IZF 1988-3

H.J.M. Steeneken
F.W.M. Geurtsen

DESCRIPTION OF THE RSG-10 NOISE DATA-BASE

CONTENTS

Rep.No. IZF 1988-3 TNO Institute for Perception,
Soesterberg, The Netherlands

Description of the RSG-10 noise data-base

H.J.M. Steeneken and F.W.M. Geurtsen

SUMMARY

Research on the effect of noise on speech recognition performance and speech communication quality requires a calibrated set of representative noise signals.
Based on a total set of 43 noise signals from military situations, a limited set of noises was selected for a data-base. The selection was done by means of a factor analysis, based on the 1/3-octave spectra of the noises.
It is shown that with 18 different noise signals (including some artificial thermal noise signals) a representative data-base can be obtained. These noise signals were recorded on PCM-tape and calibrated carefully.

Beschrijving van de RSG-10 data-base voor stoorlawaai

H.J.M. Steeneken en F.W.M. Geurtsen

SAMENVATTING

Bij onderzoek inzake de invloed van stoorlawaai op de spraakcommunicatie of op automatische spraakherkenning bestaat behoefte aan een data-base van representatieve "militaire" stoorgeluiden.
Op basis van 43 beschikbare stoorgeluiden van verschillende lawaaicondities bij de drie krijgsmachtdelen, werd een keuze gemaakt voor toepassing in de data-base. Met een factoranalyse, uitgevoerd op de 1/3-octaaf spectra van deze stoorgeluiden, werd bepaald in hoeverre met een beperkt aantal van deze stoorgeluiden een zo representatief mogelijk databestand kan worden samengesteld.
Inclusief enige artificiële ruissignalen blijkt dat met 18 verschillende stoorgeluiden mogelijk te zijn. Deze signalen zijn opgenomen in de data-base en geregistreerd op een PCM-tape. De data-base werd nauwkeurig gekalibreerd en is beschikbaar voor algemene toepassing.

1 INTRODUCTION

One of RSG-10's AC/243 (Panel 3) RSG-10 interests is the evaluation of automatic speech recognizers and speech communication channels in a military situation such as in a high noise environment. In order to obtain comparable test conditions in the different nations a noise data-base has been established. The data-base consists of a limited set of noises representative of the military environment.
For this purpose a number of nations supplied samples of noise recordings from various noise sources such as: jet-planes, helicopters, wheel carriers, tanks, and command rooms. From these noise recordings a first selection was made with respect to: recording quality, duration, type of noise, and stability. Thirty-eight different noise samples were selected as being representative of military conditions. The samples had a duration of at least four minutes. The original recordings were made on analogue tape, with a wide-band omni-directional microphone. The noise samples, obtained through this procedure can be played back and used as acoustical noise in a talker or listener environment, or electrically mixed with a speech signal.

2 DATA-BASE SELECTION

Besides the 38 selected noise samples a set of five well defined thermal noise signals have also been included.
From the 43 noise samples the 1/3-octave spectrum was measured with a spectrum analyzer (Rhode and Schwarz, type FAR). This analyser performs a 1/3-octave band analysis from the 1/3-octave band with centre frequency 80 Hz to the 1/3-octave band with centre frequency 10 kHz. The analyser was used on fast response which is comparable with a time constant of 100 ms. Five seconds of each noise sample was analysed and for each 100 ms interval a 1/3-octave spectrum was measured, hence fifty spectra were obtained.
From these spectra the equivalent level, L_{eq} (based on the power average) and the average level together with the standard deviation (of the fifty dB-values) was calculated for each 1/3-octave band and for the A-weighted and linear channel. The linear channel accounts for the wide-band signal level while the A-weighted channel includes a spectral weighting related with auditory perception. For each channel an amplitude-distribution histogram was generated, based on the 50 samples. From these histograms the L_{10} and L_{90} levels were obtained (L_{10} is the level that is exceeded during 10% of the time and similarly L_{90} is the level exceeded 90% of the time). The difference L_{10}-L_{90} is a measure for the level fluctuations.

