DOCUMENT RESUME

ED 023 209

EC 002 504

By-Ahlstrom, K.G.; And Others

Experiment in Learning to Discriminate Frequency Transposed Speech.

Uppsala Univ. (Sweden). Inst. of Education.
Spons Agency-Swedish Council for Social Science Research, Stockholm.; Tri-Centennial Fund of the Bank of Sweden, Stockholm.
Note-37p.
EDRS Price MF-$0.25 HC-$1.95
Descriptors-Articulation (Speech), Audio Equipment, Auditory Discrimination, Auditory Perception, Auditory Tests, *Auditory Training, Consonants, *Exceptional Child Research, Paired Associate Learning, *Perception, Statistical Data, Students, Student Testing, Syllables, Vowels

In order to improve speech perception by transposing the speech signals to lower frequencies, to determine which aspects of the information in the acoustic speech signals were influenced by transposition, and to compare two different methods of training speech perception, 44 subjects were trained to discriminate between transposed words or syllables. Since the subjects had normal hearing, a hearing less was simulated by including a low-pass filter and white noise in the circuit for transposing speech signals. The consonant sounds most easily identified at the first test session were /g/, /k/, /sp/ and /sk/, with /d/ and /c/ the most difficult. The subjects learned to rely in their discriminations on the code dimensions voicing, duration, and cluster consonant but not on place of articulation. Neither of the two training methods compared, anticipation and recall for paired associate learning, was found more effective. It did not seem to matter which of the two training materials, words or syllables, was used. Six figures, 18 tables, and six references are included. (GD)

report

The Institute of Education
UPPSALA UNIVERSITY
S:t Olofsgatan 12, Uppsala
SWEDEN
Head: W. Sjöstrand

36
1968

EXPERIMENT IN LEARNING TO
DISCRIMINATE FREQUENCY TRANSPOSED SPEECH

by

K.-G. Ahlström, A. Risberg and V. Lindhe

This investigation is part of a project concerning methods of training in speech perception. The project is being financed by grants form the Swedish Council for Social Science Research and the Tri-Centennial Fund of the Bank of Sweden.

Valdy Lindhe
Investigation leader

Karl-Georg Ahlström
Arne Risberg
Scientific leaders

EXPERIMENT IN LEARNING TO DISCRIMINATE FREQUENCY TRANSPOSED SPEECH

Ahlström K.-G., Risberg A. and Lindhe V. Experiment in learning to discriminate frequency transposed speech. Rep. Inst. of Educ. No. 36 1967. Forty-four Ss were trained to discriminate between transposed words or syllables. Since the Ss had normal hearing, it was necessary to simulate a hearing loss by including a low-pass filter and white noise in the circuit for transposing the speech signals. The Ss were given ten practice sessions. At sessions 1, 3, 5, 7, 9 and 10 they were given a test.
The consonant sounds most easily identified by the Ss at the first test session were /g/, /k/, /sp/ and /sk/. The sounds /d/ and /ç/ were the most difficult to identify.
The changes in sound discrimination which take place with training occur rapidly and it is doubtful whether further training would have had any appreciable effect on discrimination. The Ss learned to rely in their discriminations on the code dimensions voicing, duration and cluster consonant but not on place of articulation. No standpoint can be adopted as to which of the two training methods compared here - anticipation and recall for paired-associate learning - is the more effective. It does not seem to matter which of the two training materials - words or syllables - is used.

Background and purpose

A hearing loss is generally less pronounced in the bass portion of the frequency band. When this is the case, it seems theoretically possible to improve speech perception by transposing the speech signals to lower frequencies. The approaches to this problem have involved transposing either the whole speech spectrum (Oeken, 1963; Pimonow, 1963) or the high--frequency components of the speech signal (König-Eichler, 1954; Johansson, 1959; Ling-Druz, 1967).

The "transposer" invented by B. Johansson (Johansson, 1959) performs the last-mentioned operation. The block diagram in Fig. 1 shows the principles for this kind of transposition.

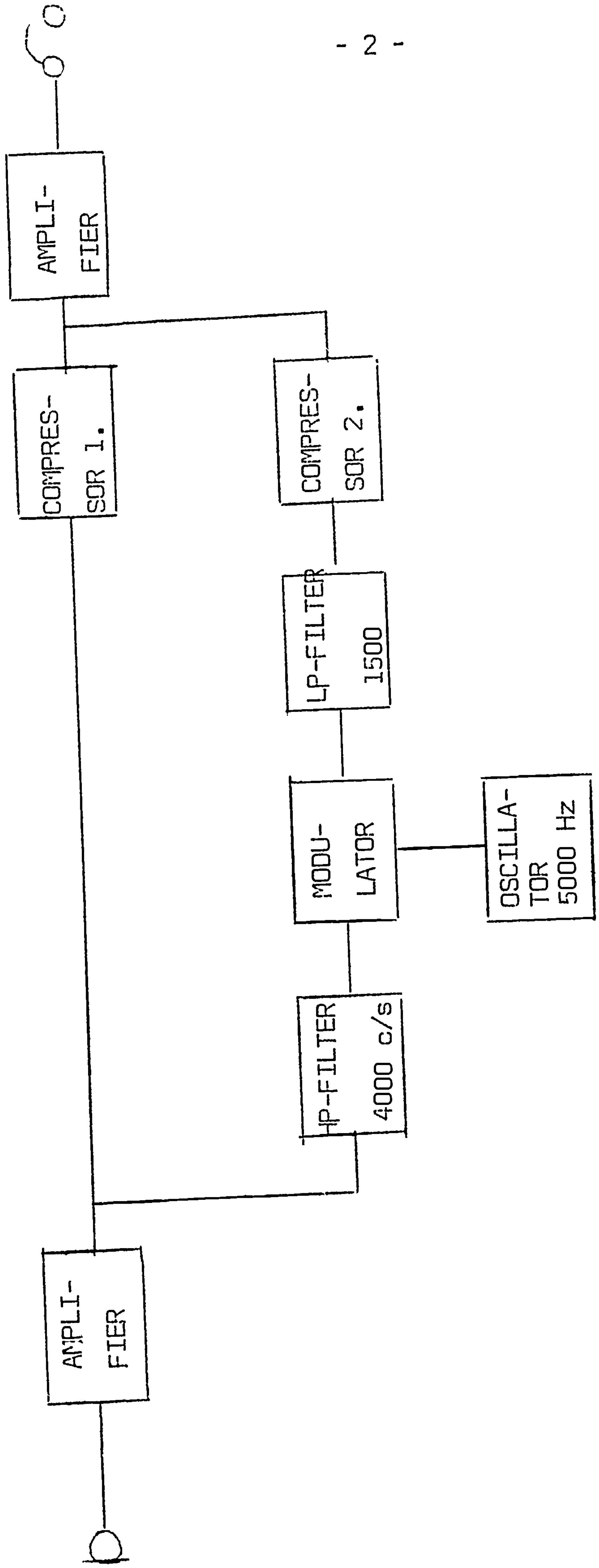

Fig. 1. Block diagram of hearing aid for transposing high-frequency sounds to lower frequencies ("transposer"). From B. Johansson (1959, 1966).

Compressor "1" normally works in its nonlimiting region and compressor "2" in the limiting region. The amplitude range of 20 dB is then compressed to a 3-dB range. The attack time is about 5 ms and release time about 50 ms.

All speech components above the cut-off frequency of the high-pass filter are transposed down to the low-frequency region. This means, for instance, that /s/ will be perceived as a low frequency voiceless sibilant.

Very few experiments have been undertaken to study the effects of transposition (Wedenberg, 1959; Johansson and Sjögren, 1965; Ling and Druz, 1967).

The main purpose of the experiment reported here was to elucidate which aspects of the information in the acoustic speech signals were influenced by transposition. Another aim was to compare two different methods of training speech perception.

Method

Simulation of hearing loss

Since the Ss had normal hearing, it was necessary to simulate a hearing loss by including a low-pass filter and noise in the equipment for transposing the speech signals. Fig. 2 contains the block diagram for the equipment.

