 0.056	0.263	0.508	− 0.622	0.467
4FAHL			− 0.745	0.102	0.324	0.611	− 0.747	0.560
LFADR				− 0.528	− 0.197	− 0.100	0.814	− 0.523
3FADR					0.238	− 0.178	− 0.203	0.180
DLFL						0.064	− 0.251	0.565
DLFH							− 0.131	− 0.063
MLD								− 0.712

(1) Extreme audiometric loss and the associated lack of frequency selectivity in the voice fundamental frequency region impairs the ability to extract fundamental frequency information, consonantal manner, and vowel quality information from speech. Subjects with little or no selectivity lose no manner information when speech is replaced by a sinusoidal voice fundamental frequency signal, and can gain considerably in the improved perception of voicing and intonation.

(2) Since low-frequency hearing loss and dynamic range are highly correlated with low-frequency auditory frequency selectivity, these two readily obtained measures, together with a measure of average dynamic range between 500 and 2000 Hz, may well be among the most useful clinical indicators of the appropriateness of a prosthesis such as the SiVo aid for an individual subject.

(3) In addition to those factors that emerge from the multivariate analysis, it is apparent that poor frequency discrimination in the upper part of the voice fundamental frequency range favors the use of frequency mapping to match speech information from speakers having a high voice fundamental frequency range.

At least in these subjects, these factors were observed in subjects with 3FA hearing losses of more than 110 dB. One factor that militates against a benefit from the SiVo aid is a high LFA hearing loss; > 95 dB is the best estimate of this limit that can be made from the present data.

TABLE X. Principal components analysis of auditory receptive measures. The table shows the eigenvalues for the first four components, the proportion of the total variance explained by each component, and the eigenvectors for the three components with eigenvalues > 1, which each account for at least 10% of the variance, and were selected for the subsequent analysis.

	Auditory component			
	1	2	3	4
Eigenvalue	4.81	1.65	1.10	0.81
Proportion of variance	0.53	0.18	0.12	0.09
LFAHL	0.41	0.06	− 0.33	
3FAHL	0.40	0.29	0.06	
4FAHL	0.43	0.22	0.08	
LFADR	− 0.40	0.17	0.34	
3FADR	0.09	− 0.55	− 0.31	
DLFL	0.18	− 0.30	0.71	
DLFH	0.17	0.55	0.15	
MLD	− 0.39	0.15	0.09	
GAP	0.31	− 0.33	0.38	

TABLE XI. Correlations between auditory principal components and speech perception measures. Significant correlations ($n = 9$) are underlined.

	Auditory component		
Speech measure	1	2	3
Question/statement			
Sx	− 0.6473	− 0.3727	− 0.0587
Sp	− 0.6499	− 0.1930	0.0673
Sx–Sp	0.4867	− 0.1027	− 0.2226
Consonant lipreading:			
SiVo Aid			
Overall, *Sx*	− 0.0140	− 0.7164	− 0.1216
Voicing, *Sx*	− 0.3685	− 0.6646	− 0.1612
Manner, *Sx*	0.2538	− 0.1114	− 0.4331
Place, *Sx*	0.5207	− 0.1703	0.0438
Amplified speech			
Overall, *Sp*	− 0.2718	− 0.1628	0.0327
Voicing, *Sp*	− 0.3466	− 0.2477	0.0855
Manner, *Sp*	− 0.2753	− 0.1815	− 0.2103
Place: *Sp*	0.1921	0.1891	0.2158
Unaided lipreading			
Overall: *LA*	0.2478	0.4035	0.3933
Voicing: *LA*	0.3208	0.1178	0.6686
Manner: *LA*	− 0.1321	0.3201	− 0.2262
Place: *LA*	0.1035	− 0.2293	0.6286
Differences between conditions			
Overall: *Sx–Sp*	0.3564	− 0.3052	− 0.1333
Overall: *Sx–LA*	− 0.1044	− 0.7726	− 0.2520
Overall: *Sp–LA*	− 0.3514	− 0.2738	− 0.0626
Voicing: *Sx–Sp*	0.0732	− 0.3330	− 0.2602
Voicing: *Sx–LA*	− 0.3873	− 0.6534	− 0.2212
Voicing: *Sp–LA*	− 0.3783	− 0.2614	0.0333
Manner: *Sx–Sp*	0.6612	0.1279	− 0.1870
Manner: *Sx–LA*	0.3261	− 0.2370	− 0.3946
Manner: *Sp–LA*	− 0.2381	− 0.2619	− 0.1496
Place: *Sx–Sp*	0.5240	− 0.3262	− 0.0787
Place: *Sx–LA*	0.4864	− 0.0827	− 0.2014
Place: *Sp–LA*	0.1177	0.3861	− 0.2856

