The development of the speech discrimination test described below must be regarded as part of a larger framework of a test series "Diagnostic Investigation of Language Acquisition" which will appear at a later stage. The need was felt to have a test at our disposal with which a quick screening could be carried out concerning the understanding of speech. The following investigation must not be taken other than an account of the work done to get more experience of one aspect of hearing. Naturally a definite pronouncement at this stage of this specific investigation would be premature.

Speech Discrimination

Development of a Test

by M.C. Dinger.

1.0 Introduction.

1.1 In present-day education an increasing use is made of language laboratories. But such aspects as adaptation of the method to the participants and fitness or aptitude of the participants to such methods of education have been little investigated. In view of this a plan arose in 1968 to develop a series of tests which were to give an insight into the characteristics of students (van Herpt, 1973). With the help of the results of these tests, combined with other data as, for instance, study results, it should be possible to make use of language laboratories more effectively.

2.0 Purpose of the investigation.

2.1 Within the framework of the diagnostic investigation concerning language acquisition*), now in progress at the Institute of Phonetic

*) This investigation is described in: 'Psychologie per computer - een pilot-investigation, Diagnostisch Onderzoek Taalverwerving'. (Psychology per Computer, a Pilot Investigation, Diagnostic Investigation Language Acquisition); by L.W.A. van Herpt, Institute of Phonetic Sciences, 1973, publication 40, Amsterdam.

Sciences, it was of great importance to develop a test, with which the hearing capacity of students making use of language laboratories can be determined. A second motive for designing such a test was the point that many investigations carried out at the Institute of Phonetic Sciences incorporate listening experiments which, in view of the validity of the results have to make use of subjects whose hearing is unimpaired. In either case the stress falls on the capacity of hearing speech-sounds. Therefore no use was made of pure tones as testing material. The aim of the development of this part of the investigation was designing a speech discrimination test which was to satisfy the following conditions:

1. The test should be made so that groups of subjects could be tested at the same time
2. The results were to be automatically processed.

3.0 Design.

3.1 Two batches of forty one-syllable tape-recorded words to be presented via head phones to the testees.
The test to consist of two parts.
Part I 40 meaningful monosyllables, divided into 4 groups of 10 monosyllables
Part II 40 meaningless monosyllables, divided into 4 groups of 10 monosyllables.
The time interval between monosyllables: 4 seconds, between groups: 12 seconds and between Parts I and II: 24 seconds. The volume varied per group of 10 items. The level of loudness for Parts I and II run parallel and are chosen in such a way that people with normal hearing will have a 100% correct score for the first of 10 items, see 6.3, (i.e. monosyllables presented loudest). The lowest level of loudness will be such that even a person with normal hearing will not understand all the words correctly, or at any rate, will do so with great difficulty.

3.2 The responses are multiple choice. For each stimulus the listener has a choice of three, which are presented on forms for automatic scoring. For these automatic scores use is made of an adapted Standard IBM 557 form. The order in which the three response possibilities are placed on the scoring forms is fully randomized.

The procedure during the test and the scoring instructions are given on printed forms.

4.0 Definition of the Problem.

The data resultingfrom the test would have to decide the possibilities and the usefulness of the discrimination test in the shape presented here. Our main interest centered on the following points.

4.1.0 The usefulness of the speech material and the alternatives given.
a) The measure of discrimination resulting from the 80 stimuli offered.
b) The measure of equality of the alternatives.

4.2.0 Delimination of the level of loudness where an optimum of discrimination occurred between testees.

4.3.0 Deciding the degree and the type of loss of hearing.
As a result of the wrong responses an evaluation would be attempted of the degree of loss of hearing and the type of this loss.
a). The degree of loss of hearing could be based on the level of loudness with which a wrongly scored response was presented
b). The type of loss of hearing could be based on the choice of the alternatives of the stimulus presented. (see 5.3.2).

4.3.1 Assessment of the validity of the test.
A comparative examination of hearing would be necessary in the shape of a tone audiogram. A number of testees with the largest and with the smallest number of mistakes, and a number from the median group would receive a request to come and have their hearing tested individually.

4.4.0 Testing of the following null hypothesis.
H_o : A speech discriminationtest with meaningful test material and a test with meaningless material presented under identical conditions show no significant difference.
Data presented in literature on the subject suggest that even a person with normal hearing always shows decreased discrimination when the stimulus material consists of nonsense-syllables (logotomes).

5.0 Material.

5.1 Part I

The starting point for the meaningful speech material was the list used at the Professor H. Burger School in Amsterdam, the so called P.H.B. list. From it those monosyllables were chosen which produced the most favourable alternatives. (see table I and II).
The right choice of alternatives is an essential part of this investigation as the investigation is conducted on the basis of forced choice scoring.

5.2 Part II

The material for the second part of the investigation was designed in collaboration with Miss J.M. van der Stelt, also of the Institute of Phonetic Sciences. It consists of meaningless monosyllables. These have been constructed in such a way that the same vowels have been used equally often in Parts I and II. In this way an attempt was made to keep both parts as equal as possible. (see tables I and II).

5.3.0 The Alternatives.

The alternatives for both parts have been formed by making use of the stimulus as before. (see tables I and II).

1. The alternatives have been formed by changing one phoneme in comparison to the stimulus offered.
2. This phoneme was chosen in such a way that an endeavour could be made to establish the type of deafness (low- or high-tone deafness) from the analysis of the types of mistakes made by the subject. For instance in Part I an item was subjected for judgment: bier [bir]. The testee could choose between: mier- bier- buur. In transcription [mir-bir-byr].

The mixing up of [i] / [y] might point to high-tone deafness; the mixing up of [m] / [b] might point to low-tone deafness.

5.3.1 The number of vowel changes and of consonant changes was kept as uniform as possible for each of the four groups. Within the group more consonant changes had to be made than vowel changes.

5.3.2 Kruizinga's *) "confusion tables" were used for the alteration of

*) Kruizinga J.H. (1955). 'Slechthorendheid en het verstaan van spraak', thesis, Excelsior, 's-Gravenhage.

consonants. For the modification of vowels Hellwag's [*2)] vowel triangle served as a point of departure. The duration of the vowel in the stimuli was also taken into account when the choice of alternatives was made.

5.3.3 The following qualifications held throughout the procedure of Part I: the alternatives had to be words in every-day use; and of Part II: the alternatives had to be absolutely meaningless.

6.0 Technical Procedure.

6.1 Tape - recordings.

The speech material was recorded at the Institute of Phonetic Sciences on an Ampex 300 recorder. The technical side of the investigation was supervised by E.O. Kappner.

To try and put the stimuli, which had to have four different sound levels, directly on tape with the right sound-level ratio seemed inadvisable as the signal to noise ratio for the lowest levels would have been very unfavourable owing to tape noise on the tape. Therefore all words were recorded at the same level and while conducting the listening test the play-back signal of the recorder was reduced with the aid of a dB-attenuator to the required level.

In this way the noise on the tape was reduced as well. The recordings were made with low-noise tape (Scotch Tape 201), the copies with low-print (Scotch Tape 138).

The requirements of the recordings were as follows:

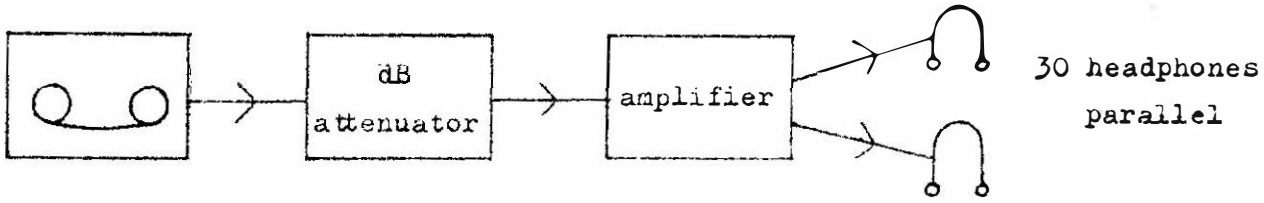
1. All words had to be pronounced clearly.
2. All words had to have the same pitch and intonation (as much as possible, anyhow). This meant retakes were necessary part of the time
3. The same loudness for the different words had to be maintained.