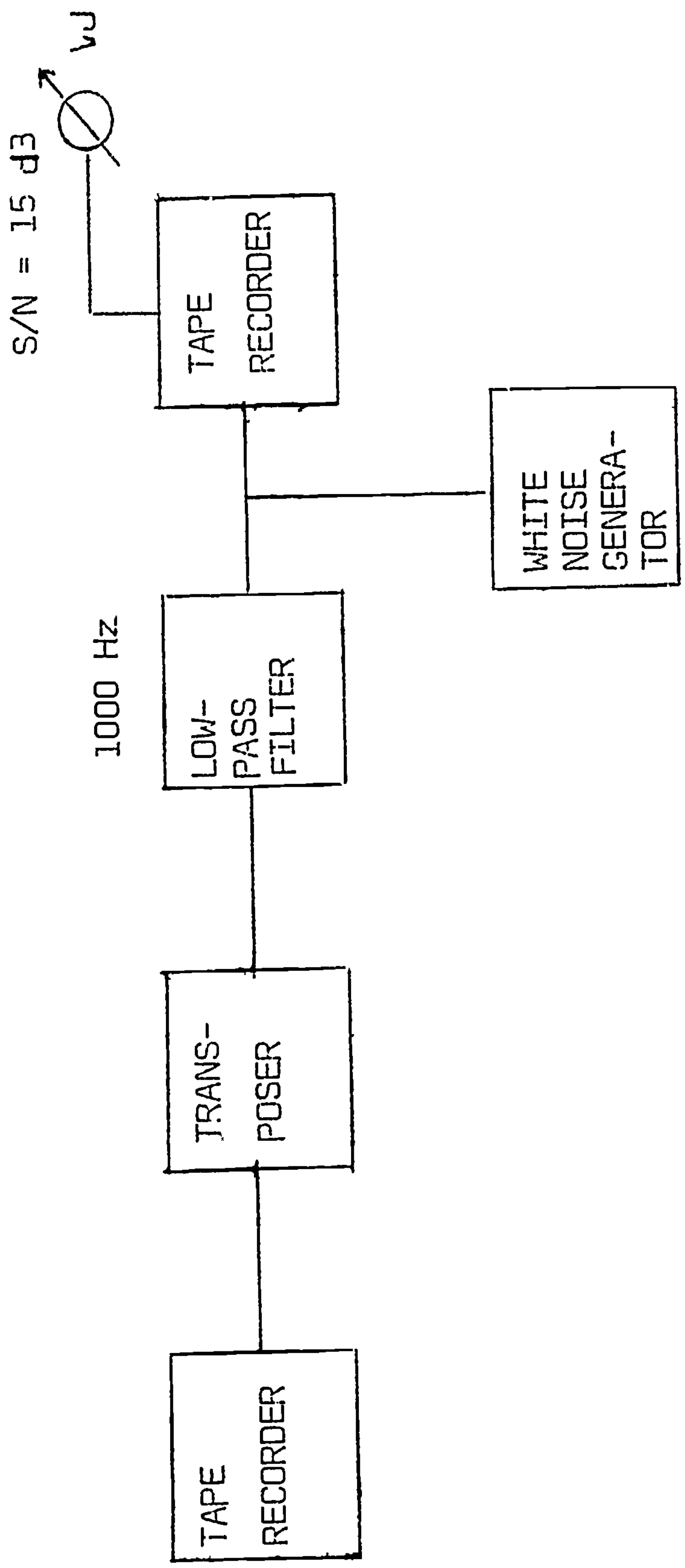

Fig. 2. Block diagram of set-up for recording the training tapes. The low-pass filter has a cut off frequency of 1000 c/s and a sharp cut off. The signal to noise ratio is read on a RMS instrument at the output from the taperecorder.

The cut-off frequency was 1000 c/s with very sharp cut off, and the signal-to-noise ratio 15 db. Without low-pass filtering a S with normal hearing would have perceived both the transposed and the non-transposed speech. Now only non-transposed components below 1000 c/s and transposed components are perceived. The testtapes were recorded through the "trans-poser" by H. Sjögren at the Institute for Technical Audiology, Karolinska Institutet, Stockholm. The subjects listened to the signal through head-phones TDH 39 at a comfortable listening level.

Design

The Ss were 44 students at the Institute of Education in Uppsala. They were randomly assigned to four groups of 11 Ss each; the groups differed with respect to practice material and training method.

Group	No. of Ss	Practice material	Training method
1	10	Words	Anticipation
2	11	Words	Recall
3	9	Syllables	Anticipation
4	8	Syllables	Recall

Two Ss trained at the same time with the same material but used different training methods. The pairing was random. The working pace was determined by the S using the anticipation method. Six Ss dropped out during the experiment because of sickness. In these cases the Ss who dropped out had been training by anticipation. The experimenter determined the working pace for the remaining S utilizing the recall method. The actual number of Ss is given in the table above.

Practice material

In a pilot study 42 syllables were presented by a ReVox G36 tape recorder to 18 Ss. The speech was transposed in the way described above. The syllables contained 14 different consonant sounds, vi . /b/, /d/, /g/, /k/, /t/, /h/, /f/, /ʃ/, /ç/, /sk/ and /sp/. Each consonant sound was combined with one of the vowels /a/, /o/ and /i/ to form syllables like /ba/, /bo/, /bi/ and so on. After a syllable had been presented, the Ss wrote down the letters they thought corresponded to the non-transposed sounds.

The conditional frequencies showing how a given transposed sound was interpreted are to be found in Table 1. It should be observed that each consonant occurred three times and each vowel 14 times per S.

Table 1. The conditional frequencies (percentages) in the pilot study. The column x contains all other letters reported by the Ss.

S's responses

Correct responses	b	d	g	p	t	k	h	s	ʃ	ç	f	sk	sp	st	a	o	i	x	tot.
b	61	7	18			3	2							2				7	100
d	43	18	37								2								100
g	6	2	90									2							100
p		6	6	3		38	34				3								100
t			6	35	14	39	6												100
k	2		2	9	5	68	2												100
h	5		2	2		16	59											16	100
s						1	3	21	33	26	15							1	100
ʃ						2	5	30	26	26	11								100
ç							4	18	38	24	15								100
f						3	24	18	20	8	27								100
sk												52	30	18					100
sp												5	73	12					100
st												36	28	36					100
a															99	1			100
o																77	23		100
i															1	41	58		100

On the basis of the data in this input-output matrix, two consonant sounds were chosen from each "sound group", viz. /d - g/, /k - t/, /s - ç/ and /sk - sp/. By combining them with the vowels /a/ and /o/, the following syllables were obtained:

da, do	sa, so
ga, go	tja, tjo (ça:, çu:)
ka, ko	ska, sko
ta, to	spa, spo

These 16 syllables were arranged in random order in a list consisting of 48 items. The sounds /d/ and /ç/ recurred five times, /g/ and /s/ seven times and the other consonant sounds six times in the list. This list will be referred to as the syllable list in the following text.

By choosing from the endings -cka, -ll, -m, -rra, -rv, -ta, -v -va and by adding the chosen ending to the syllables, the following words were obtained:

data, dom	sav, socka
gata, gom	tjav, tjocka
kall, kova	skarv, skorra
tall, tova	sparv, sporra

A word list consisting of 48 items was arranged in exactly the same way as the syllable list.

Both the syllable and the word lists were divided into two halves, one part containing item nos. 1 - 24 and the other item nos. 25 - 48. Two versions of each list were obtained by reversing the consecutive order of the two parts, viz. version A; nos. [1 - 24]→[25 - 48] , and version B; item nos. [25 - 48]→[1 - 24] .

Training methods

Two training methods were compared: the anticipation method versus the recall method for paired associate learning. In both cases the task of the S was to learn to match transposed sounds (S) with the graphical signs (G) for the corresponding non-transposed sounds.