VI. GENERAL DISCUSSION AND CONCLUSIONS

A. Overview of results

Speech pattern hearing aids appear to offer a significant potential benefit for some profoundly hearing-impaired listeners. Of 11 subjects studied here, five continue to use the present simple SiVo aid. The simplification of speech to fundamental frequency was found to significantly assist the profoundly hearing impaired in the reception of intonation, as Rosen and Fourcin (1983) and Grant (1987a) have previously shown. The benefits of simplification were especially apparent in subjects who lacked measurable auditory frequency selectivity. For three of the subjects tested, the extreme simplification of speech represented by the SiVo aid led to a significant advantage in the reception of voicing information in audio–visual consonant identification.

Consonant identification and CDT studies with additional speech pattern elements demonstrated significant advantages of adding amplitude information or both amplitude and voiceless information to voice fundamental frequency. Matched compound speech pattern information produced significant advantages over amplified speech for three of five subjects in both speech tasks, while the remaining subjects showed no disadvantage.

B. Numbers who may benefit from the SiVo aid and subsequent developments

The five subjects who continue to use the SiVo aid show 4FA hearing losses of 115 dB or more. If wearable aids were available which also coded the further speech pattern elements examined here, the potential population which may benefit from speech pattern presentation would likely be bounded by a rather lower 4FA loss of perhaps 110 dB. The best available estimates (Thornton, 1986) indicate that a 4FA hearing loss of 110 dB is found in approximately 0.03% of the population, that is, about 17 500 individuals in the UK, and about 90 000 in the USA and Canada. The population is also likely to be bounded by a LFA hearing loss of 95 dB or less, but no suitable estimate of the combined prevalence of low- and high-frequency losses is available.

C. Developments of the SiVo aid

To be generally acceptable, a speech pattern extracting aid must be robust in everyday noise and reverberant conditions. To this end, we are now implementing neural-net methods of fundamental frequency extraction which have been shown to be robust at poor signal-to-noise ratios (Howard and Huckvale, 1988a,b; Walliker and Howard, 1990). The greatest potential advantage of the speech pattern approach is that, in noisy or reverberant environments, robust pattern extraction methods such as these are likely to exceed the capacities of the profoundly hearing impaired to extract speech pattern contrasts from amplified whole speech (Faulkner *et al.*, 1991). Such developments would be likely to further enlarge the population who may benefit.

ACKNOWLEDGMENTS

We are grateful to Danavox A/S for supplying and modifying the hearing aids used as reference aids and to John Walliker for scientific and technical support. We also thank David Smith for his assistance in the collection of gap detection thresholds and for his software skills, and Alfred Ng, for the analysis of the voicing and voiceless timing patterns of the /ɑˈCɑ/ stimuli. We are grateful also to reviewers Quentin Summerfield and Ken Grant for their constructive comments on an earlier version of this paper. Finally, we thank the patients, whose tolerance and good-natured persistence contributed enormously to this work. This work was supported by the Medical Research Council of Great Britain.