The loudness of the different monosyllables, spoken by a female voice, was checked by ear as well as possible. The use of a sound-level meter is of no avail here as it does not measure loudness.

*2) Hellwag, C.F. (1781). 'De Formatione Loquelae'., dissertation, Tübingen.

6.2 Apparatus.

The Institute of Applied Linguistics of Amsterdam had offered the use of its language laboratory for the test. Here use was made of the available Tandberg Model 4 language laboratory outfit. The electronics engineer A.G. Wempe tested all headphones for this test on differences in sensitivity. These appeared to be less than 2 dB for the required frequencies. In order to avoid hum and noise in the long wire connections from the central teaching panel to the student booths it was necessary to amplify the signal at the central teaching panel and to feed the amplified signal directly into the headphones. The noise of the recorder amplifier was eliminated in this way as well. All headphones therefore had parallel connections with the central teaching panel and the recorders in the listening booths were turned off.
The interconnections of the apparatus were as follows



6.3 Loudness levels.

An essential point of investigation was to find out at which loudness level an optimal discrimination between testees became apparent. In 1969 a pilot investigation was conducted with a group of first-year Arts students. The results of this test showed that 3 output levels (items 1 - 30 inclusive) were too high. Although some mistakes were made in these blocks only the block with the lowest loudness level (items 31 - 40 inclusive) showed a discrimination between subjects. On account of this pilot investigation the 4 loudness levels of the speech discrimination test were chosen as follows:
As starting point a level was chosen which could be clearly understood (first 10 items). The loudness of the following three groups was diminished by 5, 5, and 10 dB respectively. The first word vaas [vas] of the first group had an output level of 55 dB, measured directly at the headphones with a Peekel Sound spectrometer type G.R.B.
The following observation seems called for here. Suppose that the loudness is experienced as being the same for all words when played without attenuation. It does not follow, however, that when the signal is attenuated with e.g. 35 dB this equality of loudness is maintained. After all loudness is a subjective experience. The same problem occurred as well when the recordings were made as a result of the

curvature of the Fletcher-Munson curve.
For this test this problem did not really matter as some-one with impaired hearing has the same trouble when listening to spoken signals. (speech).

7.0 Execution.

7.1 In October 1970 and 1971 the speech discrimination test was carried out with groups of 47 and 46 subjects respectively. In 1970 at the Institute of Applied Linguistics and in 1971 in the listening booths at the Institute of Phonetic Sciences in Amsterdam. The testees were all students of the day course for Speech Therapists in Amsterdam.

7.2 The loudness levels of the different groups of items were as follows. In 1970 the attenuation for Part I had been 0, 5, 5, and 10 dB per 10 stimuli respectively (see 6.3) and for Part II it was 0, 0, 5, and 10 dB*). In 1971 attenuation for Parts I and II was 0, 0, 5, and 10 dB.

7.3 The speech discrimination test lasted 7 minutes and was combined with the Seashore Measures of Musical Talents-test, which was also part of the Diagnostic Investigation of Language Acquisition.

7.4 In 1970 an air_conduction audiogram was made; in 1971 the six subjects with the lowest score and the four subjects with the highest score in the speech discrimination test underwent an audiometer test. (A large group of subjects follow the four "best" subjects, with the same number of correct scores.) The audiograms were made with a Peekel screening-audiometer type D 66 / 6936.

8.0 Results.

An outline follows below of the scores on items and alternatives with corresponding P-values and item-testcorrelation (r_{it}), which is presented in Table I for Subtest "meaningful items" (1970 and 1971) see page 104 and in Table II for Subtest "meaningless items" (1970 and 1971) see page 105
A phonetic transcription can be found on page 106.

*) The difference in loudness levels between items 11 - 20 inclusive of Part I when compared to Part II was the result of a mistake.

8.1 1970 Part I Subtest "Meaningful Items".

Total number of scores 46*) x 39**)	= 1794
Number of wrong scores	234
No scores	32
Number of wrong scores in 3rd block	42
Number of wrong scores in 4th block	132

1970 Part II Subtest "Meaningless Items".

Total number of scores 47 x 40	= 1880
Number of wrong scores	308
No scores	13
Number of wrong scores in 3rd block	78
Number of wrong scores in 4th block	179

1971 Part I Subtest "Meaningful Items".

Total number of scores 46 x 40	= 1840
Number of wrong scores	118
No scores	4
Number of wrong scores in 3rd block	30
Number of wrong scores in 4th block	43

1971 Part II Subtest "Meaningless Items".

Total number of scores 46 x 40	= 1840
Number of wrong scores	145
No scores	6
Number of wrong scores in 3rd block	39
Number of wrong scores in 4th block	81

*) 1 subject dropped out of Part I.
**) 1 item dropped out of Part I owing to a technical defect.

Table I: Subtest "Meaningful Items".**)

A: scores in 1970 - 46 testees

items 1 t/m 10 : 30 dB attenuation
items 11 t/m 20 : 35 dB "
items 20 t/m 30 : 35 dB "
items 30 t/m 40 : 40 dB "

B: scores in 1971 - 46 testees

items 1 t/m 10 : 40 dB attenuation
items 11 t/m 20 : 40 dB "
items 20 t/m 30 : 45 dB "
items 30 t/m 40 : 45 dB "

P = relative frequency of correct answers.

R_{it} = item-testcorrelation.

Stimulus	A	B	Alternative	A	B	Alternative	A	B	No score A	No score B	P A	P B	R_{it} A	R_{it} B
1. vaas	-*)	43	waas	-	3	vaag	-	0			*	0.93	0.00	0.00
2. bijt	43	46	buit	0	0	meid	1	0	2		0.93	1.00	0.40	0.00
3. hek	45	46	hok	0	0	hak	0	0	1		0.97	1.00	0.39	0.00
4. fout	44	46	goud	1	0	vat	0	0	1		0.95	1.00	0.61	0.00
5. wang	46	46	zang	0	0	hang	0	0			1.00	1.00	0.00	0.00
6. boos	45	46	boog	0	0	poos	1	0			0.97	1.00	0.07	0.00
7. bus	46	46	bos	0	0	mus	0	0			1.00	1.00	0.00	0.00
8. wol	46	46	vol	0	0	wal	0	0			1.00	1.00	0.00	0.00
9. boel	32	34	poel	14	12	doel	0	0			0.69	0.73	0.48	0.59
10. zing	36	36	zien	10	10	ving	0	0			0.78	0.78	0.14	0.23
11. vuur	35	44	vier	9	0	zuur	1	2	5		0.76	0.95	0.47	-0.07
12. maal	42	46	maai	2	0	maan	0	0	2		0.91	1.00	0.41	0.00
13. buit	45	44	bout	0	0	buik	1	2			0.97	0.95	0.14	0.42
14. bier	43	43	mier	3	1	buur	0	2			0.93	0.93	0.42	0.23
15. been	28	38	boon	1	0	peen	17	8			0.60	0.82	0.69	0.26
16. peul	44	46	pool	1	0	beul	0	0	1		0.95	1.00	0.46	0.00
17. bel	46	46	bal	0	0	bol	0	0			1.00	1.00	0.00	0.00
18. hak	39	46	jak	0	0	hap	6	0	1		0.84	1.00	0.50	0.00
19. dik	42	45	dit	2	1	tik	1	0	1		0.91	0.97	0.39	0.18
20. boot	33	38	poot	13	8	dood	0	0			0.71	0.82	0.16	0.47
21. mus	46	46	mug	0	0	mis	0	0			1.00	1.00	0.00	0.00
22. dom	46	46	dam	0	0	hom	0	0			1.00	1.00	0.00	0.00
23. lief	43	44	lieg	2	1	lies	1	0		1	0.93	0.95	0.13	0.54
24. buur	38	45	bier	2	0	muur	4	1	2		0.82	0.97	0.48	0.57
25. hoed	36	40	goed	3	0	hoek	5	5	2	1	0.78	0.86	0.57	0.34
26. lees	40	39	leus	0	1	mees	6	5		1	0.86	0.84	0.47	0.58
27. deun	38	41	dun	2	4	teun	5	1	1		0.82	0.89	0.26	0.31
28. jaar	46	44	gaar	0	0	haar	0	2			1.00	0.95	0.00	0.09
29. bijl	43	45	buil	0	1	mijl	3	0			0.93	0.97	0.30	-0.05
30. bed	42	40	pet	3	6	bek	1	0			0.91	0.86	0.01	0.17
31. pauw	22	41	paul	21	4	bouw	2	1	1		0.47	0.89	0.31	-0.04
32. wit	21	32	wip	18	13	fit	3	1	4		0.45	0.69	0.27	-0.02
33. zool	31	45	zaal	0	0	zoon	15	1			0.67	0.97	0.43	0.26
34. kop	23	33	kok	18	13	top	3	0	2		0.50	0.71	0.44	0.29
35. zien	33	38	zoen	4	1	ziel	8	6	1	1	0.71	0.82	0.41	0.59
36. bal	43	46	mal	3	0	bel	0	0			0.93	1.00	0.40	0.00
37. huur	34	45	guur	8	1	hier	2	0	2		0.73	0.97	0.65	0.26
38. doen	36	45	toen	6	1	boen	2	0	2		0.78	0.97	0.40	0.18
39. huis	45	46	huig	1	0	hijs	0	0			0.97	1.00	0.25	0.00
40. reuk	40	46	leuk	5	0	rook	0	0	1		0.86	1.00	0.30	0.00