Table I. Description of the noise samples from which the data-base was selected.

No.	Source	Remarks
1	White noise	equal energy per unit of bandwidth
2	Pink noise	equal energy per log unit of bandwidth
3	White -6 dB/oct	spectrally shaped from 250 Hz
4	White -12 dB/oct	spectrally shaped from 250 Hz
5	Speech noise	according to long-term speech spectrum
6	Destroyer 1	Engine room 130 rpm 101 dBA
7	Destroyer 2	Gear box 70 rpm 101 dBA
8	Destroyer 3	Gear box 146 rpm 101 dBA
9	Destroyer 4	Turbines 146 rpm 93 dBA
10	Destroyer 5	Boiler room 90 dBA
11	Destroyer 6	Sonar control 73 dBA
12	Destroyer 7	Operations room 70 dBA
13	M 113 1	Cabin 50 km/h 106 dBA
14	M 113 2	Driver 50 km/h 105 dBA
15	M 113 3	Cabin 40 km/h 105 dBA
16	M 113 4	Driver 40 km/h 106 dBA
17	AVGP 1	Cabin 50-60 km/h 90 dBA
18	AVGP 2	Driver 50-60 km/h 96 dBA
19	M 109 1	Driver 30 km/h 111 dBA
20	M 109 2	Gunner 30 km/h 110 dBA
21	M 109 3	Gunner 58 km/h 115 dBA
22	M 109 4	Driver 58 km/h 117 dBA
23	T 33 jet 1	200-300 Knots 108 dBA
24	Buccaneer 1	Pilot 295 Knots 200 Feet 109 dBA
25	Buccaneer 2	395 Knots 300 Feet 109 dBA
26	Buccaneer 3	300 Knots 250 Feet 108 dBA
27	Buccaneer 4	394 Knots 300 Feet 110 dBA
28	Buccaneer 5	449 Knots 300 Feet 114 dBA
29	Buccaneer 6	450 Knots 300 Feet 116 dBA
30	Buccaneer 7	502 Knots 250 Feet 114 dBA
31	Buccaneer 8	540 Knots 250 Feet 117 dBA
32	Buccaneer 9	540 Knots 250 Feet airbrakes out 117 dBA
33	Buccaneer 10	190 Knots 1000 Feet airbrakes out 109 dBA
34	Lynx 1	On platform 97 dBA
35	Lynx 2	Stationary flight 100 Knots 2000 Feet 99 dBA
36	Lynx 3	Stationary flight 100 Knots 2000 Feet 99 dBA
37	Lynx 4	Hoover door closed 100 dBA
38	Lynx 5	Hoover door open 103 dBA
39	Leopard 1 1	70 km/h 104 dBA
40	Leopard 1 2	40 km/h 99 dBA
41	Leopard 2 1	40 km/h 108 dBA
42	Leopard 2 2	70 km/h 114 dBA
43	Machine gun	Caliber 0.50

In order to select a representative limited set of noises from the total set of 43 noises we performed a principal-component analysis based on 22 1/3-octave levels. The principal-component analysis was based on a variance-covariance matrix. The 43 spectra were normalized with respect to the average of the 22 1/3-octave bands, so the principal-component analysis is based on the shape of the frequency spectrum of the different noises. The 43 noises from which the selection was made are given in Table I.

The original recording of these noises were made by:

- Royal Aircraft Establishment, Farnborough, England;
- Defence and Civil Institute for Environmental Medicine, Downsview, Canada;
- National Aeronautical Establishment, National Research Council, Ottawa, Canada;
- TNO Institute for Perception, Soesterberg, The Netherlands.

In Fig. 1 the spectral relation between the different noise samples is given in a two-dimensional space after a principal-component analysis. Dimension I explains 72.7% of the total variance, dimension II explains 11.6% of the total variance. The data-points are labelled according to the numbering of the noises in Table I.