APPENDIX

The following table is a summary of tests and test results. The compound speech pattern results are not included.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	Key
3 frequency average HL	114	112	116	112	105	107	110	113	123	117	98	114	113	dB
3 frequency average dynamic range	8	0	5	4	10	17	···	15	4	8	9	0	10	dB
Average HL at 125 and 250 Hz	94	49	92	86	75	86	70	104	102	95	53	74	82	dB
Average dynamic range at 125 and 250 Hz	10	42	15	17	21	22	···	8	9	11	35	33	18	dB
Frequency difference limen at 125 Hz	4.3	3.2	3.8	7.5	5.9	8.2	···	···	···	···	5.5	9.0	41.9	%
Frequency difference limen at 250 Hz	5.7	3.5	6.7	11.9	9.9	10.2	···	···	···	···	2.2	18.3	7.5	%
Psychoacoustic tuning curve: Masking level difference	− 1.0	14.0	8.0	0.5	10.5	7.5	···	···	···	···	10.5	9.5	4.0	dB
Gap threshold	41.4	20.0	21.4	24.1	21.3	23.0	···	···	···	···	23.6	26.1	38.9	ms
Consonants: Lips only	39.6	36.5	43.2	43.8	40.3	42.2	50.3	34.0	···	···	44.6	50.0	44.3	% correct
Consonants: Lips + Speech	49.0	57.3	82.3	54.7	91.7	86.5	67.6	65.6	···	···	89.6	78.1	81.3	% correct
Consonants: Lips + SiVo	69.8	66.7	80.2	62.0	80.2	80.2	54.3	53.6	···	···	79.2	60.4	83.3	% correct
Question/statement: Speech	56	88	92	47	98	61	···	47	···	···	100	68	78	% correct
Question/statement: SiVo	82	100	95	64	98	81	···	66	···	···	100	75	86	% correct
CDT: Lips only	26.7	55.4	57.3	73.3	···	···	···	···	···	···	31.6	43.6	···	words/min
CDT: Lips + Speech	37	78.7	77.7	70.9	···	···	···	···	···	···	70.1	60.7	···	words/min
CDT: Lips + SiVo	44	73.8	74.6	68.4	···	···	···	···	···	···	68.0	52.3	···	words/min
Amplifying aid usage	N	D	D	N	D	D	S	D	N	D	D	D	N	N = None S = Some
SiVo Aid usage	D	D	D	D	N	N	N	N	N	N	S	N	S	D = Daily

[1] Other signals of fixed intensity but with a complex spectrum, such as a pulse train, are likely to be less effective than sinusoids. The perception of pitch in these listeners is more sensitive than normal to the relative phases of the acoustic components of the signal (Rosen *et al.*, 1987), and because the effective relative phases of the components will depend upon the unknown frequency-dependent mechanical characteristics of the patient's ear, changes in signal frequency are likely to be accompanied by changes in the phase spectrum, and hence, of the temporal structure, of the signal driving the auditory nerve.

[2] In some countries, many of these patients would be considered candidates for cochlear implantation. In the UK, however, at the time the study was performed, only totally deaf patients were accepted for cochlear implantation.

[3] In Rosen *et al.* (1987), S1 was referred to as C, and S2 as M. In addition, Faulkner *et al.* (1989), and Faulkner *et al.* (1990b), used the following designations: S3–T; S4–A; S5–L; S6–S; S7–LL; S8–G; S12–LG.

[4] S10 was able to make use of his UK National Health Service BW81 aid, which had a maximum output level 2 to 3 dB higher than the unmodified 107-2 aid.

[5] The reference aid has proved a good choice. Five patients out of the eight in whom the aid was tried were using another aid at the start of the tests; only one of these five preferred his existing aid to the reference aid, and that patient (S10) proved to have such a profound audiometric loss that he was barely able to hear the reference aid. Quantitative assessments of aided lipreading of intervocalic consonants were carried out with two patients (S5 and S6) comparing their own postaural aids with the reference aid. In both cases, the results obtained using the reference aid were marginally better than with the patients' previous aids.

[6] This speaker approximates an ideal point source 30 cm behind the diaphragm, and has an excellent low-frequency amplitude and phase response extending to below 80 Hz. These characteristics are all important in the reproduction of a speech pressure waveform which closely resembles that from a live speaker. Phase distortion could have impaired the performance of the SiVo aid's fundamental frequency extractor, and could also have impaired the subjects' ability to extract pitch information from speech (Rosen *et al.*, 1987). The video recording systems used for the recording and copying of the test materials did not impose significant magnitude or phase response errors within the frequency range 80–5000 Hz.

[7] The performance of the peak-picking fundamental frequency extractor in the SiVo aid is affected by reverberation, and hence, by source-to-microphone distance. Had the SiVo aid been used at the same source-to-microphone distance as the reference amplifying aid, a marginal loss of accuracy in fundamental frequency representation would have been present, but this would not be likely to be large enough to affect the representation of the voicing pattern

[8] This result is likely to have been affected by this subject's rejection of the SiVo aid for cosmetic reasons. Results obtained later using the same signal but headphone presentation did not show poorer voicing reception with *Sx* than with speech (Fig. 5).

[9] The 2IFC thresholds obtained during the masking experiments were an average of 5.3 dB below the method of adjustment thresholds given (in dB HL) in Fig 2.

Abberton, E., and Fourcin, A. J. (1978). "Intonation and speaker identification," Lang. Speech **21**, 305–318.