In the anticipation method a transposed syllable or word, S_i, is presented to the subject, who responds, R, by trying to identify the graphical signs of the non-transposed syllable or word, after which these signs, G_i, are shown to him. Accordingly the sequence is as follows:

$$S_1 \longrightarrow R \longrightarrow G_1; \quad S_2 \longrightarrow R \longrightarrow G_2; \quad \ldots; \quad S_i \longrightarrow R \longrightarrow G_i.$$

In the application of the recall method a stimulus, S_i, is presented to the S. At the same time he is shown the syllable or word, G_i, corresponding to S_i. The correct matching, G_i, is printed in a booklet. The sequence is accordingly:

$$S_1 \longrightarrow G_1; \quad S_2 \longrightarrow G_2; \quad \ldots; \quad S_i \longrightarrow G_i,$$

Ten practice sessions were run over two weeks. Two trials on the whole list of syllables or words were given in session nos. 2, 4, 6 and 8. In session nos. 1, 3, 5, 7, 9 and 10 only one trial was given, followed by a test on all the items in the list.

The order of the items in the list was varied in each practice session by alternating versions A and B, regardless of whether the session was characterized by two trials or a trial and a test.

In groups 1 and 3, which practised by the anticipation method, a teaching machine (PAKS 1) was used. It was designed in the Speech Transmission Laboratory of the Royal Institute of Technology in Stockholm. The stimuli were recorded on one channel of a ReVox G36 tape recorder and tone codes which formed response signals to each stimulus were recorded on the other channel. These response signals were interpreted by an electronic tone--code reader connected to the tape recorder. A keyboard with five columns and eight rows (40 positions) was also connected to the tone-code reader. In each position there was a response key and an information panel. Transparent tickets bearing the relevant test for the respective experimental groups were attached to the panel. Below each panel there was a lamp which could be switched on, in order to give information as to the correct alternative (cf. Fig. 3 a and 3 b).

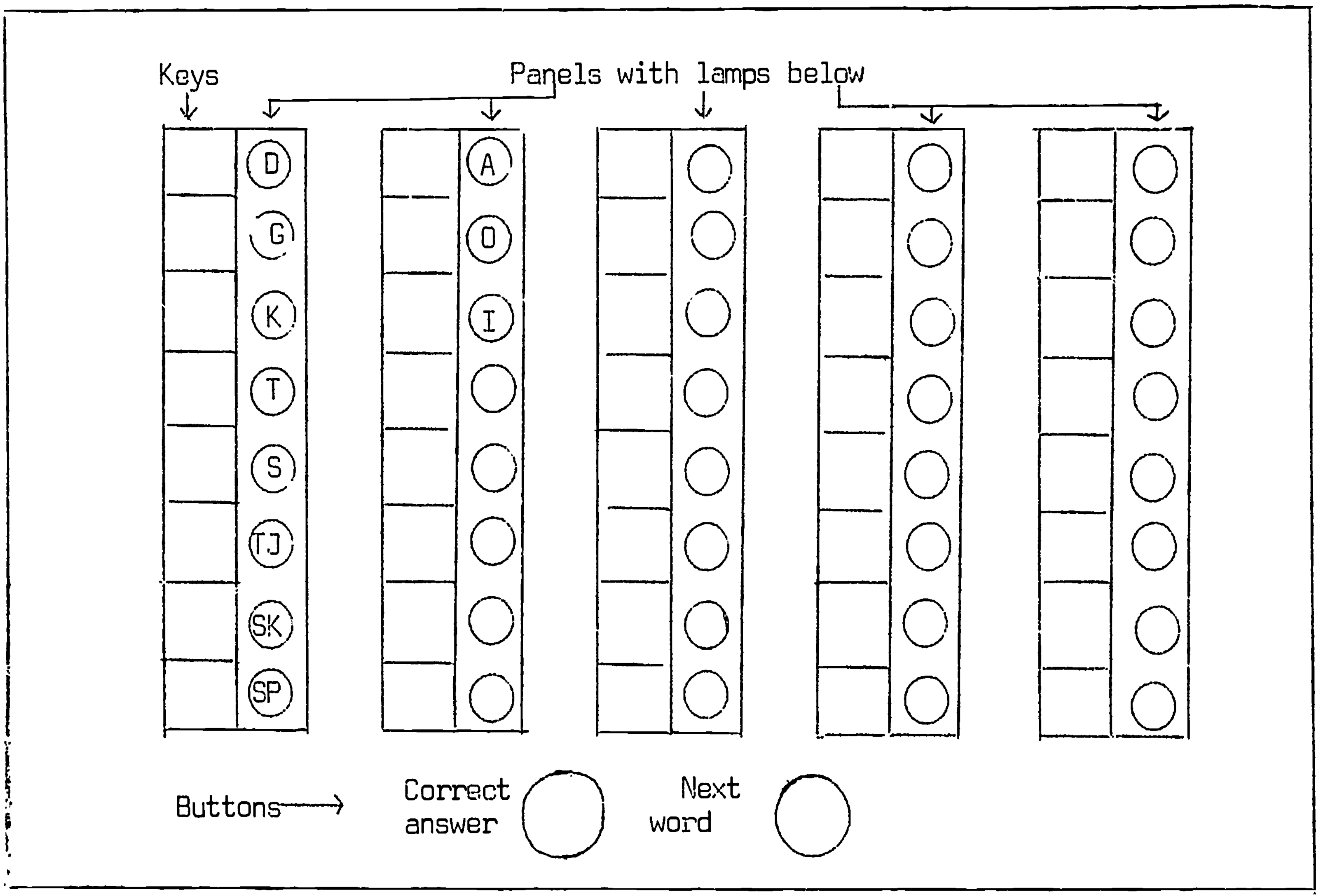

Fig. 3 a. Schematic diagram of keyboard adapted for training with syllables.