*) dropped because of technical reasons.

**) statistic processing and analysis by L.W.A. van Herpt.

Table II: Subtest "Meaningless Items".**)

A: scores in 1970 - 47 testees

items 1 t/m 10 : 30 dB attenuation
items 11 t/m 20 : 30 dB "
items 20 t/m 30 : 35 dB "
items 30 t/m 40 : 40 dB "

B: scores in 1971 - 46 testees

items 1 t/m 10 : 40 dB attenuation
items 11 t/m 20 : 40 dB "
items 21 t/m 30 : 45 dB "
items 31 t/m 40 : 45 dB "

P = relative frequency of correct answers.

R_{it} = item-testcorrelation.

Stimulus	A	B	Alternative	A	B	Alternative	A	B	No score A	No score B	P A	P B	R_{it} A	R_{it} B
1. raal	46	46	raan	1	0	rool	0	0			0.97	1.00	0.25	0.00
2. fuip	45	44	fijp	1	1	guip	1	1			0.95	0.95	0.22	0.26
3. kem	47	46	pem	0	0	kel	0	0			1.00	1.00	0.00	0.00
4. dijp	46	44	duip	0	0	bijp	1	2			0.97	0.95	0.20	0.55
5. boop	45	45	moop	2	1	beup	0	0			0.95	0.97	0.33	0.04
6. wok	45	45	vok	2	1	wot	0	0			0.95	0.97	-0.05	0.18
7. sut	46	44	sit	0	0	sup	1	2			0.97	0.95	0.10	-0.07
8. jin	47	45	jun	0	0	hin	0	1			1.00	0.97	0.00	0.73
9. kag	45	46	kog	0	0	pag	2	0			0.95	1.00	0.22	0.00
10. jaup	41	40	haup	0	0	jaut	6	6			0.87	0.86	0.21	0.39
11. noos	47	46	nees	0	0	noof	0	0			1.00	1.00	0.00	0.00
12. lan	47	46	len	0	0	lon	0	0			1.00	1.00	0.00	0.00
13. jaaf	45	45	jaf	0	0	jaas	2	1			0.95	0.97	-0.02	0.04
14. deg	45	46	neg	1	0	teg	0	0	1		0.95	1.00	0.54	0.00
15. baun	44	45	buin	1	0	daun	1	1	1		0.93	0.97	0.45	0.18
16. nuin	44	44	nijn	0	0	nuil	3	2			0.93	0.95	0.04	0.11
17. bijf	45	46	mijf	2	0	pijf	0	0			0.95	1.00	0.57	0.00
18. doeg	36	45	noeg	0	0	doef	11	1			0.76	0.97	0.52	-0.08
19. kuuf	36	42	kief	4	0	tuuf	7	4			0.76	0.91	0.40	0.10
20. fiep	34	39	siep	5	3	fiet	8	3		1	0.72	0.84	0.13	0.52
21. daai	45	44	jaai	2	2	daam	0	0			0.95	0.95	0.29	0.55
22. dauk	47	46	nauk	0	0	tauk	0	0			1.00	1.00	0.00	0.00
23. kuig	33	41	puig	14	5	kaug	0	0			0.70	0.89	0.20	0.11
24. jijm	43	44	wijm	1	2	juim	3	0			0.91	0.95	0.47	0.02
25. foem	38	43	fuum	2	1	goem	7	2			0.80	0.93	0.20	0.40
26. luup	41	40	nuup	2	2	luut	4	2		2	0.87	0.86	0.04	0.57
27. zieg	47	45	zies	0	1	zuug	0	0			1.00	0.97	0.00	0.11
28. wook	27	37	zook	1	0	woop	19	9			0.57	0.80	0.07	0.48
29. meul	37	38	mool	0	0	meun	10	8			0.78	0.82	0.29	0.51
30. jeem	34	43	joom	0	0	jeel	13	3			0.72	0.93	0.14	0.57
31. soem	35	44	uum	9	0	foem	1	1	2	1	0.74	0.95	0.36	-0.07
32. zaaf	47	45	zoof	0	0	zaf	0	1			1.00	0.97	0.00	0.04
33. buut	26	39	muut	9	2	buuk	9	4			0.55	0.84	0.39	0.55
34. hieg	8	10	jieg	33	34	gieg	5	1	1	1	0.17	0.21	0.10	0.18
35. poor	39	45	peur	4	0	toor	3	1	1		0.82	0.97	0.57	0.11
36. geug	19	39	geuf	26	7	geeg	1	0	1		0.40	0.84	0.15	0.19
37. weel	18	23	weem	28	22	zeel	0	1	1		0.38	0.50	0.26	0.10
38. wong	40	45	weng	0	0	hong	7	1			0.85	0.97	0.30	-0.01
39. fug	31	45	gug	13	1	fig	2	0	1		0.65	0.97	0.34	0.73
40. dif	28	44	bif	16	2	tif	2	0	1		0.59	0.95	0.35	-0.07

**) Statistic processing and analysis by L.W.A. van Herpt.

Phonetic transcription of Meaningful Items

1	vaas [vas]	waas* [was]	vaag [vaɣ]
2	bijt [bɛit]	buit [bʌyt]	meid* [mɛit]
3	hek [hɛk]	hok [hɔk]	hak [hak]
4	fout [faut]	goud [xaut]	vat [vat]
5	wang [waŋ]	zang [zaŋ]	hang [haŋ]
6	boos [bos]	boog [boɣ]	poos [pos]
7	bus [bœs]	bos [bɔs]	mus [mœs]
8	wol [wɔl]	vol [vɔl]	wal [wal]
9	boel [bul]	peel [pul]	doel [dul]
10	zing [ziŋ]	zien [zin]	ving [vIŋ]
11	vuur [vyr]*	vier [vir]	zuur [zyr]
12	maal [mal]	maai [mai]	maan [man]
13	buit [bʌyt]	bout [baut]	buik [bʌyk]
14	bier [bir]	mier [mir]	buur [byr]
15	been [ben]	boon [bon]	peen [pen]
16	peul [pøl]	pool [pol]	beul [bøl]
17	bel [bɛl]	bal [bal]	bol [bɔl]
18	hak [hak]	jak [jak]	hap [hap]
19	dik [dIk]	dit [dIt]	tik [tIk]
20	boot [bot]	poot [pot]	dood [dot]
21	mus [mœs]	mug [mœɣ]	mis [mIs]
22	dom [dɔm]	dam [dam]	bom [bɔm]
23	lief [lif]	lieg [liɣ]	lies [lis]
24	buur [byr]	bier [bir]	muur [myr]
25	hoed [hut]	goed [xut]	hoek [huk]
26	lees [les]	leus [løs]	mees [mes]
27	deun [døn]	dun [doen]	teun [tøn]
28	jaar [jar]	gaar [xar]	haar [har]
29	bijl [bɛil]	buil [bʌyl]	mijl [mɛil]
30	bed [bɛt]	pet [pɛt]	bek [bɛk]
31	pauw [pau]	paul [paul]	bouw [bau]
32	wit [wIt]	wip [wIp]	fit [fIt]
33	zool [zɔl]	zaal [zal]	zoon [zon]
34	kop [kɔp]	kok [kɔk]	top [tɔp]
35	zien [zin]	zoen [zun]	ziel [zil]
36	bal [bal]	mal [mal]	bel [bɛl]
37	huur [hyr]	guur [xyr]	hier [hir]
38	doen [dun]	toen [tun]	boen [bun]
39	huis [hʌys]	huig [hʌyɣ]	hijs [hɛis]
40	reuk [røk]	leuk [løk]	rook [rok]

*Dutch /w/ = [v] labiodental initially.