Ball, V. (1991). "Computer-based tools for assessment and remediation of speech," Br. J. Dis. Commun. **26**, 95–113.

Ball, V., Faulkner, A., and Fourcin, A. J. (1990). "The effects of two different speech coding strategies on voice fundamental frequency control in deafened adults," Br. J. Audiol. **24**, 393–409.

Bootle, C. M. (1986). "Computer graphics for speech therapy," Speech Hear. Lang. **2** (University College London, Dept. of Phonetics and Linguistics), 47–55.

Breeuwer, M., and Plomp, R. (1985). "Speechreading supplemented with auditorily presented speech parameters," J. Acoust. Soc. Am. **79**, 481–499.

DeFilippo, C. L., and Scott, B. L. (1978). "A method for training and evaluation of the reception of on-going speech," J. Acoust. Soc. Am. **63**, 1186–1192.

Faulkner, A., Potter, C., Ball, V., and Rosen, S. (1989). "Audio-visual speech perception of intervocalic consonants with auditory voicing and voiced/voiceless speech pattern presentations," Speech Hear. Lang. **3**, (University College London, Dept. Phonetics and Linguistics), 87–106.

Faulkner, A., Ball, V., and Fourcin, A. J. (1990a). "Compound speech pattern information as an aid to lipreading," Speech Hear. Lang. **4** (University College London, Dept. Phonetics and Linguistics), pp. 63–80.

Faulkner, A., Fourcin, A. J., and Moore, B. C. J. (1990b). "Psychoacoustic aspects of speech pattern coding for the deaf," Acta Otolaryngol. (Stockh) **Suppl. 469**, 172–180.

Faulkner, A., Rosen, S., and Moore, B. C. J. (1990c). "Residual frequency selectivity in the profoundly hearing-impaired listener," Br. J. Audiol. **24**, 381–392.

Faulkner, A., Walliker, J. R., Howard, I. S., Ball, V., and Fourcin, A. J (1992) "New developments in speech pattern element hearing aids for the profoundly deaf," Acta Otolaryngol. (Stockh) (to be published).

Fourcin, A. J. (1977). "English speech patterns with special reference to artificial auditory stimulation," in *A Review of Artificial Auditory Stimulation: Medical Research Council Working Group Report*, edited by A. R. D. Thornton (Institute of Sound and Vibration Research, University of Southampton), pp. 42–44.

Fourcin, A. J. (1990). "Prospects for speech pattern element aids," Acta Otolaryngol. (Stockh), **Suppl. 469**, 257–267.

Fourcin, A. J., and Abberton, E. (1971). "First applications of a new laryngograph," Med. Biol. Illus. **21**, 172–182.

Fourcin, A. J., Moore, B. C. J., and Douek, E. E. (1977). "Perception and presentation of artificial speech patterns," in *A Review of Artificial Auditory Stimulation: Medical Research Council Working Group Report*, edited by A. R. D. Thornton (Institute of Sound and Vibration Research, University of Southampton), pp. 24–30.

Fourcin, A. J., Rosen, S. M., Moore, B. C. J., Douek, E. E., Clarke, G. P., Dodson, H., and Bannister, L. H. (1979). "External electrical stimulation of the cochlea: Clinical, psychophysical, speech-perceptual and histological findings," Br. J. Audiol. **13**, 85–107.

Fourcin, A. J., Douek, E. E., Moore, B. C. J., Abberton, E., Rosen, S. M., and Walliker, J. R. (1984). "Speech pattern element stimulation in electrical hearing," Arch. Otolaryngol. **110**, 145–153.

Grant, K. W. (1987a). "Encoding voice pitch for profoundly hearing-impaired listeners," J. Acoust. Soc. Am. **82**, 423–432.

Grant, K. W. (1987b). "Identification of intonation contours by normally hearing and profoundly-impaired listeners," J. Acoust. Soc. Am. **82**, 1172–1178.

Grant, K. W. (1987c). "Frequency modulation detection by normally hearing and profoundly hearing-impaired listeners," J. Speech Hear. Res. **30**, 558–563.

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

Howard, D. M. (1989). "Peak-picking fundamental period estimation for hearing prostheses," J. Acoust. Soc. Am. **86**, 902–910.

Howard, D. M., and Fourcin, A. J. (1983). "Instantaneous voice period measurement for cochlear stimulation," Electron. Lett. **19**, 76–78.