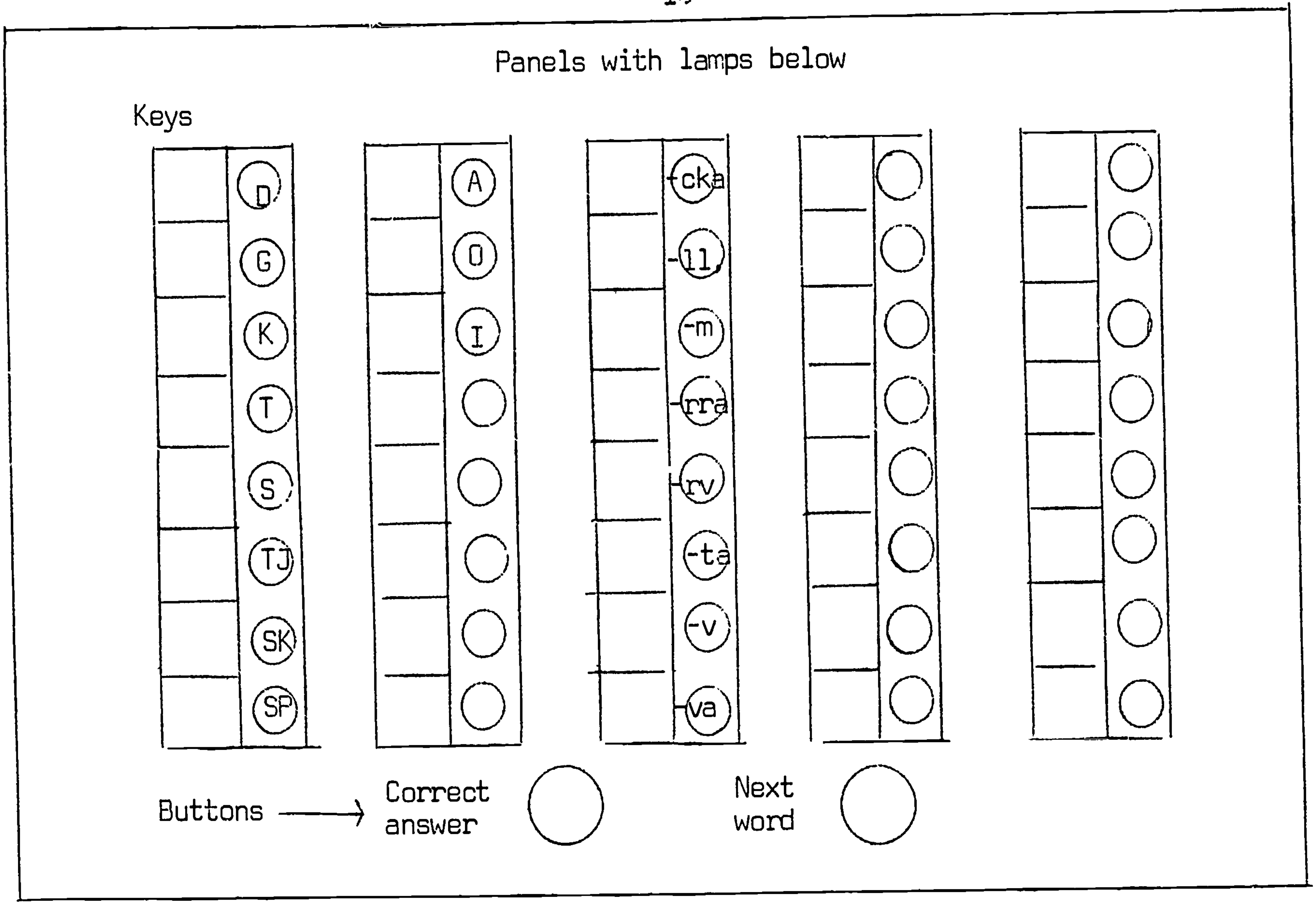

Fig. 3 b. Schematic diagram of keyboard adapted for training with words.

The apparatus functioned as follows. The S was given a stimulus from the tape recorder, which then stopped. The next step was to press the keys bearing the letters which the S considered to be the correct translation of the stimulus. After that he pressed a button marked "correct response" and lamps lit up under the correct keys. Let us suppose that the non-transposed sounds were /ska/ and the S pressed the keys marked sp and o. When he pressed the "correct response" button, the keys marked sk and a, respectively, were lit up. A new stimulus was presented when the S pressed a button marked "next word".

A S practising by the recall method was paced by the S practicing by the anticipation method. In other words, the latter S decided when to give a stimulus. The former S just listened and read the translation of the stimulus in a booklet.

Assessment of transfer effects, motivation and verbal fluency

Before the first trial and after the tenth and last practice session a transfer test was administered (the same in both cases). It consisted of 25 phonetically balanced and meaningful words. The Ss wrote down the translation they considered correct.

At the end of the fifth and the tenth sessions the Ss rated their attitude to the training on a scale containing seven steps from "Very boring" to "Very entertaining". In the last session a test measuring "verbal fluency" was given. The task was to write down in three minutes as many Swedish words as possible beginning with s and ending with a.

The various types of measurements are summarized below.

Before session 1:	Pre-test: Transfer 1
End of session 1:	Test on practice material
End of session 3:	Test on practice material
End of session 5:	Test on practice material
	Attitude test A5
End of session 7:	Test on practice material
End of session 9:	Test on practice material
End of session 10:	Test on practice mateiral
	Verbal fluency
	Attitude test A10
	Post-test: Transfer 2

Results

Changes in the number of correct responses to words and syllables

The mean numbers of correct responses in the tests on the practice material are plotted against the number of test sessions in Fig. 4.

Fig. 4. Learning curves for the experimental groups.

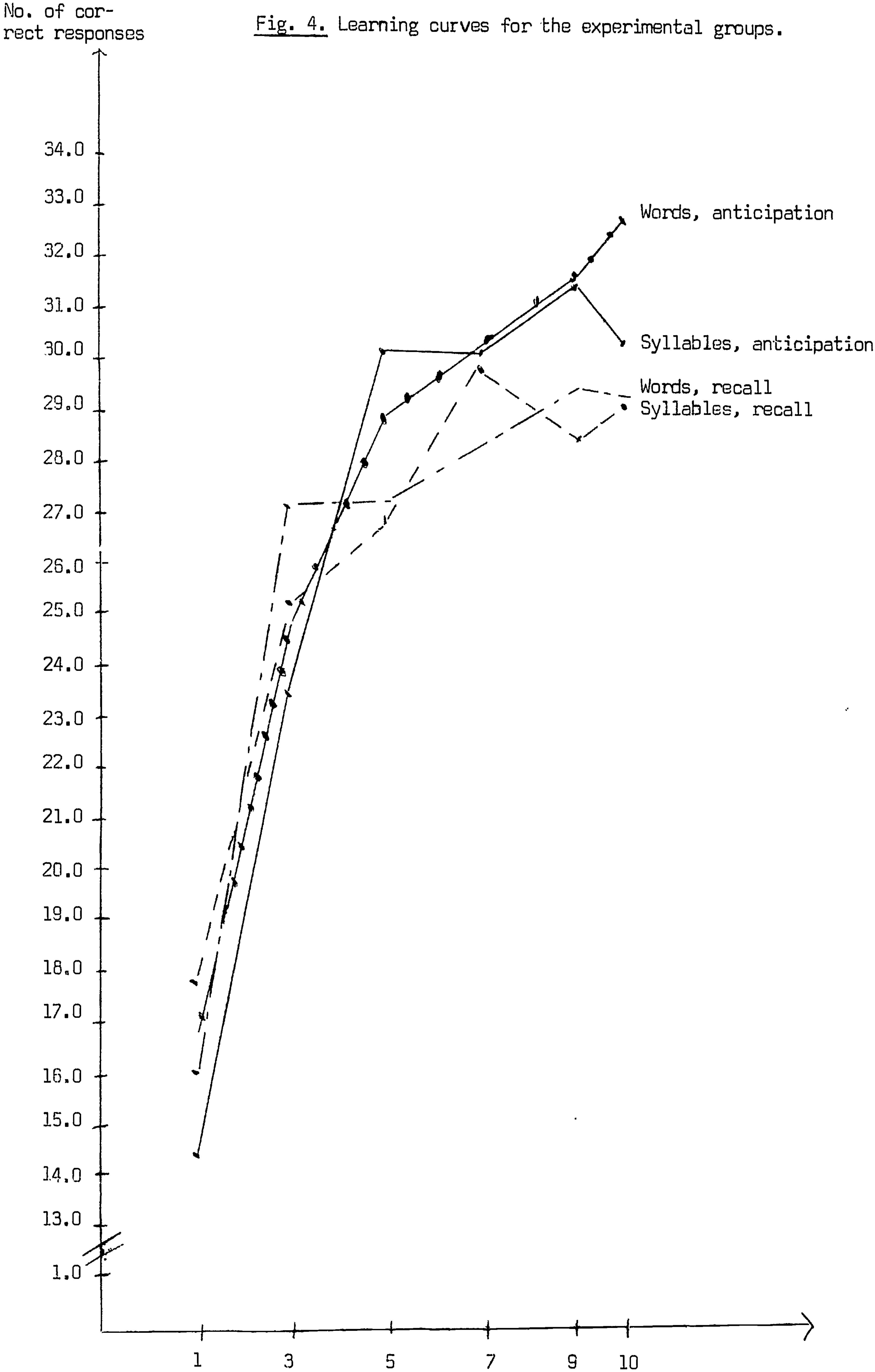

Table 2. Means and standard deviations for the experimental groups (raw scores).

	Test-session 1 M	S	Test-session 3 M	S	Test-session 5 M	S	Test-session 7 M	S	Test-session 9 M	S	Test-session 10 M	S
Words, anticipation	14.4	6.80	23.3	2.47	30.1	3.86	30.0	4.37	31.3	4.41	30.2	5.50
Words, recall	17.8	8.65	25.3	5.09	26.8	5.47	29.3	5.24	23.4	4.54	29.0	4.00
Syllables, anticipation	16.7	6.02	24.8	6.18	28.8	3.58	30.2	4.90	31.4	5.59	32.6	3.83
Syllables, recall	16.0	6.27	27.2	4.63	27.2	3.69	28.2	4.45	29.4	4.28	29.1	4.99

The data summarized in Fig. 4 and Table 2 were subjected to analyses of variance. Since the number of Ss varied between 8 and 11 in the experimental groups, the number of Ss was reduced to 8 by random choice. The results of the analyses are given in Table 3.

Table 3. Analysis of variance of raw scores in the learning tests.