*final d = [t]

*r-variants in Dutch are [r] and [R]

Phonetic transcription of Meaningless Items

1	raal [ral]	raan [ran]	rool [rol]
2	fuip [fʌyp]	fijp [fɛip]	guip [χʌyp]
3	kem [kɛm]	pem [pɛm]	kel [kɛl]
4	dijp [dɛip]	duip [dʌyp]	bijp [bɛip]
5	boop [bop]	moop [mop]	beup [bøp]
6	wok [wɔk]	vok [vɔk]	wot [wɔt]
7	sut [sʉt]	sit [sɪt]	sup [sœp]
8	jin [jɪn]	jun [jœn]	hin [hɪn]
9	kag [kaχ]	kog [kɔχ]	pag [paχ]
10	jaup [jaup]	haup [haup]	jaut [jaut]
11	nooa [noa]	nees [nes]	noof [nof]
12	lan [lan]	len [lɛn]	lon [lɔn]
13	jaaf [jaf]	jaf [jaf]	jaas [jas]
14	deg [dɛχ]	neg [nɛχ]	teg [tɛχ]
15	baun [baun]	buin [bʌyn]	daun [daun]
16	nuin [nʌyn]	nijn [nɛin]	nuil [nʌyl]
17	bijf [bɛif]	mijf [mɛif]	pijf [pɛif]
18	doeg [duχ]	noeg [nuχ]	doef [duf]
19	kuuf [kyf]	kief [kif]	tuuf [tyf]
20	fiep [fip]	siep [sip]	fiet [fit]
21	daai [dai]	jaai [jai]	daam [dam]
22	dauk [dauk]	nauk [nauk]	tauk [tauk]
23	kuig [kʌyχ]	puig [pʌyχ]	kaug [kaʉχ]
24	jijm [jɛ̈im]	wijm [wɛim]	juim [jʌym]
25	foem [fum]	fuum [fym]	goem [χum]
26	luup [lyp]	nuup [nyp]	luut [lyt]
27	zieg [ziχ]	zies [zis]	zuug [zyχ]
28	wook [wok]	zook [zok]	woop [wop]
29	meul [møl]	mool [mol]	meun [møn]
30	jeem [jem]	joom [jom]	jeel [jel]
31	soem [sum]	suum [sym]	foem [fum]
32	zaaf [zaf]	zoof [zof]	zaf [zaf]
33	buut [byt]	muut [myt]	buuk [byk]
34	hieg [hiχ]	jieg [jiχ]	gieg [ɣiχ]
35	poor [por]	peur [pør]	toor [tor]
36	geug [χøχ]	geuf [χøf]	geeg [χeχ]
37	weel [wel]	weem [wem]	zeel [zel]
38	wong [wɔŋ]	weng [wɛŋ]	hong [hɔŋ]
39	fug [fœχ]	gug [χœχ]	fig [fɪχ]
40	dif [dɪf]	bif [bɪf]	tif [tɪf]

Table III.

Subtest "Meaningful Items".

1970	attenuation	average P-value	1971	attenuation	average P-value
1st block	30 dB	0.921	1st block	40 dB	0.944
2nd block	35 dB	0.858	2nd block	40 dB	0.944
3rd block	35 dB	0.905	3rd block	45 dB	0.929
4th block	40 dB	0.707	4th block	45 dB	0.902

Subtest "Meaningless Items".

1970	attenuation	average P-value	1971	attenuation	average P-value
1st block	30 dB	0.958	1st block	40 dB	0.962
2nd block	30 dB	0.895	2nd block	40 dB	0.961
3rd block	35 dB	0.830	3rd block	45 dB	0.910
4th block	40 dB	0.615	4th block	45 dB	0.817

8.2 A striking difference occurs between the scores of the groups taking part in the experiment in 1970 and 1971 respectively, both for the subtest "meaningful items" and for "meaningless items". Although the loudness levels in 1971 were lower than in 1970 , the results of the 1971-group were better. It is possible that the place where the test was conducted has something to do with it. The language laboratory where the test was conducted in 1970 is situated on the side of the street, so that street-noise may have influenced the discrimination of the items. In 1971, however, the test was conducted in the noise-reduced booths of the Institute of Phonetic Sciences. This idea is corroborated by the fact that in 1970 "no scores" occurred 32 times in Part I and 13 times in Part II, whereas in 1971 there were 4 "no scores" in Part I and 6 in Part II. The degree of difficulty of the items seems, on the whole, to be on one level, independent of the loudness level chosen. The group of 1971 scored better, but finds the items as difficult or as easy as the 1970-group. (see tables I and II).

8.3 Both groups make more mistakes in the meaningless items than in the

subtest "meaningful items", this in accordance with data found in literature on the subject. When judging "meaningless items" the subject is deprived of the possibility of using the deductive powers he possesses based on his linguistic knowledge. The hypothesis stated in 4.4.0 was not tested statistically:

a) owing to a mistake made with the attenuation of the loudness level in 1970 (see 7.2).

b) owing to the relatively small number of mistakes made in 1971.

However, a trend is clearly observable: the subtest "meaningless items" has a greater number of mistakes than the subtest "meaningful items", (see 8.1).

8.4 The connection between loudness level and degree of difficulty of items is apparent from the average P-values per loudness level. (see Table III, page 108.

8.5 Every time the 4th block of both test parts appears to present most difficulties, even of the same loudness level was used as in the 3rd block. It is possible that the 4th block happens to have the items which are most difficult to discriminate in both subtests. It might also point to a certain amount of fatigue of the listeners, although this is doubtful when considering the duration of the test, ca. 10 minutes.

8.6.0 A closer look at the scores shows the following:

8.6.1 Subtest "meaningful Items".

Most mistakes were made in the 4th block, both in 1970 and in 1971, notwithstanding the fact that the loudness levels in 1971 of blocks 3 and 4 were identical (see 8.5). Items which are difficult to discriminate are: item 32 wit, with preference shown for wip in both tests (18, 13) as against fit (3.1)*), item 34 kop with strong preference shown for the alternative kok (18, 13) as against top (3.0).

*) Scoring on alternative items will be indicated as follows: (,), the first number indicating the 1970 score, the second number that of 1971.

A notable discrepancy between scores of 1970 and 1971 are found in:
item 31 pauw: paul (21.4), bouw (2.1)
item 37 huur: guur (8.1), hier (2.0)
item 33 zool: zoon (15.1), zaal (0.0).

Some items appear to have been scored correctly by (almost) all subjects, e.g. item 21 mus, item 22 dom, item 36 bal, and item 39 huis. The possibility that either the stimulus is redundant or that the alternatives chosen do not function cannot be ruled out.

If we compare the results with those of the pilot-investigation it becomes clear that the subjects' reaction to the stimuli and their preference for certain alternatives remained practically unaltered.

8.6.2 Subtest "Meaningless Items".

Poor discrimination occurs in blocks 3 and 4:
item 28 wook with preference for woop (19.9) as compared to zook (1.0),
item 36 geug with preference for geuf (26.7) as compared to geeg (1.0).
The very poor discrimination of item 34 hieg cannot be explained. In 1970 and 1971 39 and 36 wrong responses occurred. Jieg (33, 34) was heard by 33 and 34 subjects respectively as against 5 and 1 who scored gieg (5.1). The very first time, during the pilot-investigation, the item was poorly responded to, notwithstanding the fact that the loudness level was higher that year.