Howard, I. S., and Huckvale, M. A. (1988a). "Speech fundamental period estimation using a trainable pattern classifier," in *Proceedings of Speech '88: 7th FASE Symposium* (Institute of Acoustics, Edinburgh), pp. 129–136.

Howard, I. S., and Huckvale, M. A. (1988b). "Training feature detectors for use in automatic speech recognition," in *Proceedings of Speech '88: 7th FASE Symposium* (Institute of Acoustics, Edinburgh), pp. 1365–1372.

Miller, G. A., and Nicely, P. E. (1955). "An analysis of perceptual confusions among some English consonants," J. Acoust. Soc. Am. **27**, 338–352.

Milne, J (1977). *Heineman Guided Readers Handbook* (Heineman, London).

Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory filter-bandwidths and excitation patterns," J. Acoust. Soc. Am. **74**, 750–753.

Moore, B. C. J., and Glasberg, B. R. (1986). "Comparisons of frequency selectivity in simultaneous and forward masking for subjects with unilateral hearing impairments," J. Acoust. Soc. Am. **80**, 93–107.

Nordmark, J. O. (1968). "Mechanisms of frequency discrimination," J. Acoust. Soc. Am. **44**, 1533–1540.

Patterson, R. D., and Moore, B. C. J. (1986). "Auditory filters and excitation patterns," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 123–178.

Rosen, S., Faulkner, A., and Smith, D. A. J. (1990). "The psychoacoustics of profound hearing impairment," Acta Otolaryngol. (Stockh) **Suppl. 469**, 16–22.

Rosen, S., and Fourcin, A. J. (1983). "When less is more: further work," Speech Hear. Lang. 1 (University College London, Dept. Phonetics and Linguistics), 3–27.
Rosen, S., and Fourcin, A. J. (1986). "Frequency selectivity and the perception of speech," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 373–487.
Rosen, S., Fourcin, A. J, Abberton, E., Walliker, J. R., Howard, D. M., Moore, B. C. J., Douek, E. E., and Frampton, S. (1985). "Assessing assessment," in *Cochlear Implants*, edited by R. A. Schindler and M. M. Merzenich (Raven, New York), pp. 479–498.
Rosen, S., Fourcin, A. J., and Moore, B. C. J. (1981). "Voice pitch as an aid to lipreading," Nature **291**, 150–152.
Rosen, S., Moore, B. C. J., and Fourcin A. J. (1979). "Lipreading with fundamental frequency information," Proceedings of the Institute of Acoustics Autumn Conference, Windermere 1979, paper 1A2, pp. 5–8.
Rosen, S., and Nevard, S. (1987). "A headphone monitoring system for low frequency psychoacoustics," Br. J. Audiol. **21**, 108–109 (Abstract).
Rosen, S., Walliker, J. R., Fourcin, A. J., and Ball, V. (1987). "A microprocessor-based acoustic hearing aid for the profoundly impaired listener," J. Rehabil. Res. Devel. **24**, 239–260.
SAS Institute Inc. (1987). *SAS/STAT™ guide for personal computers, Version 6 Edition* (SAS Institute Inc., Cary, NC).
Shailer, M. J., and Moore, B. C. J. (1983). "Gap detection as a function of frequency and bandwidth," J. Acoust. Soc. Am. **74**, 467–473.
Stock, D., and Rosen, S. (1986). "Frequency discrimination and resolution at low frequencies in normal and hearing-impaired listeners," Speech Hear. Lang. 2, (University College London, Dept. Phonetics and Linguistics), pp. 193–222.
Summerfield, A. Q. (1983). "Audiovisual speech perception, lipreading and artificial stimulation," in *Hearing Science and Hearing Disorders*, edited by M. E. Lutman and M. P. Haggard (Academic, London), pp. 131–182.
Thornton, A. R. D. (1986). "Estimation of the number of patients who might be suitable for cochlear implant and similar procedures," Br. J. Audiol. **20**, 221–230.
Tyler, R. S. (1986). "Frequency resolution in hearing-impaired listeners," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 373–487.
Walliker, J. R., and Howard, I. S. (1990). "The implementation of a real time speech fundamental period algorithm using multi-layer perceptrons," Speech Commun **9**, 63–71.
Weber, D. L. (1977). "Growth of masking and the auditory filter," J. Acoust. Soc. Am. **62**, 424–429.
Wier, C. C., Jesteadt, W., and Green, D. M. (1977). "Frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **61**, 178–184.