Source of variation	Sums of squares	df	Mean squares	F
Between Ss	1 890.37	31	60.98	
Between training methods (A-R)	4.37	1	4.37	0.07
Between materials (W-S)	19.37	1	19.37	0.29
Interaction (A-R) (W-S)	26.51	1	26.51	0.40
Error between Ss	1 840.12	28	65.72	

To be continued.

Source of variation	Sums of squares	df	Mean squares	F
Within Ss	7 517.83	160	46.99	
Between test sessions (Ts)	4 715.48	5	943.10	29.58xxx
Interaction (A-R) (Ts)	177.66	5	35.53	1.11
Interaction (W-S) (Ts)	24.56	5	4.93	0.15
Interaction (A-R)(W-S)(Ts)	27.02	5	5.40	0.17
Error within	4 463.38	140	31.88	
TOTAL	9 408.20	191		

n = 32, A-R = anticipation-recall, W-S = words-syllables, Ts = test sessions.

Only one of the main effects is significant; there are systematic changes with practice. The mean proportions of correct responses at the first test session vary between 0.29 and 0.37 in the four groups. The corresponding proportions at the last test session vary between 0.60 and 0.68. Accordingly, the proportion of correct responses has nearly doubled at the end of the practice period. Fig. 4 shows, however, that the learning curves are rapidly approaching their asymptotic levels. In other words, it is probable that further practice would have had only slight influence on the performances.

Changes with practice in the identifications of consonant sounds

Tables 4 A - 4 F give the input - output matrices (or confusion matrices) containing the conditional frequencies in percentages for consonant sounds after practice sessions 1, 3, 5, 7, 9 and 10 for groups 1 and 2, practising words. Tables 5 A - 5 F contain the same conditional frequencies for groups 3 and 4, practising syllables.

Table 4: Conditional frequencies for consonant sounds in groups 1 and 2, practising words.

Subjects' responses (columns); Correct responses (rows)

A. Session 1: After 15 min. practice

	d	g	t	k	s	tj	sk	sp	x
d	20	58		1			1		20
g	23	57	1				1		18
t	1	3	29	38	1	1	2	1	24
k			43	52	0.5	0.5	2		2
s	2	2	6	5	29	20	6		30
tj	1		2		40	18	12	1	26
sk			1	2	3		39	37	18
sp							16	61	23

B. Session 3: After 45 min. practice

	d	g	t	k	s	tj	sk	sp	x
d	43	53					1		3
g	36	61	1.5						1.5
t	1.5	1	42	49	1.5	1	1		3
k			47	49		1.5	0.5	0.5	1.5
s	2	1	1	2	55	30	1		8
tj	1	1			56	39			3
sk			2				67	28	3
sp							22	76	2

C. Session 5: After 1 hour 15 min. practice

	d	g	t	k	s	tj	sk	sp	x
d	40	56					0.5		3.5
g	38	60							2
t			52	43			0.5		4.5
k			43	54			2		1
s		3	1		56	36			4
tj	1				55	42			2
sk			2	1			57	38	2
sp					1		17	80	2

	d	g	t	k	s	tj	sk	sp	x
d	45	54				1			
g	42	58							
t			47	53					
k			34	66					
s					70	30			
tj					50	50			
sk			1.5	1.5			64	31	2
sp							1.5	84	0.5

D. Session 7:
After 1 hour 45 min.
practice

	d	g	t	k	s	tj	sk	sp	x
d	52	48							
g	39	61							
t			47	53					
k			34	64	1				1
s					68	31			1
tj			1		49	49			1
sk			1	1			70	28	
sp							10	90	

E. Session 9:
After 2 hours 15 min.
practice

	d	g	t	k	s	tj	sk	sp	x
d	53	47							
g	45	55							
t	1		52	47					
k			31	68					1
s					70	30			
tj					52	48			
sk			1				71	28	
sp							10	90	

F. Session 10:
After 2.5 hours
practice

Table 5: Conditional frequencies for consonant sounds in groups 3 and 4, practising syllables.

Subjects' responses (columns); Correct responses (rows)

A. Session 1: After 15 min. practice

	d	g	t	k	s	tj	sk	sp	x
d	15	25	13	11	4	2	14	9	7
g	11	49	10	2	2	6	7	6	7
t	3	1	33	44	3	1	3	2	10
k	1	5	25	55	4	2	2		6
s	1	7	2	9	23	22	10	7	13
tj	1	2	5	5	29	18	18	10	12
sk			2		2	4	59	22	11
sp	1	3	1	4	3	2	40	35	11

B. Session 3: After 45 min. practice

	d	g	t	k	s	tj	sk	sp	x
d	30	30	12	5	5		10	8	
g	11	62	6	5	2		2	10	2
t	3	1	38	58					
k	2	2	30	62		1		1	2
s		3	1	6	50	37			3
tj			2	3.5	36	45	3.5	5	5
sk					2	1	69	26	2
sp					1	1	36	59	3

C. Session 5: After 1 hour 15 min. practice

	d	g	t	k	s	tj	sk	sp	x
d	52	28	2	4	4		4	6	
g	16	65	11	3	1		2	1	1
t	4	1	45	50					
k	2	1	29	65	2				1
s		1	3	3	55	33			5
tj			3	3	33	56		4	1
sk							82	16	2
sp				1		1	33	64	1

	d	g	t	k	s	tj	sk	sp	x
d	56	19	10	5			5	5	
g	19	70	3	1.5	1.5	1.5	0.5	3	
t	5	1	49	44			1		
k	1	2.5	28	64	1			2.5	1
s			2	4	65	28		0.5	1.5
tj			4		35	55			6
sk							80	20	
sp					1		34	65	

D. Session 7:
After 1 hour 45 min.
practice

	d	g	t	k	s	tj	sk	sp	x
d	53	29	7	5			1	5	
g	19	71	6	0.5				3	0.5
t	1		50	49					
k	3	4	27	65		1			
s	1	1	3	3	61	27	2	1	1
tj			1		40	56	1		2
sk							78	22	
sp							25	75	

E. Session 9:
After 2 hours 15 min.
practice

	d	g	t	k	s	tj	sk	sp	x
d	55	34	6.5		1.5		1.5	1.5	
g	19	69	6	1.5	1		1	1	1.5
t	2	1	40	56					1
k	3	3	25	66		1		2	
s			4	5	63	27	1		
tj			3	1	41	55			
sk							80	20	
sp							34	66	

F. Session 10:
After 2.5 hours
practice

The mean values of the diagonals, i.e. the average numbers of correct identifications of a consonant sound in the matrices, are plotted against the number of practice sessions in Fig. 5.