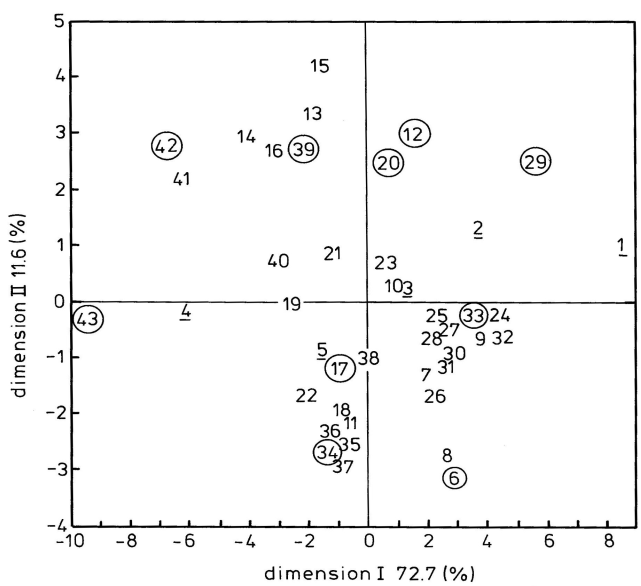

Fig. 1 Two-dimensional subspace for the original 38 military and 5 artificial noises (underlined). The data

In Fig. 2 the factor pattern of the transformation of the original 22 dimensions to a two-dimensional subspace is given. Dimension I clearly indicates a selection between "high" and "low" frequencies (the original dimension 1 corresponds to the 1/3-octave band with centre frequency 80 Hz and original dimension 22 with 1/3-octave band with centre frequency 10 kHz). Dimension II separates between the mid-frequencies and the lower and upper part of the spectrum.

The encircled data points in Fig. 1 are the selected "military" noises for the noise data-base. For this selection the following considerations were taken into account:

- the selected noises must be fairly well distributed in the space given in Fig. 1;
- for all representative military situations given in Table I at least one noise sample should be selected;
- all artificial noises (no. 1-5) are useful for calibration purposes and for a simple experimental design.

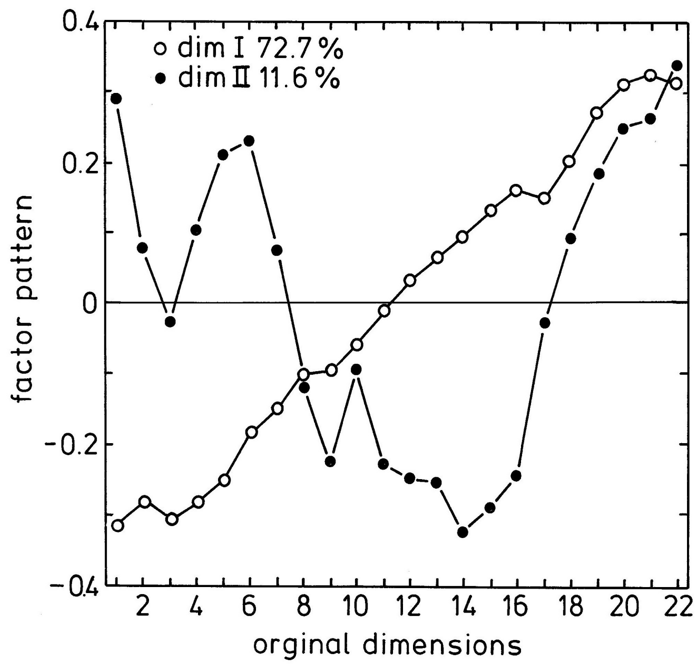

Fig. 2 Factor pattern from the principal component analysis of the noises of Table I for transformation to a two-dimensional space.

The sample noises selected on the basis of these criteria are listed in Table II. A sinusoidal signal (1000 Hz) has been added for the purpose of level definition. Two noises were added after the selection of the noise data-base:

- noise from a HF radio channel (required for studies on HF communication links);
- STITEL test signal as a reproducible voice babble: this time-varying signal has an equal long-term frequency spectrum as noise type number 6.