The subtest "meaningless items", as well as the subtest "meaningful items", contains a number of items which (almost) all subjects respond to correctly:
item 1 raal, item 3 kem, item 11 noos, item 12 lan, item 22 dauk, item 27 zieg, and item 32 zaaf.

When we consider the items which were wrongly scored and see which were the alternatives chosen by the testees, it becomes clear that, taking the sound substitutions on the whole, there is a tendency to substitute consonants rather than vowels. It becomes clear, that the influence which the substitution of one phoneme exercises on the totality of sound impression which a monosyllable evokes, is much greater, than the impression left by the actual characters of the sounds would have led us to suppose. What is meant here is the influence

which a consonant or vowel exercises on the following or preceding consonant or vowel. Subjects are inclined to make more consonant substitution than vowel substitutions. This might be explained as follows:

a) The use of a linear amplifier, which first influences the highest and the lowest frequences unfavourable when attenuation takes place as a result of the curvature of the Fletcher-Munson curve.

b) The amplitude of the vowels is higher as a rule than the amplitude of the consonants.

9.0 Audiometry.

When comparing the results of the speech discrimination test with the data of the tone audiometry (see 9.3), no correspondence appears between the results of the two types of tests. That is to say, subjects with the highest number of wrong responses(S) for the speech discrimination test do not show up worst in the tone audiograms. Nor have the best subjects (B) the best tone audiograms.

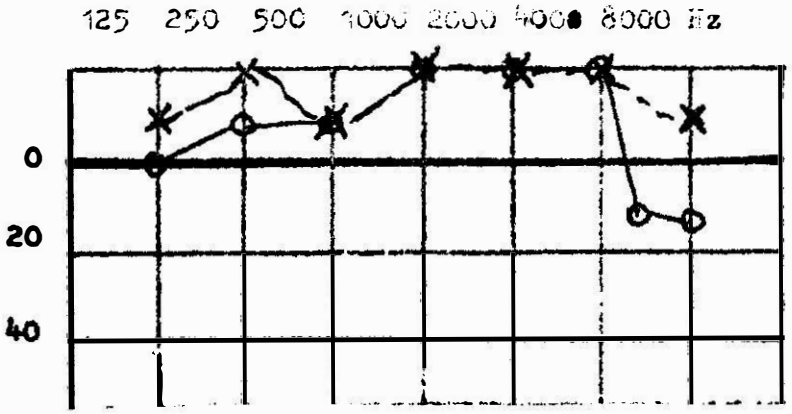
9.1 The following should be taken note of: With the aid of tone audiometry the ability to hear pure tones is measured. The results are rendered in a tone audiogram: a graphic representation of loss hearing in dB when compared to a normal auditory organ*), set out as a function of the frequency. The tone threshold shows the point where a subject actually hears the pure tone in 50% of the cases when it is produced. No verdict can be given as regards the ability of understanding speech when basing it on a tone audiogram only. Speech is, after all, a complex signal and certain pathological conditions, such as recruitment or certain central factors, will affect the understanding of speech rather than of pure tones adversely. In order to test the capacity of understanding speech use is made of speech audiometry. This speech audiogram is usually made as follows: the patient, wearing headphones, is subjected to a number of tape-recorded monosyllabic or polysyllabic words. The intensity level is attenuated per group of 10 words. The patient is requested to repeat what he has heard. The result is put in a diagram of which the horizontal shows the intensity in decibels, the vertical the percentage of correct responses. The shape of the graph gives the

*) 'Normal' defined as international zero-level.

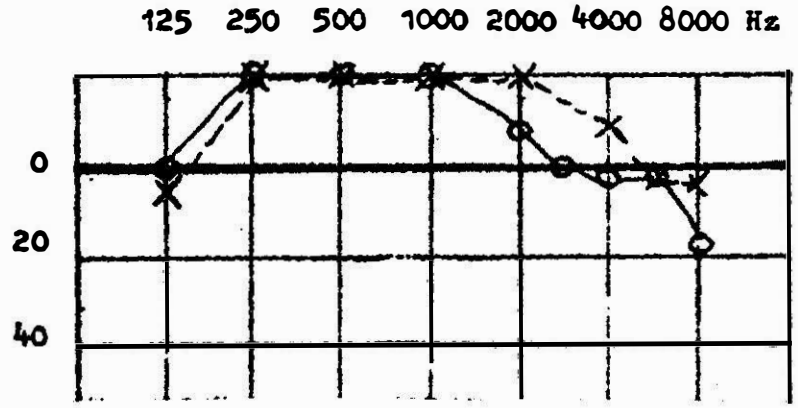
investigator certain information about the type of hearing loss and possibilities of revalidation.

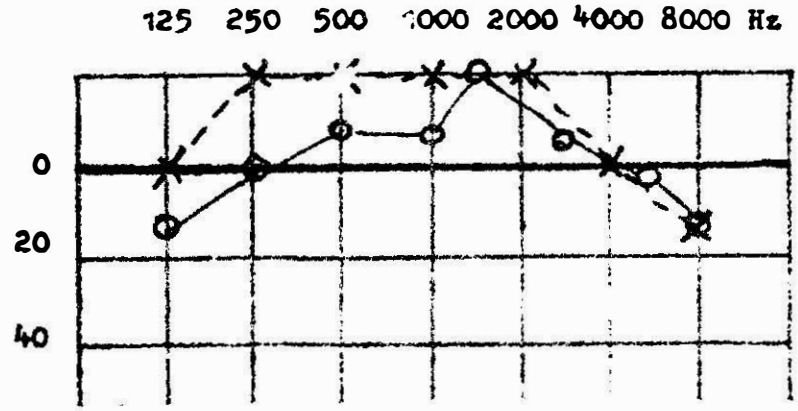
).2 The speech discrimination test as described here, is, therefore, a modification of the usual speech audiometry (for closer analysis of loss of hearing special types of speech audiometry are used). As none of the subjects participating in the speech discrimination test appeared to have an abnormal tone audiogram, no essential deviations were to be expected from the discrimination of speech based on the capacity of hearing. The divergences which the subjects of normal hearing showed in their ability to discriminate speech in the test conducted here, depend apparently on different factors, of which nothing can be said on the basis of this test.

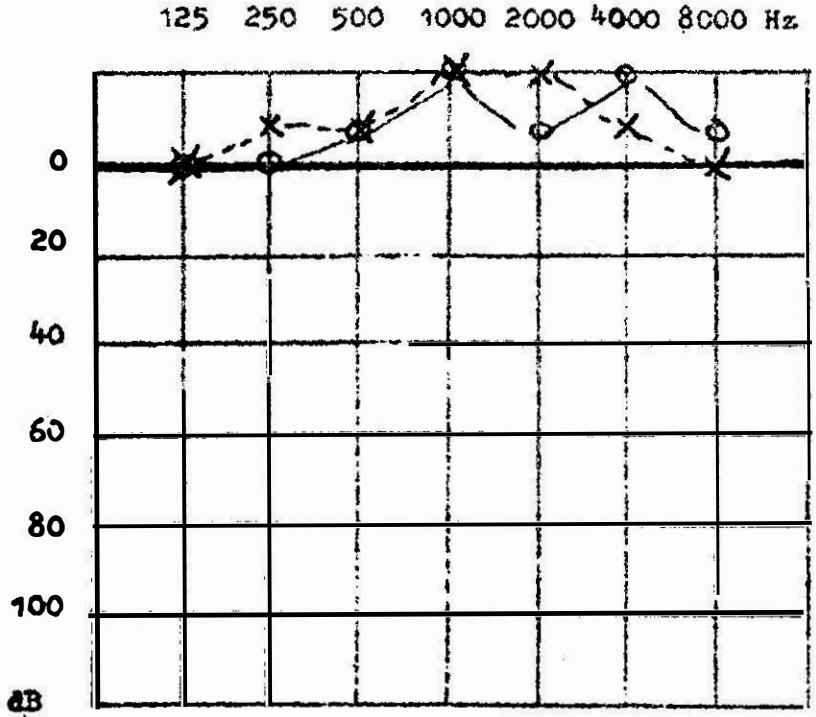
(B)



TONE AUDIOGRAM
of the best (B)
testees in the speech-
discriminationtest
1970.