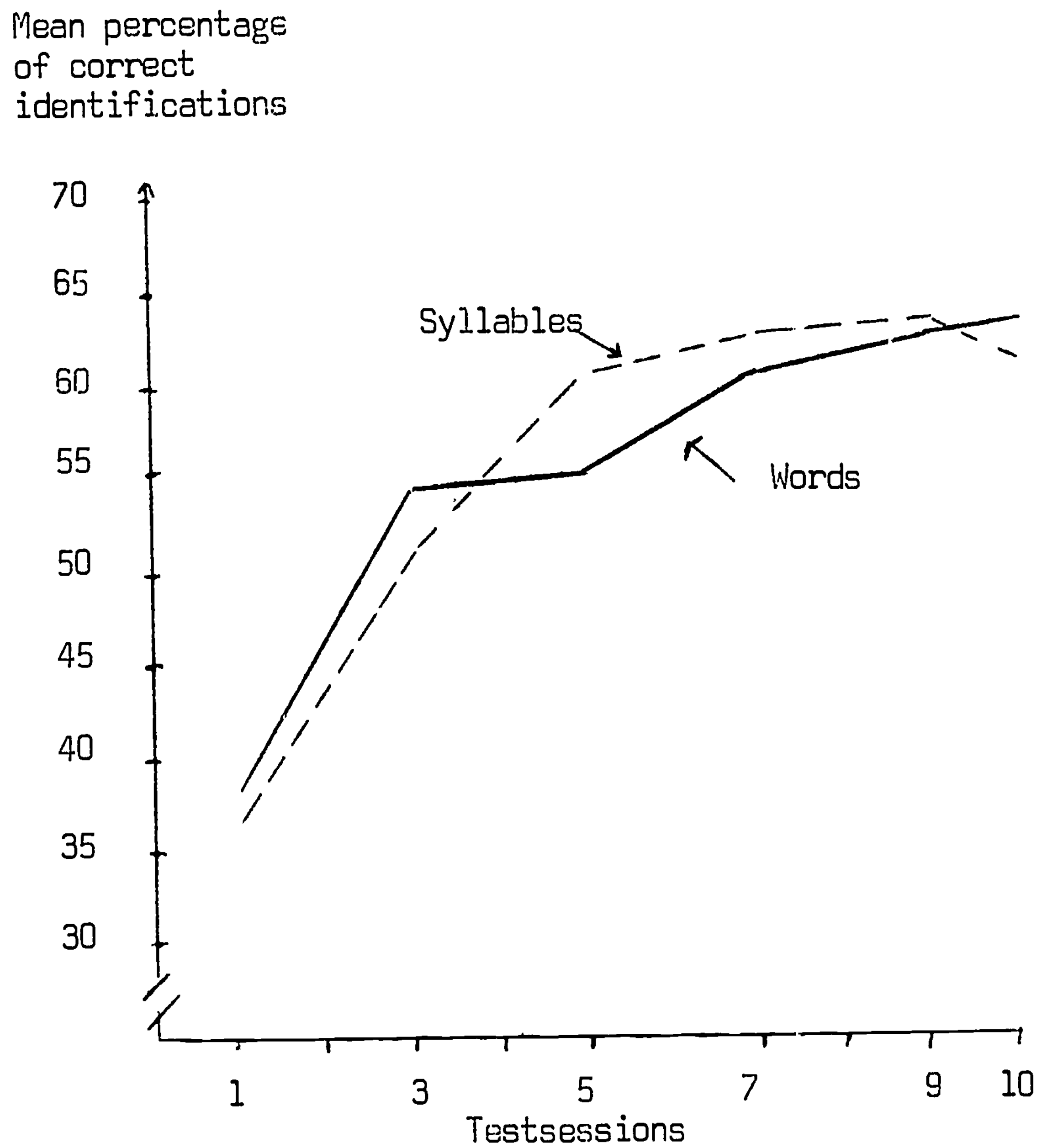

Fig. 5. The average numbers (percentages) of correct identifications of a consonant sound.

It is evident from Tables 4 and 5 and Fig. 5 that the two materials are about equal in difficulty and that the identification of a consonant sound generally improves with practice. Some differences between the matrices for the word groups and syllable groups should, however, be commented on.

The x category contains responses other than the possible eight syllables or words. The number of wrong responses is higher in the word groups at the initial stage of learning, probably because it is easier to learn what the alternative syllables are than what the alternative words are. After session 5 the frequencies in this category are about the same for both materials.

The matrices corresponding to the tests at practice level 10 show some small differences between the materials. A plausible explanation of this is that the Ss´ responses to the stimuli are influenced by their relative frequencies in the Swedish language, i.e. that a S prefers to guess at the most frequent word when uncertain how to discriminate between two words. The words "sparv" (sparrow) and "sporra" (spur) are, for instance, more frequent in Sweden than "skarv" (splice) and "skorra" (grate). A S who cannot discriminate between /sp/ and /sk/ is more apt to guess at /sp/ than /sk/ in the word material, where sp is a more frequent response to /sk/ than sk is to /sp/. The opposite is true among the syllables, where /ska/ and /sko/ are meaningful words but /spa/ and /spo/ are not.

Changes in discriminations between consonant sounds

Parallel to the changes with training in the identification of consonant sounds, changes may occur in the discriminations between sounds. Tools for quantifying the degree of discrimination in each of the input-output matrices are provided by information theory (cf., for instance, Attneave, 1959). The S is regarded in the following discussion as a channel with an input and an output of information.

Let $\hat{H}(x)$ be the amount of information at the input or the average uncertainty per element with reference to the classification $\underline{x}$. An estimate of H(x) in "bits" is

$$\hat{H}(x) = \overset{i}{\Sigma}\, p_i \quad \log_2 \frac{1}{\hat{p}_i} = \log_2 n - \frac{1}{n} \overset{i}{\Sigma}\, n_i \quad \log_2 n_i.$$

The corresponding estimate of the amount of information at the output is

$$\hat{H}(y) = \overset{i}{\Sigma}\, p_j \log_2 \frac{1}{\hat{p}_j} = \log_2 n - \frac{1}{n} \overset{i}{\Sigma}\, n_j \log_2 n_j.$$

The estimate of the amount of information necessary to locate an element in both classifications, or the joint uncertainty, is

$$\hat{H}\,(x,y) = \overset{i,j}{\Sigma}\, \hat{p}_{ij} \quad \log_2 \frac{1}{\hat{p}_{ij}} = \log_2 n - \frac{1}{n} \overset{i,j}{\Sigma}\, n_{ij} \quad \log n_{ij}.$$

A measure of contingency between input and output is

$$\hat{T}(x;y) = \hat{H}(x) + \hat{H}(y) - \hat{H}(x,y).$$

In other words, T(x;y) measures the reduction in the amount of information required to locate an element in one calssification, if it has been located in the other. It is the amount of information transmitted from stimulus to response. When there is no relation whatever between the S's responses and the stimuli, $\hat{T}(x;y) = 0$ "bits". The maximum $\hat{T}(x;y)$, given a certain input, is equal to $\hat{H}(x)$, It should be noted that $\hat{T}(x;Y)$ does not reflect the degree of correspondence with an externally arranged pairing of stimuli and responses. Transmission may "... be perfect even when some or all responses are 'wrong' by external standards, provided the subject is perfectly consistent in his 'errors' and maintains his own one-to-one correspondences between responses and stimuli" (Attneave, 1959,p, 50).

A confusion matrix was established for each S at each level of practice (session nos. 1, 3, 5, 7, 9 and 10) and $\hat{T}(x;y)$ was found for each matrix. The means and standard deviations in the four experimental groups at each test session are given in Table 6.

Table 6. Means and standard deviations of $\hat{T}(x;y)$. n = 38.

Test session	Words				Syllables			
	Anticipation		Recall		Anticipation		Recall	
	M	S	M	S	M	S	M	S
1	1.700	0.243	1.527	0.606	1.427	0.848	1.504	0.938
3	2.062	0.128	2.034	0.098	1.812	0.500	1.785	0.087
5	1.997	0.134	2.017	0.099	1.973	0.686	1.886	0.387
7	2.130	0.100	2.056	0.093	1.989	0.424	1.976	0.663
9	2.171	0.137	2.135	0.033	2.060	0.500	1.950	0.469
10	2.199	0.[illegible]20	2.121	0.024	1.996	0.480	1.946	0.489

An analysis of variance with the same eight Ss in each group as in the earlier analysis gave the results summarized in Table 7.

Table 7. Analysis of variance of $\hat{T}(x;y)$ values.

Source of variation	Sums of squares	df	Mean squares	F
Between Subjects	3.665	31	0.118	
Between A and R	0.171	1	0.171	2.036
Between W and S	1.150	1	1.150	13.690^{xxx}
Interaction (A-R)(W-S)	0	1	0	0
Error variance between Ss within groups	2.344	28	0.084	

(To be continued)

Source of variation	Sums of squares	df	Mean squares	F
Within Ss	11.255	160	0.070	
Between testsessions (TS)	6.333	5	1.267	21.584xxx
1. Interaction (A-R)(TS)	0.016	5	0.003	0.051
2. Interaction (W-S)(TS)	0.136	5	0.027	0.460
3. Interaction (A-R)(W-S)(TS)	0.042	5	0.042	0.716
4. Error variance	8.224	140	0.0587	
Total	14.920	191		

n = 32 A-R = Anticipation-Recall W-S = Words-Syllables TS = Testsessions

Only two main effects were significant ($p < 0.001$); between test sessions and between words and syllables. Thus, the degree of discrimination increased significantly with training and the groups that trained with words discriminated significantly better between consonant sounds than the groups that trained with syllables.