Table II Description of the noise samples selected for the RSG-10 Noise Data-base.

no.	source	description	Table I	tape counter
1	Sinusoid	1000 Hz, 0 dB		0'00"
2	Pink noise	equal energy per 1/3-oct	2	1'00"
3	White noise	equal energy per Hz bandwidth	1	5'20"
4	White -6 dB/oct	cut off freq. 250 Hz, -6 dB/oct	3	9'30"
5	White -12 dB/oct	cut off freq. 250 Hz, -12 dB/oct	4	13'40"
6	Speech noise	average speech spectrum	5	17'50"
7	M 109	30 km/h 110 dBA	20	22'00"
8	Buccaneer	Pilot 190 Knots 1000 Feet 109 dBA	33	26'10"
9	Leopard 2	70 km/h 114 dBA	42	30'20"
10	Wheel carrier	50-60 km/h 90 dBA	17	34'30"
11	Buccaneer	450 Knots 300 Feet 106 dBA	29	38'40"
12	Lynx	Platform 97 dBA	34	42'50"
13	Leopard 1	70 km/h 104 dBA	39	47'00"
14	Operation room	opsroom of destroyer 70 dBA	12	51'10"
15	Destroyer	engine room 101 dBA	6	55'20"
16	Machine gun	calibre 0.50 repeated	43	59'30"
17	HF radio	noise from HF radio channel	-	1h04'00"
18	STITEL	STI test signal	-	1h08'00"

3 DATA-BASE CALIBRATION

The selected noises and calibration signals are available on a digital PCM tape. This tape contains a 4-minute sample of each signal. From the recorded signals the spectrum and overall levels were measured as described before. These spectral specifications are given in Appendix B.

The levels are relative to the RMS level of the sinusoidal signal, which makes level adjustment quite easy.

The pink noise signal can be used to verify the frequency response of the play-back equipment. A flat 1/3-octave spectrum should be obtained. This method of frequency response verification is in accordance with the guidelines given by RSG-10 in document AC/243 (Panel 3) D/243 dated March 14, 1985.

In order to verify the selection of the noises in this data-base a second principal-component analysis was performed on the normalized spectra. A two-dimensional representation of the results is given in Fig. 3 for dimension I and dimension II. For all noises the 1/3-octave spectrum was determined twice on two different noise samples. Each sample with a duration of five seconds taken from the first minute of the noise recording and of the last minute of the noise recording. The 1/3-octave spectra were based on the L_{eq} of 50 spectra, sampled every 100 ms. The corresponding datapoints are labelled according to the identification number given in Table II.

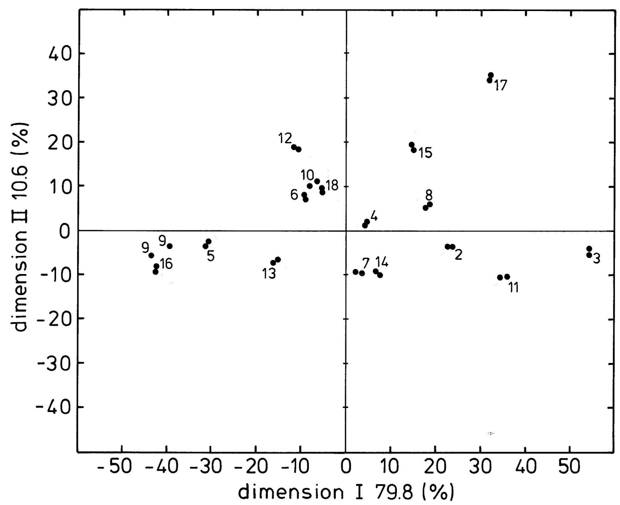

Fig. 3 Two-dimensional subspace for the selected noises of the noise database. The noise spectra are determined twice which results in two data points for each noise and indicates the longterm reproducibility.