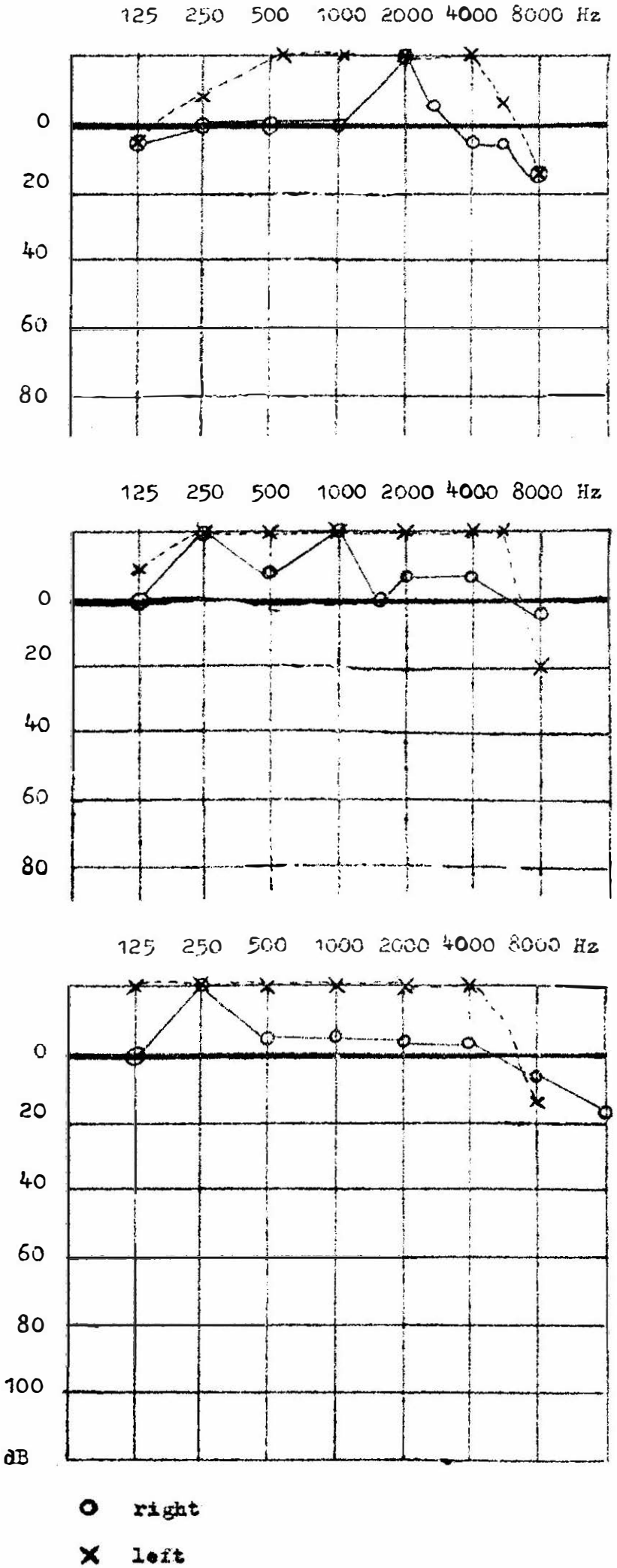




o right

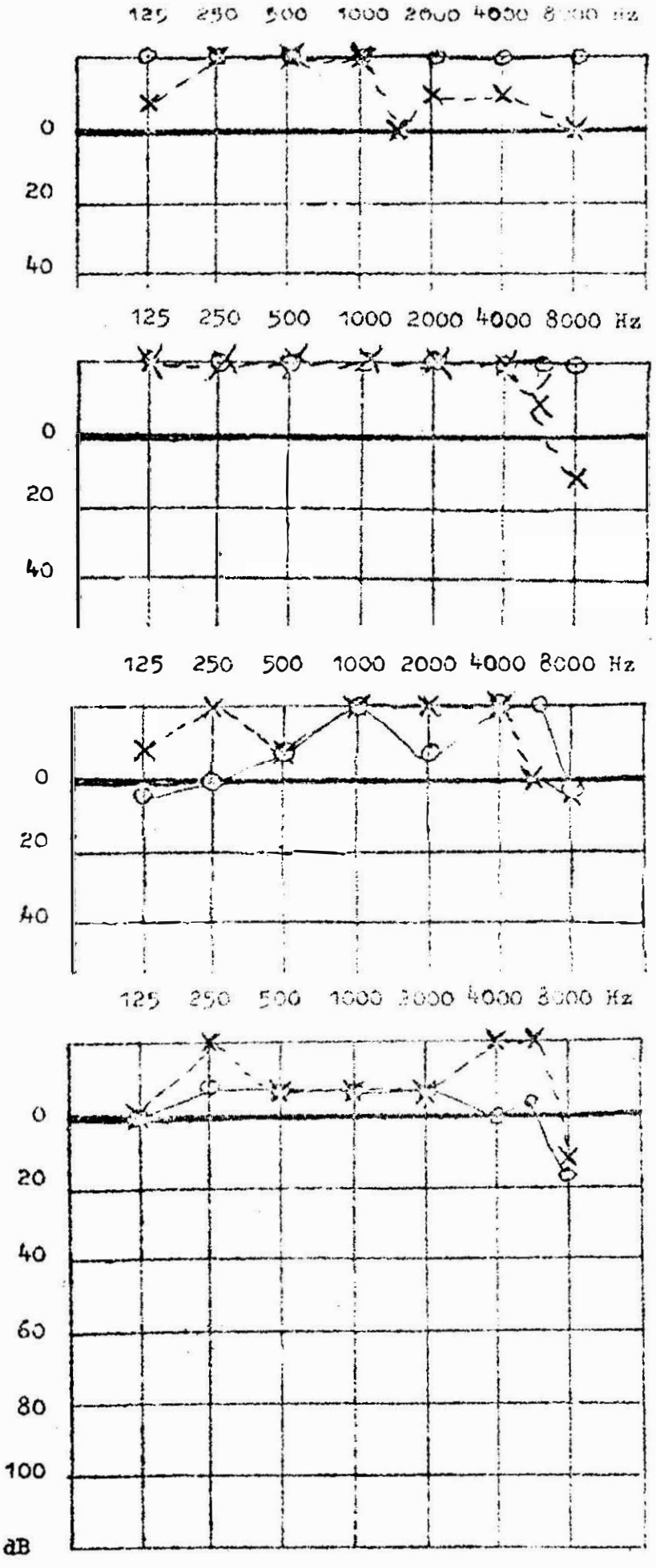
× left

(B)



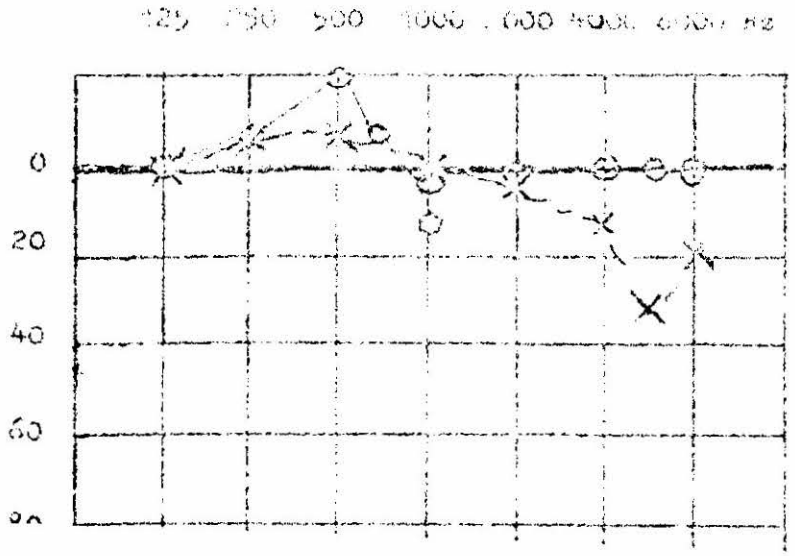
TONE AUDIOGRAM
of the best (B) testees in the speech-discriminationtest 1970.