A closer inspection of the matrices in Tables 4 and 5 reveals that the primary basis for discrimination between the consonant sounds is the mode of articulating them. Tables 8 and 9 below are identical with Tables 4 F and 5 F. Here the mode of articulation is marked with a thick line, while the place of articulation is marked with a dashed line through the relevant cells. It will be seen that sounds with the same mode of articulation cannot be discriminated from each other (cf., for instance, the sounds /d/ and /g/), while sounds with different modes of articulation are almost perfectly discriminated from each other.

Table 8. Percentual results of the word groups in the final test.

	d	g	t	k	s	tj	sk	sp	x
d	53	47							
g	45	55							
t	1		52	47					
k			31	68					1
s					70	30			
tj					52	48			
sk			1				71	28	
sp							10	90	

Table 9. Percentual results of the syllable groups in the final test.

	d	g	t	k	s	tj	sk	sp	x
d	55	34	6.5		1.5		1.5	1.5	
g	18	69	6	1.5	1		1	1	1.5
t	2	1	40	56					1
k	3	3	25	66		1		2	
s			4	5	63	27	1		
tj			3	1	41	55			
sk							80	20	
sp							34	66	

In order to relate the changes in discrimination to specific acoustical characteristics, presumably discriminated by the listener, we shall adopt the same system of articulatory features as was utilized by Miller and Nicely (1955) in their study of perceptual confusions among consonants under various conditions of noise, low-pass filtering and high-pass filtering. We know from their study that place of articulation and duration are disturbed by low-pass filtering, while voicing and nasality are resistant to disturbances by low-pass filtering down to about 600 c/s. Accordingly, it would be interesting to find out if this is true even in transposition.

Table 10. The code dimensions of the consonant sounds.

Consonant sound	Voiced (1) Voiceless (0)	Duration Long (1) Short (0)	Mode Front (1) Back (2)	Clustered (1) Single (0)
s	0	1	1	0
tj	0	1	2	0
t	0	0	1	0
k	0	0	2	0
d	1	0	1	0
g	1	0	2	0
sp	0	1	1	1
sk	0	1	2	1

Four sets of confusion matrices were constructed, each consisting of two matrices. The sets corresponded to the code dimensions. The first matrix in a set contained the responses at session 1 and the other matrix the responses at session 10. The matrices are presented in Tables 11 A and B, 12 A and B, 13 A and B and 14 A and B.

Table 11. Voicing. A. Relative frequencies in percentages at session 1.
B. Relative frequencies in percentages at session 10.

	0	1	Σ
0	96.68	3.32	100.00
1	32.99	67.01	100.00

Voicing. A.

	0	1	Σ
0	99.16	0.84	100.00
1	5.40	94.60	100.00

Voicing. B.

Table 12. Duration. A. Relative frequencies in percentages at session 1.
B. Relative frequencies in percentages at session 10.

	0	1	Σ
0	89.09	10.91	100.00
1	9.76	90.24	100.00

Duration. A.

	0	1	Σ
0	98.56	1.44	100.00
1	1.45	98.55	100.00

Duration. B.

Table 13. Mode. A. Relative frequencies in percentages at session 1.
B. Relative frequencies in percentages at session 10.

	1	2	Σ
1	46.98	53.02	100.00
2	41.97	58.03	100.00

Mode. A.

	1	2	Σ
1	63.35	36.65	100.00
2	34.86	65.14	100.00

Mode. B.

Table 14. Cluster consonants. A. Relative frequencies in percentages at session 1. B. Relative frequencies in percentages at session 10.

	0	1	Σ
0	89.58	10.42	100.00
1	8.31	91.69	100.00

Cluster consonants. A.

	0	1	Σ
0	99.33	0 67	100.00
1	0.25	99.75	100.00

Cluster consonants. B.

At the first session the voiceless consonant sounds (0) were perceived as voiceless in 96.68 per cent of all cases and as voiced in 3.32 per cent. After practice the corresponding figures were 99.16 per cent and 0.84 per cent respectively. In other words, practice sharpened discrimination. If we consider the voiced sounds (1), the changes are even greater.

Clear-cut changes with training are found in all code dimensions, but the _Ss_´ ability to discriminate on the basis of place of articulation is poor, even after ten sessions of training.

The improvement in transmission of duration found here may be due to the frequency transposition and may be also to the amplitude compression.

Transfer effects

A test consisting of phonetically balanced words was administered before the first trial and after the last trial. The mean performance at the pre-test was $\underline{M} = 5.0$ and at the post-test $\underline{M} = 7.5$. In order to study if all four groups were characterized by the same increments, an analysis of covariance was run partiallying the pre-test scores from the post-test scores. The product moment correlation between pre-test and post-test in the pooled groups is 0.59. The results will be found in Table 15.

Table 15. Analysis of covariance of transfer data. Pre-test kept constant.

Source of variation	Sums of squares	df	Mean squares	*F*
Between groups	1.27	3	0.42	-
Within groups	159.48	33	4.83	
Total	160.75	36		

The differences between the groups, as regards changes in the transfer test, with practice, are not significant. Accordingly, it is impossible to tell if the overall change in mean performance is an effect of real transfer. It may also be a result of adjustment to the testing situation, for instance, a warming-up effect.

The product moment correlation for the pooled groups between the two attitude effects is 0.60. The correlation between the attitude test and the number of correct responses at session 5 is r = 0.09 and at session 10 r = 0.16.

Attitudes to training

All the Ss rated their attitudes to training at the end of sessions 5 and 10. The differences between the mean values in the four groups on each occasion were tested by analysis of variance. The results are presented in Tables 16 and 17.

Table 16. Analysis of variance of attitude-test data at session 5.

Source of variation	Sums of squares	df	Mean squares	*F*
Between groups	17.98	3	5.99	12.22^{xx}
Within groups	16.32	33	0.49	
Total	34.30	36		

The differences are significant ($p < 0.01$). By pairwise comparison of each group mean with every other and *t*-testing the differences, it was found that the attitudelevels were highest under the condition of anticipation of syllables and lowest under the condition of recall of words.

Table 17. Analysis of variance of attitude-test data at session 10.

Source of variation	Sums of squares	df	Mean squares	*F*
Between groups	7.31	3	2.44	1.78
Within groups	45.19	33	1.37	
Total	52.5	36		

In this case no differences between groups could be found. The product moment correlation between the two attitude measurements is 0.60 in the pooled groups. The correlations between the attitude test and the number of correct responses at session 5 and session 10 are 0.09 and 0.16 respectively. Since none are significantly greater than $r = 0$, it is not probable that the *Ss*´ attitudes have influenced the performances.

Verbal fluency

The verbal-fluency task consisted in writing down in three minutes as many Swedish words as possible beginning with s and ending with a. The correlations between this test and a sample of the other variables are reported in Table 18.

Table 18. Correlations between verbal fluency and other variables.

	Attitude		Transfer		Practice		
	5	10	1	10	1	5	10
Verbal fluency	0.15	0.20	0.28	0.42^{xx}	-36^{x}	0.04	0.16

Verbal fluency is positively correlated ($p < 0.01$) with the post-test on transfer. A slight positive relation is also noticeable at the pre--test. The reason for this may be that Ss with high verbal fluencies are more able to utilize the perceived cues for guessing according to the statistical structure of Swedish words. It may seem puzzling, however, why verbal fluency is negatively related to the number of correct responses at the first session of training.

The reasons may be that the Ss were not informed at test session 1 that the test stimuli were the same as the stimuli just presented in the preceding trial. Accordingly, Ss with high verbal fluencies may have relied on the same guessing habits as in the transfer test. If this were the case, they would be mislead, because the statistical structure of the syllables and words is not the same as the structure of Swedish words in general.

Summary and discussion

The consonant sounds most easily identified by the Ss at the first test session were /g/, /k/, /sp/ and /sk/. The sounds /d/ and /ç/ were the most difficult to identify. This is demonstrated in Fig. 6, in which the increases and decreases in the relative frequencies of identifications from session 1 to session 10 are plotted against the frequency with which a given sound was identified at session 1.