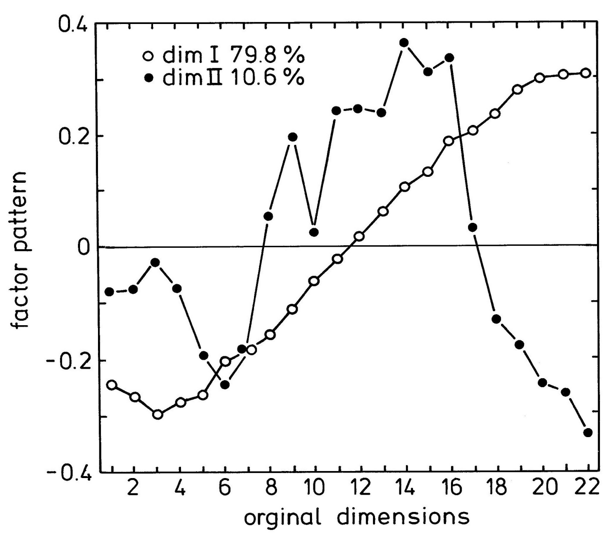

Fig. 4 Factor pattern from the principal component analysis of the noises of Table II for transformation to a two-dimensional subspace.

Fig. 4 gives the factor pattern used for the transformation from the original 22 individual 1/3-octave levels to the two-dimensional subspace. This subspace explains 90.4% of the total variance. In Appendix A a table with the factor pattern is given. With this table new spectra can be transformed to the present subspace of the selected noise data-base. For this transformation the average of the 1/3-octave spectrum should be normalized to 0 dB and then multiplied by the factor patterns of the first and second dimension as given in Appendix A. After the transformation the point of gravity correction (Appendix A) must be added. This procedure is performed for the noise samples number 17 and 18 of Table II which were added later to the data-base.

4 APPLICATION OF THE DATA-BASE

The noise samples of this data-base can be divided into two groups, one group obtained from acoustical noise recording in a number of military environments and a second group obtained from a thermal noise source. The noises of the second group are selected in such a way that the range of different spectra of the noises from the first group is covered.
The long-term stability of the group of thermal noises is higher than for the noises obtained in a practical situation. This is quantified by the differences between L_{90} and L_{10} as given in the specification of the noise spectra in Appendix B. The level fluctuation is also quantified by the standard deviation around the mean of the fifty dB-values of the spectrum measurements. An indication of the long-term stability is given in Fig. 3 where all data points in the reduced spectrum space are given twice, i.e. one data point was obtained from the first minute of the noise recording and the other data point from the end of the recording during the fourth minute.
The noises can be used to simulate an acoustical environment by play-back through a public address system in a diffuse acoustical environment. In this situation a talker or a listener is placed in the same environment. The noise level can either be set according to the practical situation from which the noise was obtained or varied in steps in order to obtain the performance of a system under test as a function of the noise level.
Besides the overall level of the noise the ratio between speech and noise may be of interest. To determine the speech to noise ratio the method described by Steeneken and Houtgast (1986) can be used.
A second application of the noise data-base is to add the noises to pre-recorded speech signals. In this case a correction for the speech microphone response should be made. Normally a spectral shaping is sufficient. In order to predict the signal-to-noise ratio in a practical situation the vocal effort of the talker should be taken into account. The speech

level can be corrected according to the method given by Wheeler and Rawlinson (1985) or Webster (1984).

The data-base is available on digital tape (Sony PCM Betamax and DAT) and on analogue tape (compact cassette). The recording is about one hour long (4 minutes for each noise).

Acknowledgements

Some of these recordings were made by B. Crabtree of DCIEM, Downsview, W. Bowen and K. Smith of RAE, Bedford, and M. Hunt of RAE/NRC, Ottawa.

The manuscript was reviewed by Andrew Varga of RSRE, Malvern, England.

REFERENCES

AC/243 (Panel 3) D/243 (1985). Recommendations of analogue tape recording for data-bases.

Steeneken, H.J.M. and Houtgast, T. (1986). A comparison of some methods for measuring speech levels. Report IZF 1986-20, TNO Institute for Perception, Soesterberg, The Netherlands.

Webster, J.C. (1984). Noise and communication. In: D.M. Jones and A.J. Chapman (Eds.), Noise and Society, John Wiley & Sons.

Wheeler, P.D. and Rawlinson, R.D. (1985). An investigation into the influence of side-tone level on vocal performance. Institute of Sound and Vibration Research, report AC 530A, Southampton, England.