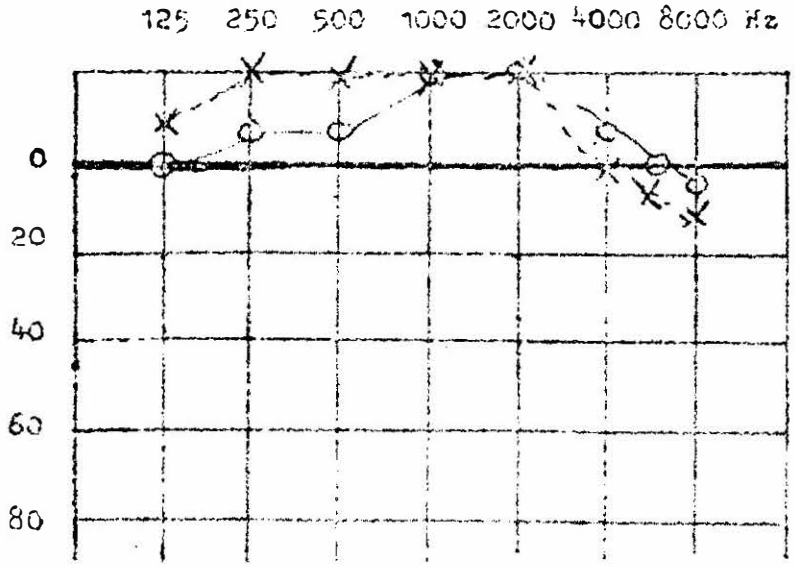
(W)

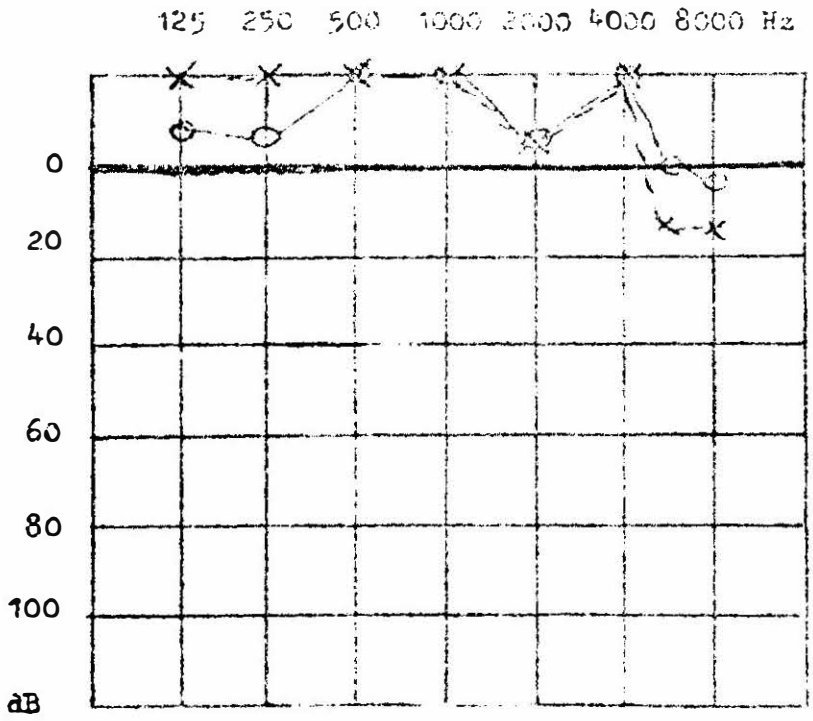


TONE AUDIOGRAM of the worst (W) testees in the speech-discriminationtest 1970.

(W)