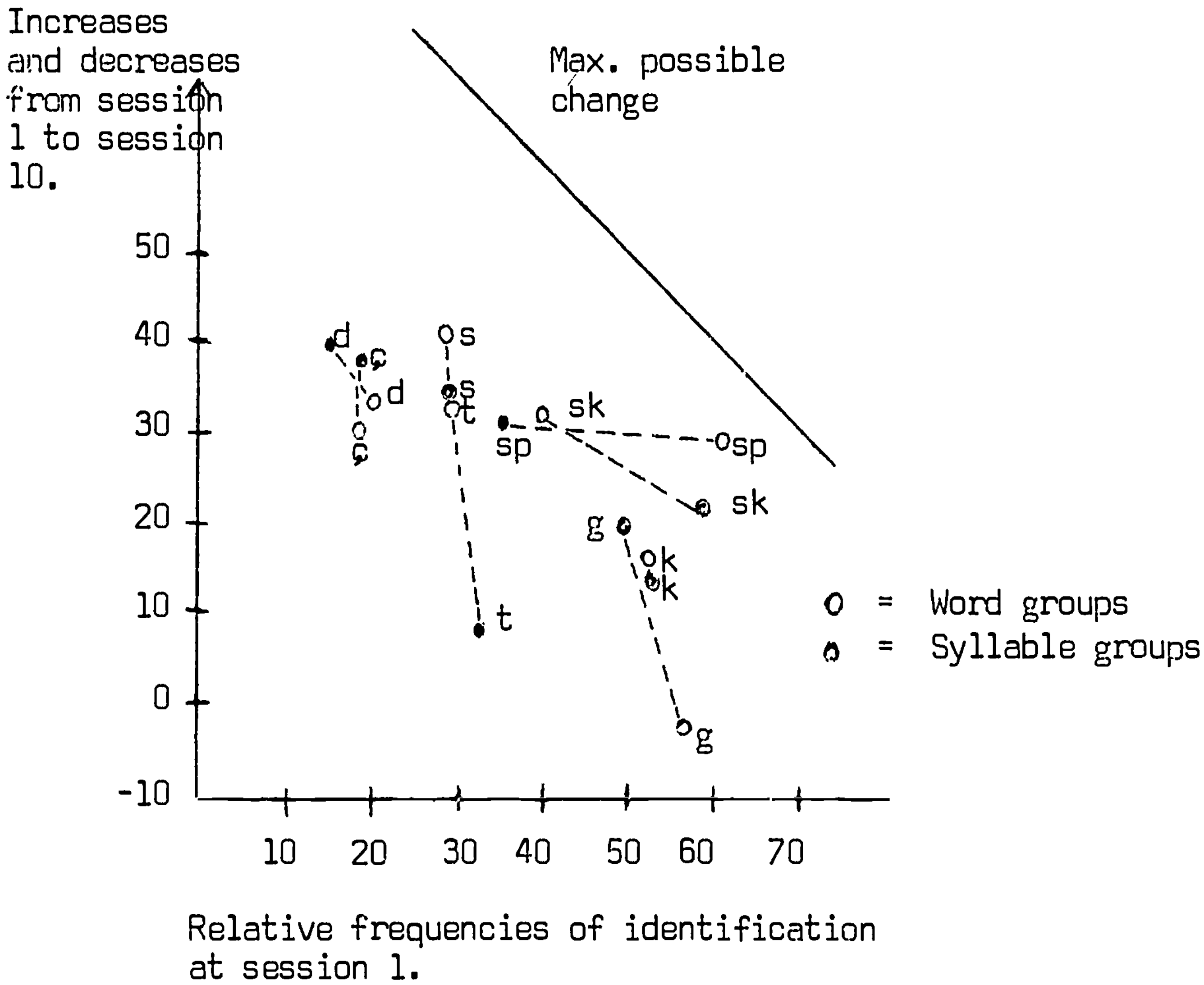

Fig. 6. Effects of training on identifications of consonant sounds. The greatest improvements with practice in relation to the maximum possible gain are found in the sounds /sp/, /sk/ and /s/.

The changes in sound discrimination which take place with training occur rapidly and it is doubtful whether further training would have had any appreciable effect on discrimination. The initial value of the amount of information transmitted is in the range 1.4 - 1.7 "bits". The primary reason for this is probably that the Ss had normal hearing. A S with normal hearing is able to recognize the low-frequency sounds which have not been transposed, and it is possible that he can shift his reference system when the frequency range of the fricatives is shifted. For a person with serious impaired hearing since birth, the initial amount of transmitted information would be close 0 "bits", because he has no reference system and thus all sounds are new to him.

An interesting point is that the Ss could learn to rely in their discriminations on the code dimensions voicing, duration and cluster consonant but not on place of articulation. A training experiment should be performed in which transposition with and without compression is compared with low-pass filtering, in order to record more specifically the factors which determine the perception of this code dimension.

No standpoint can be adopted as to which of the two training methods - anticipation and recall - compared here is the more effective. Anticipation seems to b. somewhat - but insignificantly - more effective (cf. Fig. 4). Other studies with paired-associate materials show that the anticipation method is more effective when the stimuli are difficult to code, and the recall method is more effective when the number of responses exceed the memory span of the S (Barch and Levine, 1967). In other words, the outcome of a comparison between the two training methods may depend on the number of stimuli, the type of stimuli, the cut-off frequency of the low-pass filter and, in the case of Ss with impaired hearing, the degree of hearing impairment and linguistic competence.

It does not seem to matter which of the two training materials - words or syllables - is used. But syllables appear to be somewhat - though insignificantly - easier to learn.

The results of the training sessions vary with the S´s degree of verbal fluency. A S with high scores in the S-A test tends to get high scores in the training sessions. In future experiments different verbal abilities could be combined in order to find out which of them are relevant for this type of training experiment.

REFERENCES

Attneave, Fred: Applications of Information Theory to Psychology. Holt, Rinehart and Winston, Inc., U.S.A. 1959.

Barch, A.M. and Levine, J.R.: Presentation Method in Auditory Identification Learning. Journal of Verbal Learning and Verbal Behavior, vol. 6, 1967, 282 - 288.

Edwards, A.L.: Experimental Design in Psychological Research. Holt, Rinehart and Winston, Inc., U.S.A. 1965.

Johansson, B.: Proceedings of Third International Congress on Acoustics. Stuttgart 1959, vol. 2, 655 - 57.

Johansson, F. and Sjögren, H.: The Effect on Normal Hearing Subjects and Profoundly Deaf Children by Coding Voiceless Consonants Using Frequency Transposition. Paper B69 presented at the Fifth International Congress on Acoustics, Liège, 1965.

König, F. and Eichler, H.: Ein neues Verfahren zur Hörverbesserung bei hochgradiger Perceptionsschwerhörigkeit durch Hörapparate. Archiv Ohren- usw. Heilk, u.Z. Hals- usw. Heilk. 165, 1954, 326 - 331.

Lindqvist, E.F.: Design and Analysis of Experiments in Psychology and Education. Houghton Mifflin Company, U.S.A. 1953.

Ling, Daniel: Three Experiments on Frequency Transposition. Speech held at The Conference on Speech Analyzing Aids for the Deaf, June 1967, at the Hearing and Speech Center, Graduate School, Gallaudet College, Washington D.C.

Miller, F. and Nicely, P.: An Analysis of Perceptional Confusions Among Some English Consonants. The Journal of the Acoustical Society of America, vol. 27, No. 2, March 1955.

Oeken, F.W.: Frequenztransposition zur Hörverbesserung bei Innenohrschwerhörigkeit. Archiv Ohren- usw. Heilk. u.Z. Hals- usw. Heilk. 181, 1963 418 - 425.

Pimonow, L.: L'application de la parole synthétique dans la corrective auditive. Acustica 12, 1962, 73 - 82.

Wedenberg, Erik: Auditory Training of the Severly Hard of Hearing Using a Coding Amplifier. Proceedings of the Third International Congress on Acoustics. Stuttgart 1959.