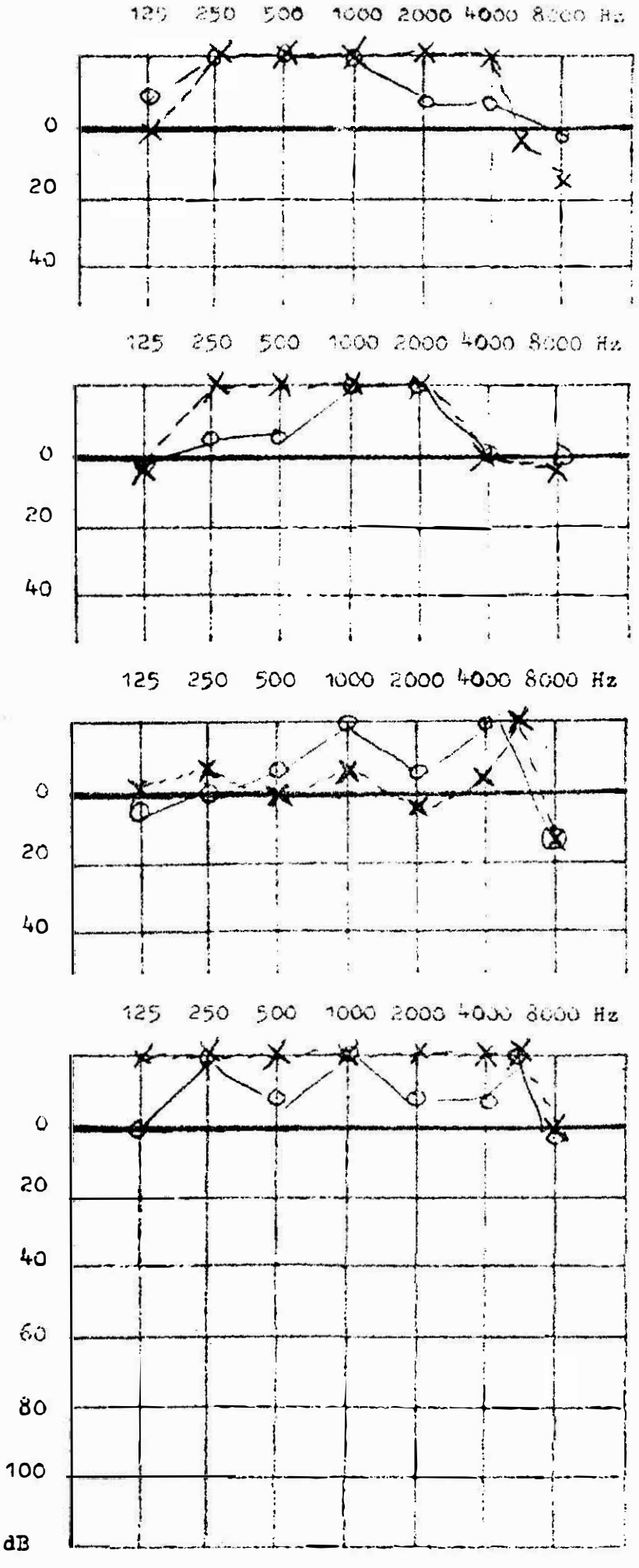


o right

x left

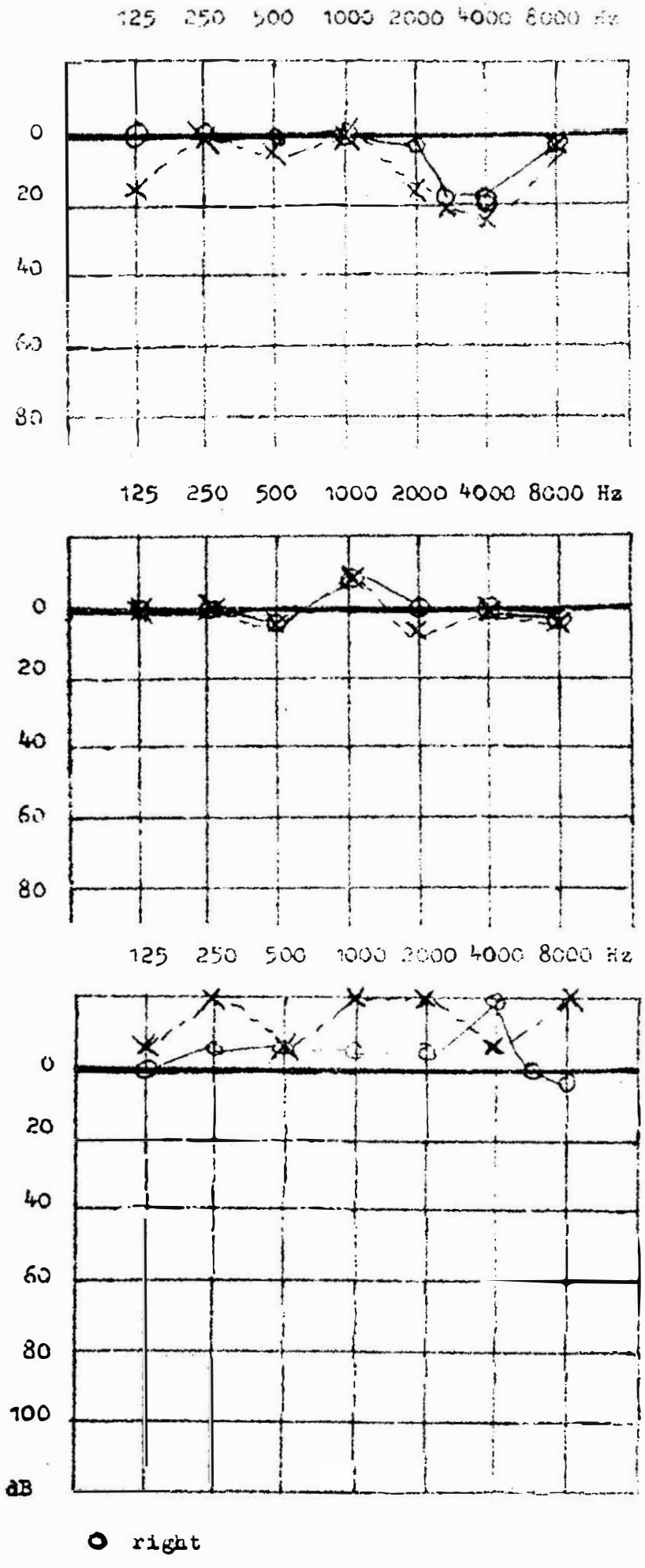
TONE AUDIOGRAM of the worst (W) testees in the speech-discriminationtest 1970.

(B)



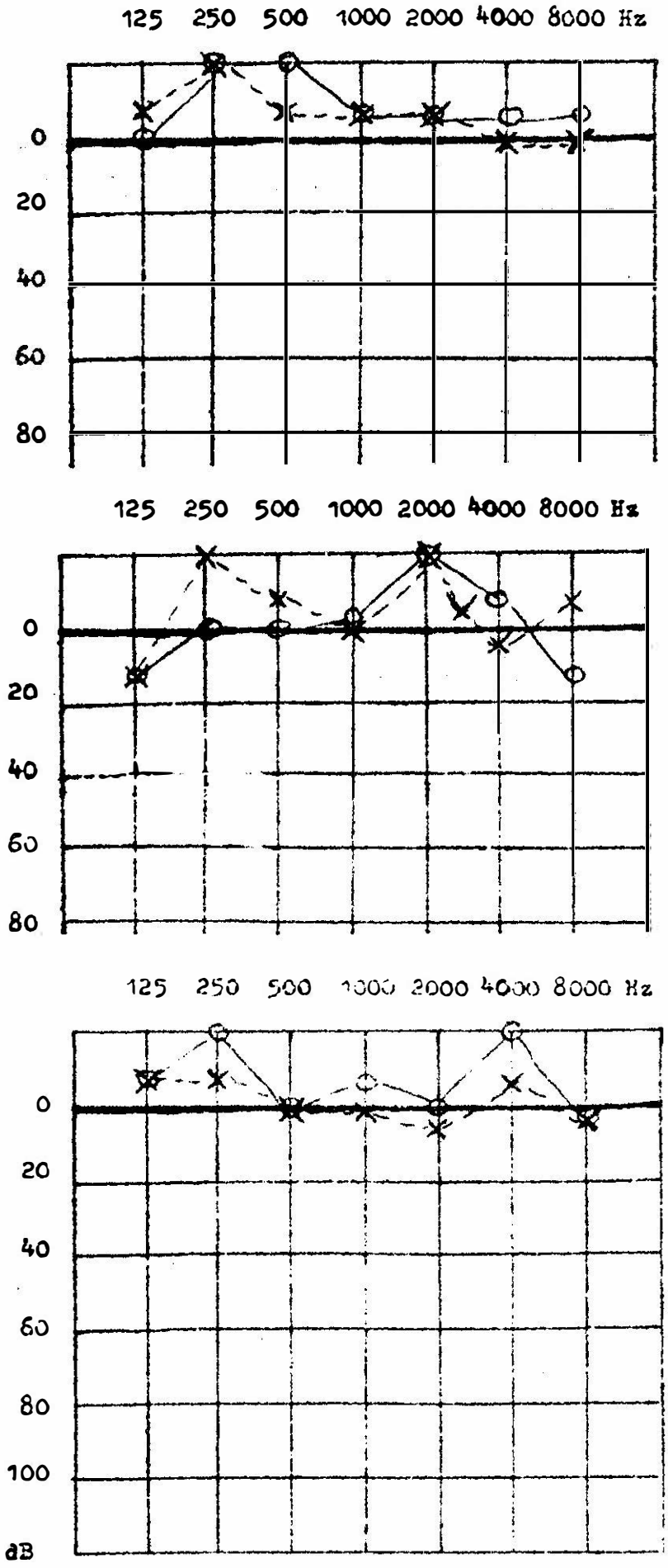
TONE AUDIOGRAM
of the best (B) testees in the speech-discriminationtest 1971.

(w)



TONE AUDIOGRAM
of the worst (w) testees in the speech-discriminationtest 1971.

(W)



TONE AUDIOGRAM
of the worst (W) testees in the speech-discriminationtest 1971.

10.0.0 Conclusion.

10.1.0 The objection that might be raised against the speech discrimination test in this form, is, that the results state only whether certain speech stimuli at a certain loudness level are perceived well or are not perceived, but that they give no information as to the way in which different results could have occurred between subjects. In other words, it is not quite clear what is being measured with this test. This is partly the result of the fact that only subjects with normal hearing participated. Consequently it is impossible to give a decisive answer on the points brought up under 4.1.0 to 4.3.0 inclusive. Some remarks have been made under the heading 'Results'.

10.2.0 The value and the usefulness of the test might be assessed more fully in the following manner:

10.2.1 The speech discrimination test taken with a new group of subjects in such a way that the whole corpus of the test is offered on the same loudness level. With the new findings on hand it will be possible to assess which items do not discriminate and which alternatives are not functional. A comparison with the results of 10.2.2 will no doubt be of some interest.

10.2.2 The speech discrimination test taken with a group of subjects with impaired hearing. The subjects would have to satisfy the following conditions:

a) hearing loss will have to be to the same degree for the whole group
b) hearing loss will have to be conform for the whole group. (e.g. perceptive loss without recruitment).

In order to assess a) and b) a tone audiogram would have to be made of each subject.
Dependent on the results of this test a criterion could be set up by which the border - line sufficient / insufficient discrimination of speech is settled with regard to possible hearing-loss.

10.2.3 A follow-up investigation of language laboratory students. This could give an indication regarding the connection between the ability of discriminating speech and the ability of acquiring a correct pronunc-

iation of a foreign language.

10.2.4 A comparison of the data of the speech discrimination test with those of the SEASHORE-test, which, as was stated before, was conducted in combination with the speech discrimination test.

10.2.5 A comparison of the test data of the speech discrimination test, presented in the form of a multiple choice test and also presented in a free-choice situation, conducted with two matched groups of subjects. Execution of the free-choice test would only be a matter of a different way of scoring.

10.3.0 With these new data on hand an assessment could be made regarding the following points:

a) 1. the speech discrimination test used as a means of testing
   2. if so, what category of testees can be subjected to it.

b) some alterations to be made in the speech discrimination test, viz. a change of some items or of some alternatives

c) the discrimination test to be maintained in its present form.